

Novel Organoid Models for the Functional Validation of Pancreatic Cancer Genomic Variants

A thesis submitted for the degree of PhD by:

Shannon Rebecca Nelson, BSc., MSc.

November 2021

This work was carried out under the supervision of

Dr. Naomi Walsh,

Dr. Denise Harold,

National Institute for Cellular Biotechnology,

School of Biotechnology,

Dublin City University

&

Prof. John Crown,

St. Vincent's University Hospital

[This page is intentionally left blank]

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of PhD, is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

*Signed:*_____

*Date:*_____

ID Number: 17210149

[This page is intentionally left blank]

*This thesis is dedicated to my late mother, Mary Nelson.
Pursuing a PhD was a dream my mother always had for me.
I am glad I have had the chance to do this in your memory, Mom.
I hope I've made you proud.*

[This page is intentionally left blank]

Acknowledgements

I am thankful to Assistant Professor Naomi Walsh for giving me the opportunity to complete this PhD. Thank you to Assistant Professor Denise Harold for all of your help and guidance over the years, and to Professor John Crown, for your support throughout the project. Thank you to everyone in the Walsh Cancer Research Group: Fred, Caoimhe, Sara, and Letizia for all the clean room chats, and for being an extra set of hands when I needed them.

During my time in the NICB, I had the chance to work with a number of amazing researchers, many of which I am now lucky to call friends. Thank you everyone in the DCU Cancer Research group – Denis, Alex, AnnMarie, Laura Breen, Alacoque, Nicola, Laura Ivers and Dalal. In particular, I would like to thank Neil, because no matter how busy you were – and you were always incredibly busy – you always made time if I needed help. Throughout the course of my PhD you have become a great friend and mentor. To everyone in open office for making the NICB such a great place to work – Taylor, Ciara, John and Charles – for all the scone Fridays, games nights and Fantasy Football leagues. Thank you to everyone who was there just to have a chat when I needed one – Justine, Joanne, Clair, Giuseppe and Creina. I would also like to thank those whose work behind the scenes meant this research could be done – Donal, Mairead, Gillian and Josephine.

I would like to thank all the members of the DREAM Lab in Aarhus, Denmark – Alun, Lin, Lars, Joanna, Betinna and Hui who welcomed me to the team. Thank you to Trine and Mette for being such a good friends.

To Berta and Ali, thank you both for being so dependable and being there whenever I needed you, whether it was for a coffee, a chat, or just some tough love. Your constant support, encouragement, and motivation helped me through the hardest days PhD, and for that I thank you.

Thank you to my friends and family, including my Dad; my sisters and brother: Sarah, Owen, Sinéad, Kirah, Niamh and Amanda for always reminding me how

proud Mom would be, and my nieces and nephews: Laura, Lilly, Shane, AJ, John (Duder) and Aveah who never failed to put a smile on my face. Thank you to George and Jackie, for all the days out and fun! Thank you to Karen, for being there when I need someone to listen, and for always giving me great advice. Thank you to Seán for taking all of my two-hour long calls, even though you hate being on the phone. And a huge thank you to Nana Phil, after a long and stressful day you were always there to give me never ending encouragement.

David, I would not have been able to do this without you. Thank you for being there for me through every high and every low of the PhD. For always knowing the difference between when I really needed your encouragement for that final push to get something over the line, and when I needed to take a break and forget about my work for a few hours and making me take that time off. Thank you for being my rock. I am so lucky to have had your love and support during my PhD, and it will take a long time for me to repay you for everything you have done for me.

Table of Contents

Chapter 1. Introduction	1
1.1. Cancer overview	1
1.2. Pancreatic cancer overview	1
1.2.1. The pancreas	1
1.2.2. Pancreatic cancer development	3
1.2.3. The cell cycle	4
1.3. Biomarkers	5
1.4. Treatments	5
1.4.1. Resection	5
1.4.2. Chemotherapy	6
1.5. Genomic variants of PDAC	8
1.5.1. Somatic mutations	8
1.5.2. Genomic subclasses of PDAC	14
1.5.3. Predisposition gene mutations	18
1.5.4. GWAS risk loci	22
1.5.5. GWAS pathway analysis	24
1.6. Diabetes	27
1.6.1. MODY	27
1.6.2. The role of the <i>HNF1</i> family in PDAC	32
1.7. Role of Apoptosis in DNA repair pathway	35
1.8. Models of PDAC research	37
1.8.1. Established PDAC cell line cultures	37
1.8.2. Primary tumour cultures	39
1.8.3. Patient derived xenografts	39
1.8.4. Organoids	40
1.9. Methods used in assessing functionality of genes in PDAC	49
1.9.1. CRISPR	49
1.9.2. CUT&RUN – Cleavage Under Targets and Release Using Nuclease	51
1.9.3. Dual luciferase reporter assays	55
1.10. Aims	57
Chapter 2. Methods	59
2.1. Cell culture	59
2.1.1. Cell lines	59
2.1.2. Cell lines and culture conditions	59
2.2. Wnt3a conditioned media	62

2.2.1.	Conditioned media collection	62
2.2.2.	TOPFlash assay	62
2.3.	PDAC organoid culture	63
2.3.1.	Organoids & CLOs	63
2.3.2.	Media and plate preparation	65
2.3.3.	PDAC PDX organoid culture establishment	68
2.3.4.	Establishment of isogenic matched 2D primary cell line	72
2.3.5.	CLO establishment	72
2.3.6.	Species confirmation PCR	73
2.4.	Mutation analysis	74
2.5.	Proliferation assays	76
2.5.1.	Cell line proliferation assay	76
2.5.2.	Organoid and CLO cell proliferation assay	76
2.6.	Toxicity assays	76
2.6.1.	Cell line drug toxicity assays	76
2.6.2.	Organoid and CLO drug toxicity assay	76
2.6.3.	Apoptosis assay	77
2.7.	Staining and imaging	77
2.7.1.	Brightfield imaging	77
2.7.2.	Haematoxylin and eosin stain	77
2.7.3.	Immunofluorescence	78
2.7.4.	Immunohistochemistry	80
2.8.	Western blotting	83
2.9.	RNA extraction	84
2.9.1.	cDNA synthesis	84
2.9.2.	Quantitative reverse transcription PCR (RT-qPCR)	85
2.9.3.	RNA-sequencing	86
2.10.	<i>In vivo</i> xenograft study	88
2.10.1.	Tumour induction via subcutaneous implant	88
2.10.2.	Tumour measurements	89
2.10.3.	Tumour retrieval and processing	89
2.10.4.	Sectioning	89
2.11.	CUT&RUN	89
2.11.1.	CUT&RUN assay	89
2.11.2.	Sequencing	90
2.12.	SNP selection	91
2.12.1.	Literature search	91
2.12.2.	Identification of proxy variants	91
2.12.3.	<i>In silico</i> analysis of SNPs	91
2.12.4.	Prioritisation of SNPs	92

2.12.5.	Position weight matrix (PWM)	93
2.13.	Dual luciferase reporter assays	94
2.13.1.	SNP cloning	94
2.13.2.	Restriction enzyme digestion	96
2.13.3.	Ligation	97
2.13.4.	Transformation	98
2.13.5.	PCR screening	98
2.13.6.	Site directed mutagenesis	98
2.13.7.	Dual luciferase reporter assay	100
2.14.	CRISPR	101
2.14.1.	Golden Gate Assembly	101
2.14.2.	Lentivirus protocol	104
2.14.3.	Transduction	105
Chapter 3.	Establishment and characterisation of novel <i>in vitro</i> PDAC models – PDX derived organoids and cell line organoids	107
3.1.	Introduction	107
3.1.1.	Aims	108
3.2.	Results	109
3.2.1.	Conditioned media	109
3.2.2.	Development of PDX organoids and isogenic matched primary cell lines	111
3.2.3.	Confirmation of primary cell lines and organoid culture derived from human PDAC PDX tumour cells	113
3.2.4.	Generation of CLOs as novel models to study PDAC	114
3.2.5.	Characterisation of novel models	115
3.2.6.	Proliferation assays	119
3.2.7.	Gemcitabine response assays	124
3.2.8.	Stem cell marker expression	128
3.2.9.	Immunofluorescent analysis of PDAC, stem cell and cancer stem cell markers	131
3.2.10.	Transcriptomics	140
3.2.11.	Subtyping of PT291 cancer models	148
3.2.12.	Xenograft <i>in vivo</i> study	150
3.3.	Discussion	162
Chapter 4.	Experimental and <i>in silico</i> assessment of the role of MODY pathway genes in PDAC	169
4.1.	Introduction	169
4.1.1.	Aims	170
4.2.	Results	171

4.2.1.	Functional analysis of SNPs and genes associated with the MODY pathway identified through GWAS pathway analysis	171
4.2.2.	Role of genes in the MODY pathway in the development and progression of PDAC	205
4.2.3.	Optimisation of use of lentiviral CRISPR to generate knockout organoids	232
4.2.4.	CUT&RUN of <i>HNFLA</i> and <i>HNFB</i> in CLO models	234
4.3.	Discussion	239
Chapter 5. Assessing the role of the ATM/DNA repair pathway on drug treatment response		243
5.1.	Introduction	243
5.1.1.	Aims	245
5.2.	Results	246
5.2.1.	Functional analysis of SNPs and genes associated with ATM and DDR pathways identified through GWAS pathway analysis	246
5.2.2.	Role of CASP7 in PDAC	273
5.2.3.	Targeting the DDR pathway in non- <i>BRCA</i> mutant PDAC cells using a novel compound capable of mimicking the effect of <i>BRCA2</i> mutation	288
5.3.	Discussion	307
Chapter 6. Summary and future work		315
6.1.	Summary	315
6.2.	Conclusions	316
6.3.	Future work	316
Chapter 7. References		319
Chapter 8. Appendix		361
8.1.	Appendix A	361
8.2.	Appendix B	363
8.3.	Appendix C	367
8.4.	Appendix D	368

Abbreviations

2D	Two-dimensional
3D	Three-dimensional
ADEX	Aberrantly differentiated endocrine exocrine
ADM	Acinar-to-ductal metaplasia
ARTP	Adaptive rank truncated product
BCA	Bicinchoninic acid
BP	Base pairs
CAF	Cancer-associated fibroblast
CAS	CRISPR-associated
CASPASES	Cystine-aspartate proteases
CEA	Carcinoembryonic antigen
ChEC	Chromatin endogenous cleavage
CHFM	Complete human feeding media
ChIC	Chromatin immuno-cleavage
ChIP-SEQ	Chromatin immunoprecipitation with massively parallel sequencing
CLOs	Cell line organoids
COMPASS	Changes and Characteristics of Genes in Patients with Pancreatic Cancer for Better Treatment Selection
CRISPR	Clustered regularly interspaced short palindromic repeats
crRNA	Small CRISPR-RNA
CT	Computed tomography
CUT&RUN	Cleavage under targets and release using nuclease
DamID	DNA adenine methyltransferase identification
DDR	DNA damage response
DMEM	Dulbecco's Modified Eagle Medium
DSB	Double-stranded break
ECM	Extracellular matrix
eQTL	Expression quantitative trait loci
FAMMM	Familial atypical multiple mole and melanoma syndrome
FAP	Familial adenomatous polyposis

FFPE	Formalin fixed paraffin embedded
FOLFIRINOX	Irinotecan, oxaliplatin, fluorouracil, and leucovorin
gBRCAm	Germline <i>BRCA</i> mutated
GEMM	Genetically engineered mouse model
GEPIA2	Gene expression profiling interactive analysis 2
GTEX	Genotype-tissue expression
GWAS	Genome wide association studies
HBOC	Hereditary Breast and Ovarian Cancer
HDR	Homology directed repair
HRR	Homologous recombination repair
HTS	High throughput screening
IGV	Integrative genomics viewer
IHC	Immunohistochemistry
INDEL	Insertion or deletion
LD	Linkage disequilibrium
LTG	Laboratory of Translational Genomics
MMR	Mismatch repair genes
MODY	Maturity onset diabetes of the young
MRI	Magnetic resonance imaging
MSI-H	Microsatellite instability-high
NCCN	National Comprehensive Cancer Network
NCI	National Cancer Institute
NES	Normalised effect size
NGS	Next-generation sequencing
NHEJ	Non-homologous end joining
NHGRI-EBI	National Human Genome Research Institute-European Bioinformatics Institute
OCT	Optimal cutting temperature
OR	Odds ratio
OS	Overall survival
pA-MNase	Protein-A micrococcal nuclease
p-adj	Adjusted p-value

PAM	Protospacer adjacent motif
PANC4	Pancreatic Cancer Case-Control Consortium
PanIN	Pancreatic intra-epithelial neoplasia
PANSCAN	Pancreatic Cancer Cohort Consortium
PBS	Phosphate buffered saline
PCC	Pearson correlation coefficient
PDAC	Pancreatic ductal adenocarcinoma
PDX	Patient derived xenograft
PJS	Peutz-Jegher syndrome
POLO	Pancreas cancer OLaparib Ongoing
PolyHEMA	Poly-2-hydroxyethyl methacrylate
PP	Pancreatic polypeptide
PWM	Position weight matrixes
RT-qPCR	Reverse transcription quantitative PCR
RCT	Randomised control trials
ROCKi	Rho kinase inhibitor
ROS	Reactive oxygen species
RPMI	Roswell Park Memorial Institute-1640 medium
SCID	Severe combined immune deficiency
SFM	Serum free media
sgRNA	Single guide RNA
SNPS	Single nucleotide polymorphisms
STR	Short tandem repeat
T1DM	Type 1 diabetes mellitus
T2DM	Type 2 diabetes mellitus
TALENS	Transcription activator like effector nucleases
TF	Transcription factor
TIL	Tumour infiltrating lymphocyte
TPM	Transcripts per million
UCSC	University of California Santa Cruz
VNTR	Variable number of tandem repeat
ZFNs	Zinc finger nucleases

[This page is intentionally left blank]

Research Outputs

Publications

Nelson, S.R.; Zhang, C.; Roche, S.; O'Neill, F.; Swan, N.; Luo, Y.; Larkin, A.M.; Crown, J.; Walsh, N. Modelling of pancreatic cancer biology: Transcriptomic signature for 3D PDX-derived organoids and primary cell line organoid development. *Sci. Rep.* **2020**, *10*, 1–12.

Nelson, S.R.; Walsh, N. Genetic Alterations Featuring Biological Models to Tailor Clinical Management of Pancreatic Cancer Patients. *Cancers* **2020**, *12*, 1233.

Nelson, S.R.; Roche, S.; Cotter, M.; Garcia, P.A.; Reitmeier, D.; Zollbrecht, E.; O'Neill, F.; Clynes, M.; Doolan, P.; Medha, J.P.; Swan, N.; Larkin, A.M.; Walsh, N. Genomic Profiling and Functional Analysis of let-7c miRNA-mRNA Interactions Identify SOX13 to Be Involved in Invasion and Progression of Pancreatic Cancer. *J. Oncol.* **2020**

Invited talks

EACR Travel Fellowship from Ireland to Denmark, IACR/EACR Early Career Research Symposium and Workshop – Motility Experience Session, Irish Association for Cancer Research, Galway, Ireland 2020

Oral presentations

Turning up the volume on a silent killer – the development of novel models of pancreatic cancer, The Patrick G. Johnston Award for Excellence in Cancer Research Outreach, Irish Association for Cancer Research, Galway, Ireland 2020

Primary cell line organoids (CLOs) as a novel and alternative 3D organotypic cell model for pancreatic cancer, Breakthrough Cancer Research Session, Irish Association for Cancer Research, Galway, Ireland 2020

Poster presentations

Nelson, S.R.; Andrieu, C.; Crown J.; Harold, D.; Walsh, N. Analysis of the role of maturity onset diabetes of the young (MODY) pathway transcription factors in pancreatic cancer. *Irish Society of Human Genetics Livestreamed Annual Scientific Meeting, 2020*.

Nelson, S.R.; Roche, S.; O'Neill, F.; Larkin, A.; Crown, J.; Walsh N. Assessment of pancreatic cancer 3D organotypic cultures as models for functional preclinical analysis.

Conference Proceedings: AACR Annual Meeting 2019, Atlanta, GA, USA 2019. DOI:10.1158/1538-7445.AM2019-44.

Nelson, S.R.; Roche, S.; O'Neill F.; Larkin A.; Crown J.; Walsh N. Assessment of expression of tumour-initiating cell markers in pancreatic ductal adenocarcinoma organoids and isogenetically matched cell lines. *Annual meeting for the Irish Association for Cancer Research, Belfast, Northern Ireland, 2019.*

Nelson, S.R.; Reitmeier, D.; Anton-Garcia, P.; O'Neill, F.; Roche, S.; Larkin, A.; Walsh N. Genomic profiling and functional analysis of miRNA let7-c:SOX13 mRNA interactions involved in invasive progression of pancreatic cancer. *Annual meeting for the Irish Association for Cancer Research, Dublin, Ireland, 2018.*

Nelson, S.R.; Roche S.; O'Neill F.; Larkin A.; Crown J.; Walsh N. Optimisation of protocols for the establishment of pancreatic cancer organoids and cell lines from patient derived xenografts (PDX) tumour models *30th EORTC-NCI-AACR SYMPOSIUM, Dublin, Ireland 2018.*

Awards

DCU Faculty of Science and Health Outstanding Graduate Researcher of the Year 2021;
Professor Patrick Johnston IACR award for excellence in cancer research outreach IACR 2020;

IACR AOIFA 2020 Conference Award for €1,000 to attend the EACR conference.

Funding

EACR Travel Fellowship Grant: €2,500. Developing methods for the use of CRISPR in the genome editing of pancreatic cancer organoids, Aarhus University, Denmark

AOIFA-IACR Seed-Funding Award: €6,000. Assessing the role of transcription factors *HNFI*A and *HNFI*B in the development and progression of pancreatic cancer

Novel organoid models for the functional validation of pancreatic cancer genomic variants

Shannon Rebecca Nelson

Abstract

With a five-year survival rate of 9%, pancreatic ductal adenocarcinoma (PDAC) has the one of the worst prognoses of all cancers. Some limitations in the understanding of the disease are due to the lack of representative *in vitro* patient tumour models. In order to overcome this unmet preclinical need, this thesis outlines the establishment of a method for the development of organoid and isogenic matched primary cancer cell line models from patient derived xenograft (PDX) tumours. In order to create a patient-reflective yet versatile *in vitro* model, the matched primary cell line was developed further and subsequently generated organoids, termed cell line organoids (CLOs). These CLOs represent the phenotypic and transcriptomic profile of the original organoids and PDX tumour.

Recent genome wide association studies (GWAS) and pathway analyses have implicated genes and single nucleotide polymorphisms (SNPs) from the maturity onset diabetes of the young (MODY) gene set and the Pujana *ATM* Pearson correlation coefficient (PCC) network in the development of PDAC. The biological functionality of the genomic variants identified from the GWAS-enriched pathways were assessed using *in silico* methods and experimental dual luciferase reporter assays. Genes in the MODY pathway such as hepatocyte nuclear factor-1 alpha/beta (*HNFI*A and *HNFI*B) act as transcription factors. Their role in cancer progression was assessed through single and double CRISPR knockouts in PDAC primary cell cultures. CUT&RUN (cleavage under targets and release using nuclease) was performed to identify genes regulated by *HNFI*A and *HNFI*B TFs. Additionally, targeting the DNA damage response (DDR) pathway in non-*BRCA* mutated PDAC was assessed using a novel drug which mimics the effect of a *BRCA*2 mutation.

In conclusion, this thesis shows the development of novel, adaptable organoid models for functional validation of genomic variants in PDAC. Furthermore, biological investigation of GWAS pathway identified SNPs and genes from the MODY and Pujana *ATM* PCC pathways highlights the powerful nature of these tools in identifying genomic variants associated with PDAC. It also highlights the importance of functional experimental analysis to provide better understanding of their role in the development and progression of PDAC.

[This page is intentionally left blank]

Chapter 1. Introduction

1.1. Cancer overview

Cancer is a leading cause of death in every country in the world [1]. Overall, the rate of cancer incidence and mortality is growing worldwide, this is due to both an ageing population and changes in the prevalence of cancer risk factors [2]. In 2020, there were an estimated 19.3 million new cases, and 10 million cancer deaths, with lung cancer the leading cause of cancer related death [3]. Due to reductions in smoking, and improvements in diagnosis and treatment, the rate of cancer death has declined continuously from its peak in 1991 through 2018 [4]. While ongoing research has resulted in drastically decreased mortality rates for number of cancers such as breast, prostate and colorectal, improvements in diagnosis, treatment and survival are required for other cancers such as lung and pancreatic cancer [4].

1.2. Pancreatic cancer overview

With a five-year survival rate of 9%, pancreatic cancer has one of the worst outcomes of all cancers. Due to its rapid progression and fatal outcome, long term survivors are limited to those with resected early-stage tumours [5,6]. As approximately 80% of pancreatic cancer patients are diagnosed after the disease has metastasised, most people diagnosed are ineligible for resection, the only curative treatment. It is the third leading cause of cancer related death in the Western world, and by 2030 it is estimated that pancreatic cancer will surpass colorectal cancer to become the second most fatal cancer in the United States [4]. The most common form of pancreatic cancer is pancreatic ductal adenocarcinoma (PDAC), which occurs in the exocrine pancreas, with the remaining 5% of cases in the endocrine pancreas [7]. Epidemiological factors, including smoking, obesity, type II diabetes mellitus and acute pancreatitis account for approximately 25% of cases of PDAC [8–11].

1.2.1. The pancreas

The pancreas is a complex organ which is essential for life. It is located behind the stomach and it plays a central role in blood sugar regulation and digestion by

production of hormones and enzymes. It is comprised of two components, the endocrine and exocrine portions, which are both morphologically and physiologically distinct (Figure 1.1) [12]. The exocrine portion of the pancreas comprises 95% of the organ and is made of acinar cells which mainly produce digestive enzymes, including proteases, amylases, and lipases. These enzymes are delivered to the duodenum via ductal cells [13]. The endocrine pancreas is made up of highly organised and vascularised micro-organs called the Islets of Langerhans. These islets consist of specialist endocrine cells (α , β , δ , and pancreatic polypeptide (PP) cells), which all produce specific hormones. β -cells are the most abundant and secrete insulin; α , δ , and PP produce glucagon, somatostatin, and pancreatic polypeptide, respectively [12].

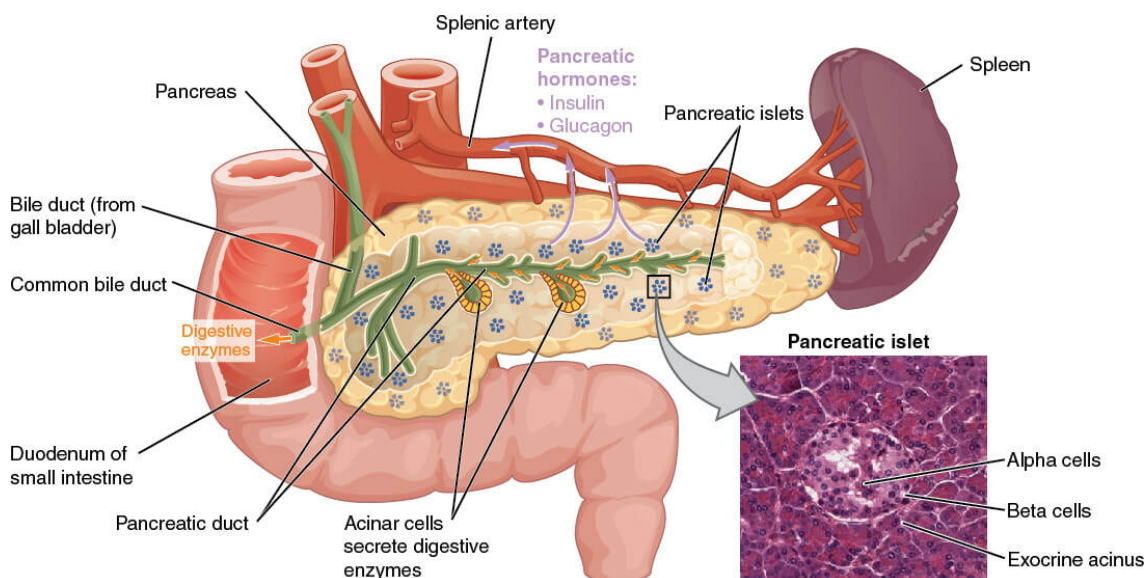


Figure 1.1: Diagram of the pancreas and surrounding organs. The micrograph reveals pancreatic islets. Taken without permission from University of Michigan Medical School © 2012 [14].

1.2.2. Pancreatic cancer development

Pancreatic cancer develops when the enzyme-producing acinar cells transdifferentiate into ductal-like cells in response to environmental conditions, including tissue damage, inflammatory signals and stress. This process is called acinar-to-ductal metaplasia (ADM) [15]. During ADM, cells become more susceptible to pro-oncogenic events, such as activating mutations in the proto-oncogene *KRAS*, resulting in the cells transforming into the precursor lesions of PDAC – pancreatic intra-epithelial neoplasia (PanIN). This transformation is considered to be the initiating genetic event for the development of PDAC, and is followed by a sequential progression, involving the loss or deactivation of several tumour suppressor genes including *CDKN2A*, *TP53* and *SMAD4* (Figure 1.2) [16].

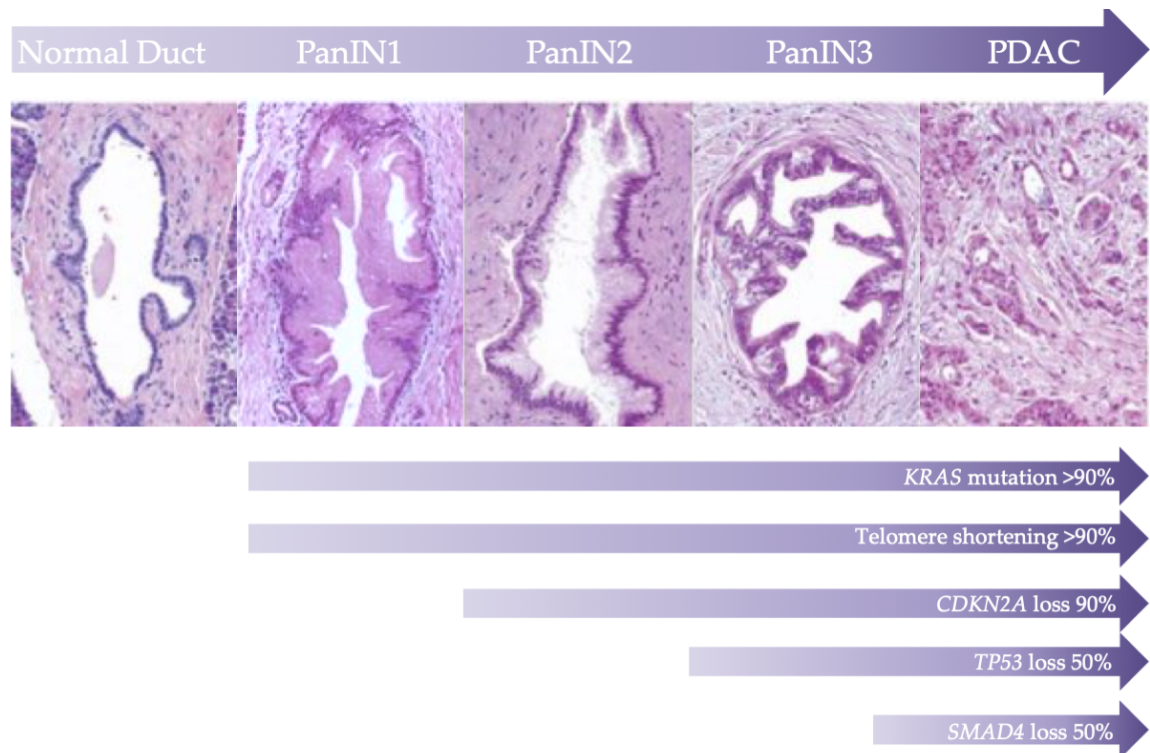


Figure 1.2: Illustration of the progression and development of PDAC, starting with normal ducts from PanIN stages to advanced carcinoma. Arrows below indicate genomic events occurring as the disease progresses, the percentage of patients with the mutation, and when these events typically begin. Haematoxylin and eosin stained tissue sections images modified without permission from Hezel et al. [17].

1.2.3. The cell cycle

The cell cycle is a process in which regulatory proteins direct the cell through a specific sequencing of events, resulting in mitosis and the production of two daughter cells [18]. Cyclin-dependent kinases (CDKs) form complexes with cyclin proteins and regulate the cell's progression through the four-stages of the cell cycle – growth phase (G1), DNA synthesis (S), preparation to divide (G2) and mitosis (M) (Figure 1.3) [19]. The cyclin-CDK complexes interact with downstream targets including Rb and E2F and are regulated by a number of proteins including TP53, P21, P16 and CDC25. The cell cycle is often dysregulated in cancer due to mutations in oncogenes or tumour suppressors which impact cell cycle regulation [18].

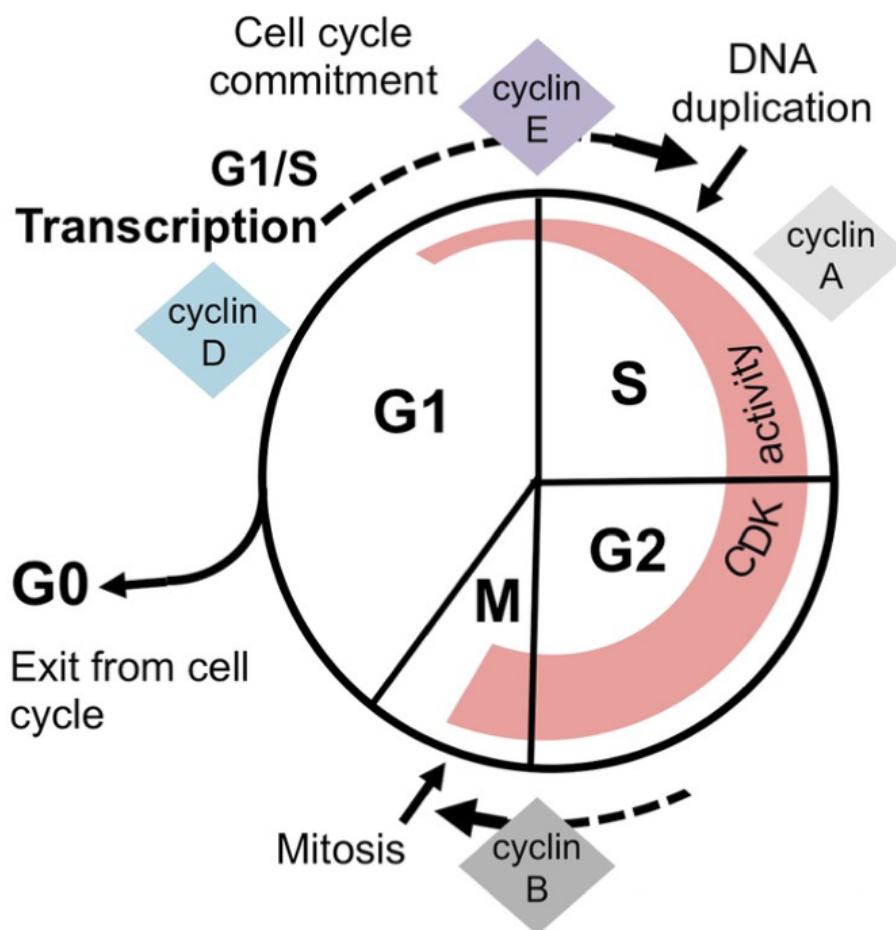


Figure 1.3: The four-stages of the cell cycle – growth phase (G1), DNA synthesis (S), preparation to divide (G2) and mitosis (M) – are driven by cyclin-dependent kinases (CDKs)-cyclin complexes. Taken without permission from Bertoli and de Bruin [19].

1.3. Biomarkers

There is currently an unmet clinical need for biomarkers for the early detection of PDAC. The only biomarker recommended by the National Comprehensive Cancer Network (NCCN) guidelines is CA19-9, a sialylated Lewis antigen present in glycosphingolipids and glycoproteins. However, it has a number of limitations. These include: a low positive predictive value, with false positives in patients with obstructive jaundice, and false negatives in patients with a Lewis blood-type negative phenotype [20]. It may be positive in patients with non-malignant disease, such as pancreatitis. Also, only 65% of patients with PDAC have increased levels of CA19-9 [21]. Carcinoembryonic antigen (CEA) is also used clinically for diagnosing PDAC. However, it also has a low positive predictive value, with a sensitivity and specificity of 45% and 75% respectively [22].

A number of researchers are working to improve the predictive value of these biomarkers. Kim *et al.* [23] have validated the use of the novel biomarker thrombospondin-2 (THBS2) in combination with CA19-9 using an ELISA to distinguish PDAC from pancreatitis, with a specificity of 98% and a sensitivity of 87%. This biomarker combination can be used for the identification of patients with early-stage, resectable PDAC. Jenkinson *et al.* [24] identified that thrombospondin-1 (THSBI) can be used as biomarker of PDAC. Their study found there was a significant decrease in THSBI levels up to 24 months prior to diagnosis of PDAC. They also found a link between reduced serum THSBI levels, and PDAC-associated diabetes mellitus.

1.4. Treatments

1.4.1. Resection

The only curative treatment for PDAC is resection with adjuvant chemotherapy. However, most pancreatic cancer tumours are found at a late, inoperable stage. In order for the tumour to be resectable, it must have no arterial or venous involvement; not be attached to other organs and have no distant metastases. Given these criteria, less than 20% of patients are eligible for resection [25]. The mean survival time of patients who have had a resection and adjuvant

chemotherapy with gemcitabine and capecitabine is 26 months, with a five-year survival rate of 30% [26]. Of the patients who are operated on, 60% relapse within 12 months, most likely due to micro-metastases which were not detected during the diagnostic computed tomography (CT) scan [27]. When tumours are less than <2 cm in diameter the five-year survival rate is 50% and almost 100% when the tumour size <1 cm [28,29].

1.4.2. Chemotherapy

Only 25-30% of PDAC patients treated with chemotherapeutic drugs respond, however, they do eventually become resistant to treatment due to the development of resistance mechanisms, such as deficiencies in drug uptake, alteration of drug targets, activation of DNA repair pathways, and resistance to apoptosis [30]. Gemcitabine is the mainstay of modern chemotherapy for pancreatic cancer, it acts to disrupt DNA replication, and causes activation of the S-phase checkpoint, resulting in the prevention of DNA synthesis. It is transported into cells by the human equilibrative nucleoside transporter 1 (hENT-1) protein, which allows the passage for pyrimidine nucleosides in and out of the cell [31]. Levels of hENT-1 expression appear to be indicative of survival times; in treatment with gemcitabine, patients with low hENT-1 expression had a median survival of 17.1 months, and patients with high hENT-1 expression had a median survival of 26.2 months [32]. Patients with low levels of h-ENT expression should not be treated with pyrimidine nucleosides such as gemcitabine, and should be treated with 5-FU-based regimens [33].

Epidermal growth factor receptor (*EGFR*) has previously been shown to be over expressed in PDAC, with increased expression correlated to metastasis and poor prognosis [34]. A phase III gemcitabine-based combination study with erlotinib, an EGFR inhibitor, showed small but significant survival rates (6.2 vs 5.9 months) compared to gemcitabine alone [35]. On the basis of this clinical trial, the FDA approved erlotinib in combination with gemcitabine chemotherapy for first-line use in locally advanced and metastatic pancreatic carcinoma in 2005. A gemcitabine-nab-paclitaxel combination was also given FDA approval in 2013;

however, the treatment regime has a high occurrence of adverse events [36,37]. A drug combination study FOLFIRINOX (irinotecan, oxaliplatin, fluorouracil, and leucovorin) showed significant overall survival (11.1 vs. 6.8 months) and one-year survival (48.4% vs. 20%) over gemcitabine alone, however there was also an increased occurrence of adverse effects [38]. FOLFIRINOX was approved by the FDA as a first-line treatment for PDAC in 2011. All FDA approved treatments for the treatment of PDAC are listed in Table 1.1.

Table 1.1: FDA approved PDAC chemotherapies with published survival benefits in palliative and adjuvant settings, year refers to the year of publication of palliative chemotherapy trials for advanced PDAC. Modified from Katayama et al. [39].

Year	Intervention	Palliative Survival (months)	Adjuvant Survival Benefit (months)
1962	Fluorouracil	4.41	11 vs. 20
1997	Gemcitabine	4.41 vs. 5.65	20.2 vs. 22.8
2005	Erlotinib + gemcitabine	5.91 vs. 6.24	No benefit
2011	FOLFIRINOX ¹	6.8 vs. 11.1	35.0 vs. 54.4
2013	Nab-paclitaxel + gemcitabine	6.7 vs. 8.5	36.2 vs. 40.5*
2015	Liposomal irinotecan	4.2 vs. 6.1	—
2019	Olaparib	18.1 vs. 18.9*	—
2020	NALIRIFOX ²	6.1 vs. 4.2	—

¹FOLFIRINOX: Leucovorin Calcium, Fluorouracil, Non-liposomal Irinotecan Hydrochloride, Oxaliplatin; ²NALIRIFOX: Leucovorin Calcium, Fluorouracil, Liposomal Irinotecan Hydrochloride, Oxaliplatin; *denotes data taken from interim analysis of an ongoing clinical trial.

1.5. Genomic variants of PDAC

Research into the genetic landscape of the disease, including familial cancer syndromes, inherited predisposition loci and somatic mutations is vital to identifying those at risk of developing PDAC. Recent genomic analysis has resulted in the division of the disease into different subtypes. Understanding the genetic background of the disease can allow for improved selection of therapies for patients.

1.5.1. Somatic mutations

Studies into the PDAC genome have shown that there are approximately 60 alterations per tumour; most of which are point mutations [40]. Activating mutations of *KRAS* are nearly universal, and inactivation of *TP53*, *SMAD4* and *CDKN2A* occur at rates of >50% (Figure 1.4) [41].

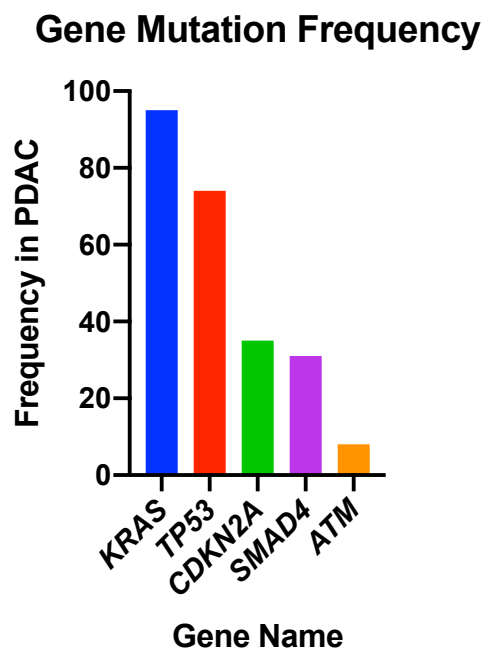


Figure 1.4: Frequency of the most common somatic gene mutations (*KRAS*, *TP53*, *CDKN2A*, *SMAD4* and *ATM*) in PDAC [42,43].

1.5.1.1. *KRAS*

KRAS is a molecular switch, when bound to guanosine-5'-triphosphate (GTP) regulates cell proliferation, differentiation, apoptosis and cell signalling. The *KRAS* mutation is near universal in PDAC, with 94% of tumours possessing the mutation

[43]. Activating point mutations in codon 12, 13, or 16, (most commonly G12D), result in reduced GTP hydrolysis. Cases with *KRAS* mutations at codon 61 give a favourable prognosis, as there is less extracellular-signal-regulated kinase (ERK) activation [44]. The major *KRAS* effector pathways are outlined in Figure 1.5.

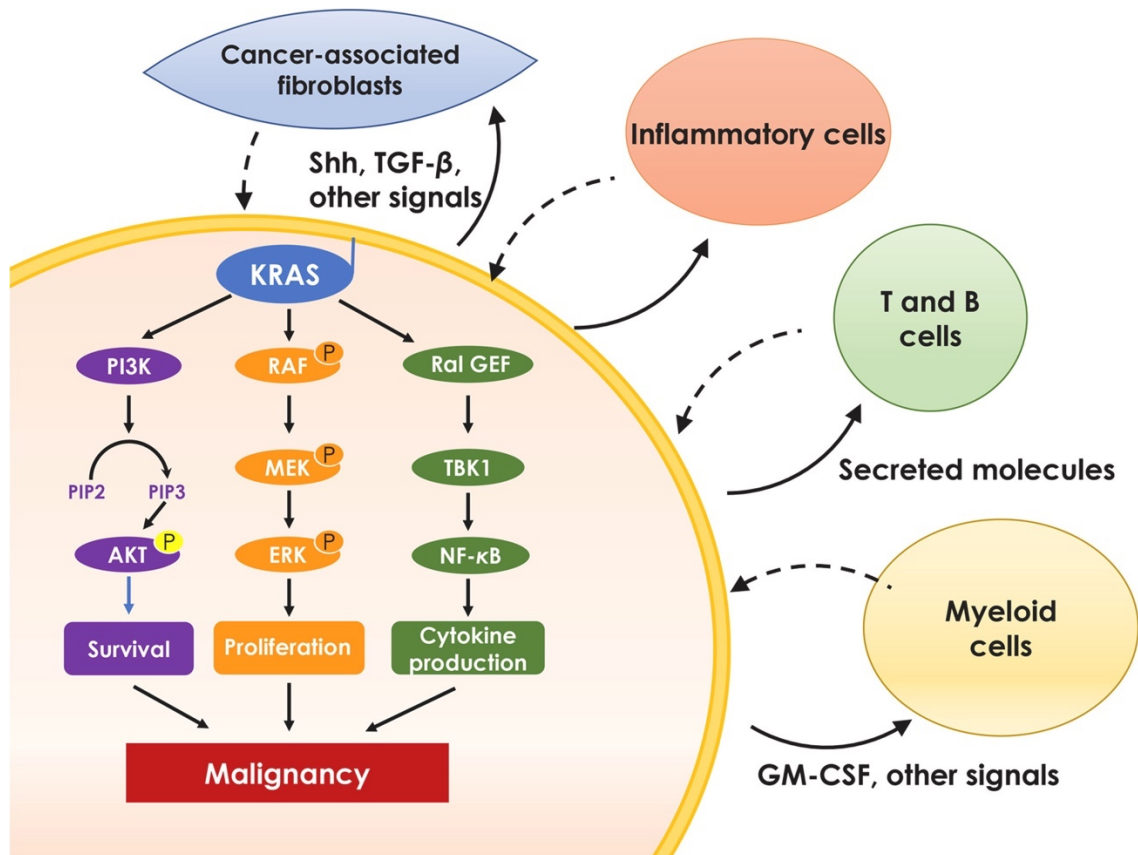


Figure 1.5: The major *KRAS* effector pathways. Oncogenic *KRAS* activates intracellular PI3K, MAPK or RAL-GEF pathways to promote cell survival, proliferation and cytokine secretion. The stromal cells, such as fibroblasts, innate and adaptive immune cells, are also activated in a paracrine manner, and in turn in turn promote cancer malignancy. Taken without permission from Liu et al. [45].

In pancreatic cancer *KRAS* mutations are believed to be the early events in neoplastic transformation [46]. Oncogenic *KRAS* is not sufficient to initiate the carcinogenesis process; this relies on the downstream activation of *Raf-1*, *Rac*, *Rho* or *PI3K* [47]. RAS proteins are modified by farnesyl transferase, an enzyme which adds 15-carbon farnesyl lipid to the carboxyl-terminal cysteine of RAS. This modification is shown to be essential for RAS membrane association and transformation [48]. Clinical trials looking into targeting *KRAS* have shown little

promise. A Phase III trial of R115777, a selective inhibitor of farnesyl transferase in combination with gemcitabine failed to show increased life expectancy in comparison to gemcitabine plus placebo [49]. Patients with *KRAS* mutations have a median survival time of 17 months compared to 30 months for those without *KRAS* mutations [46]. Approximately 3% of PDAC cases are due to microsatellite instability or altered chromosome ploidy [50]. This is usually due to mutations in the mismatch repair genes (MMR) genes *MSH2* and *MSH6*. Typically, *KRAS* is wild type in cancers with these mutations.

1.5.1.2. *CDKN2A*

Cyclin-dependent kinase inhibitor 2A (*CDKN2A*), is inactivated in 95% of PDAC cases, by homozygous deletion, mutation of alleles or promoter hypermethylation resulting in gene silencing [51–53]. *CDKN2A* inhibits cyclin-dependent kinases 4/6 (*CDK4/CDK6*) and thereby activates the retinoblastoma (Rb) family of proteins, which blocks the transition from G1 to S-phase [54], as shown in Figure 1.6. The role of *CDKN2A* is discussed in more detail in Section 1.5.3.3.

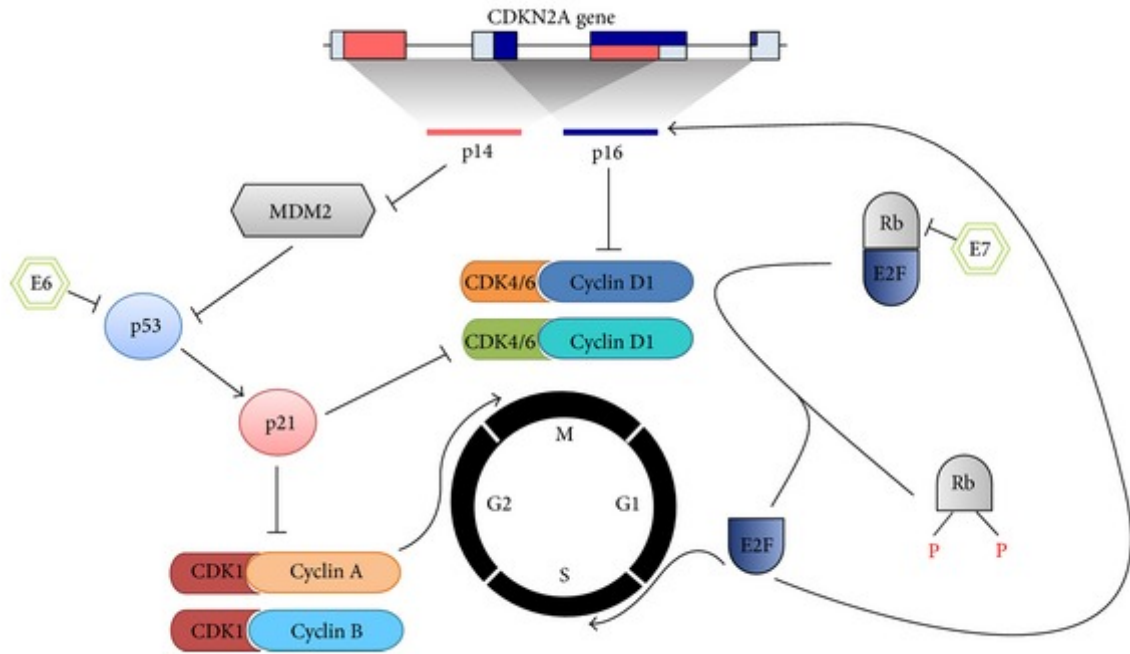


Figure 1.6: Cell cycle arrest by CDKN2A. The CDKN2A gene encodes two alternatively spliced transcripts, p16INK4A and p14ARF. p16INK4A inhibits the CDK4/6-cyclin D1 complexes, preventing phosphorylation (P) of retinoblastoma (Rb) proteins, enabling binding and inactivating the E2F transcription factors, when are required for progression to S-phase and upregulation of p16INK4A. p14ARF activates the tumour suppressor gene TP53 by inhibiting MDM2, which inactivates TP53 by ubiquitin-mediated degradation. Active TP53 induces the expression of p21, a negative cell cycle regulator which is an inhibitor of the CDK1-cyclin A/B complexes, thereby preventing the progression from G2 phase to metaphase. Taken without permission from Al-Kaabi et al. [55].

1.5.1.3. TP53

TP53, known as the guardian of the genome, is the most frequently mutated gene in cancer [56]. It acts as a tumour suppressor, and has roles in apoptosis, genomic stability, inhibition of angiogenesis and arrest of cell growth [57]. It is mutated in 75% of PDAC tumours, mainly by point mutations [58]. TP53 controls cell cycle at the G1/S interface and plays a vital role in inducing programmed death in response to DNA damage [5]. The major TP53 effector pathways are outlined in Figure 1.7.

p53 Signaling

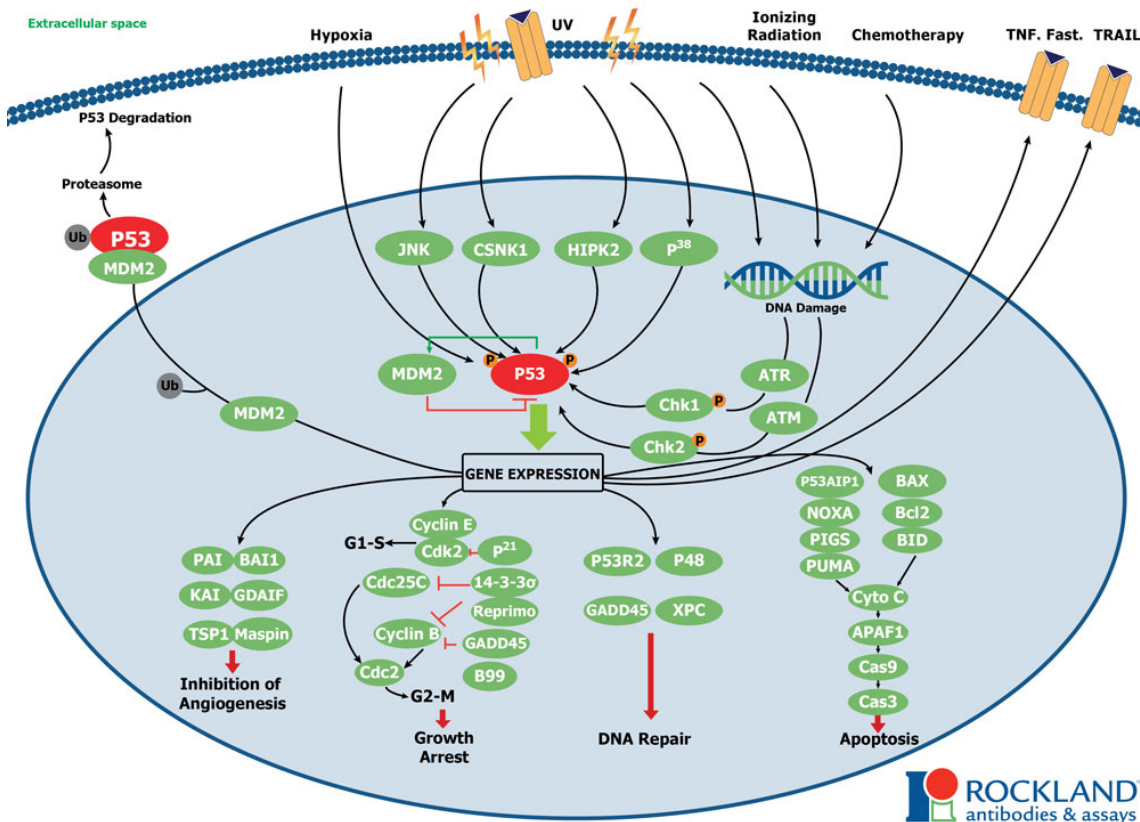


Figure 1.7: TP53 is inactivated by MDM2 by ubiquitin-mediated degradation, this is inhibited by stress-induced acetylation and phosphorylation (P) of MDM2 by the kinases ATM and c-Abl. Activated TP53 will induce a cell cycle arrest to allow either repair and survival of the cell or apoptosis. Activated TP53 binds DNA and activates expression of several genes including WAF1/CIP1 encoding for p21. p21 binds to the G1-S/CDK (CDK2) and S/CDK complexes inhibiting their activity. When p21 forms a complex with CDK2 the cell cannot continue to the next stage of cell division.

Taken without permission from Rockland Immunochemicals Inc. [59].

A study by Weissmueller *et al.* [60] found mutant TP53 induced platelet-derived growth factor receptor b (*PDGFRb*). Knockdown of *PDGFRb* in PDAC cell lines resulted in reduced invasion of the cells. Mutant TP53 was also found to inhibit p73, which represses *PDGFRb*. The study found that increased expression of *PDGFRb* in PDAC patient samples correlates with a worse outcome for patients.

1.5.1.4. SMAD4

SMAD4 is a transcription factor in the transforming growth factor beta (TGFβ) signalling pathway and is inactivated in 50% of advanced pancreatic cancers. It

acts with TGF β 1 as a tumour suppressor to regulate pancreatic cell cycle arrest and apoptosis, mediated by targets such as p21, which causes G1 cell cycle arrest [61]. PDAC patients with biallelic deletion of *SMAD4* more frequently had metastasis than those with wild type *SMAD4* [62]. The major SMAD4 effector pathways are outlined in Figure 1.8.

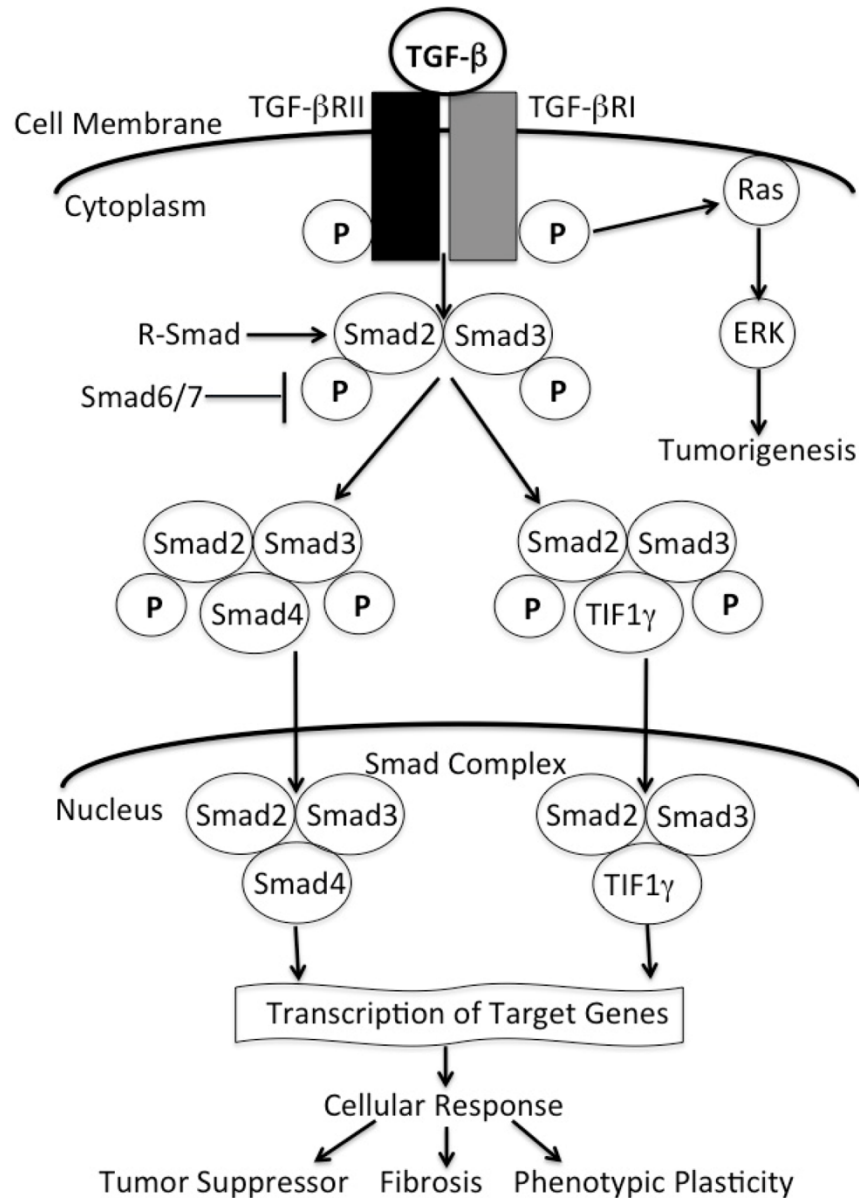


Figure 1.8: TGF- β binds a complex of transmembrane receptor serine/threonine kinases in the cell surface and induces transphosphorylation (P) of the receptors. SMAD2/3 are phosphorylated and form a complex with a SMAD4. Activated SMAD complexes translocate into the nucleus, where they regulate transcription of target genes, through physical interaction and functional cooperation with DNA-binding transcription factors. Taken without permission from Ahemd et al. [63].

1.5.1.5. ATM

Whole genome sequencing of PDAC tumours identified other genes which are frequently mutated, such as ataxia telangiectasia mutated (ATM), a serine/threonine kinase with a role in DNA double-stranded break repair [58]. Recent data based on large-scale sequencing studies reported up to 8% of ATM mutations in PDAC cohorts [43].

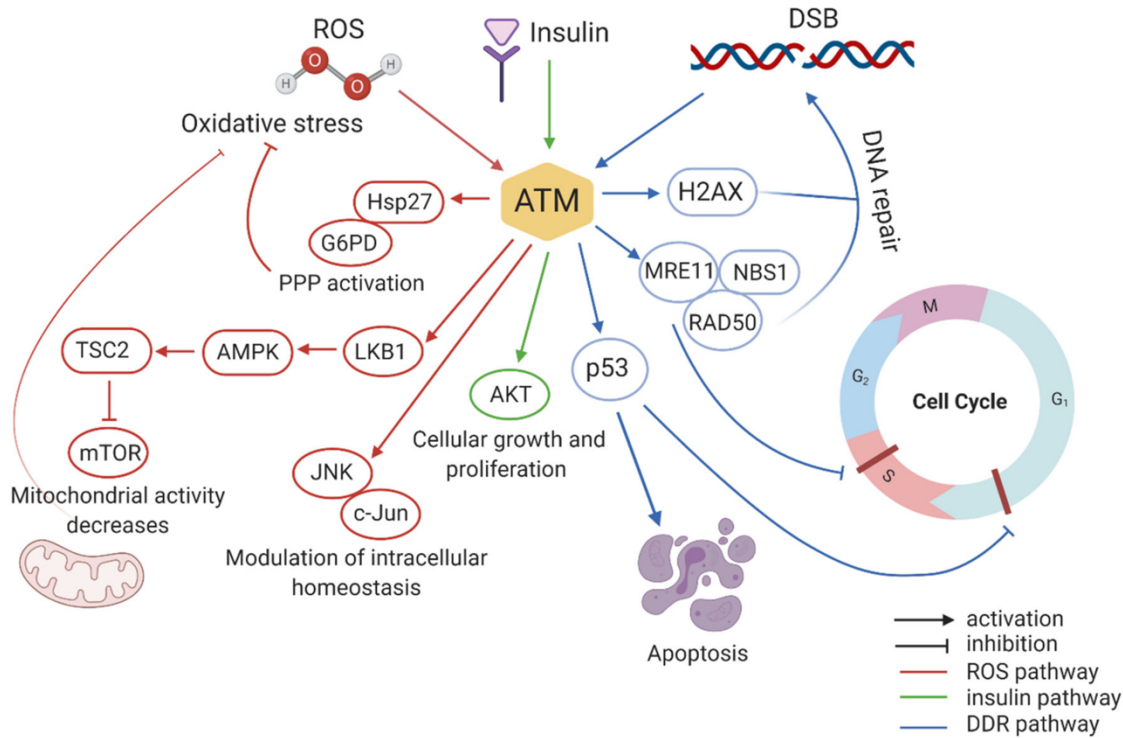


Figure 1.9: A schematic of the different ataxia telangiectasia mutated (ATM) signalling pathway to coordinate cell cycle arrest, DNA repair or apoptosis in response to stress. Once activated, ATM phosphorylated downstream DNA damage repair proteins, H2AX to initiate DNA repair; MRE11 (meiotic recombination II) and TP53 to halt the cell cycle; AKT (protein kinase B) to activate cellular growth and proliferation; JNK (c-JUN N-terminal kinase family) contributing to oxidative stress; mTOR pathway which decreases mitochondrial activity, and the Hsp27 (heat shock protein 27) which acts as an antioxidant. Taken without permission from Pizzamigilo et al. [64].

1.5.2. Genomic subclasses of PDAC

1.5.2.1. Structural variation

Whole genome sequencing of PDAC by Waddell *et al.* [42] resulted in the reclassification of the disease into four different subtypes dependent on the

structural variation of the genome. These are represented as: stable (<50 structural variants) – 20% of cancers; scattered (5-200 structural variants) – 36% of cancers; locally rearranged (>200 structural variants clustered on < 3 chromosomes) – 14% of cancers; or unstable genomes (>200 structural variants distributed across the genome) – 30% of cancers (Figure 1.10). The unstable genome was found to have defects in the DNA damage response (DDR) genes and associated with a response to platinum-based therapy.

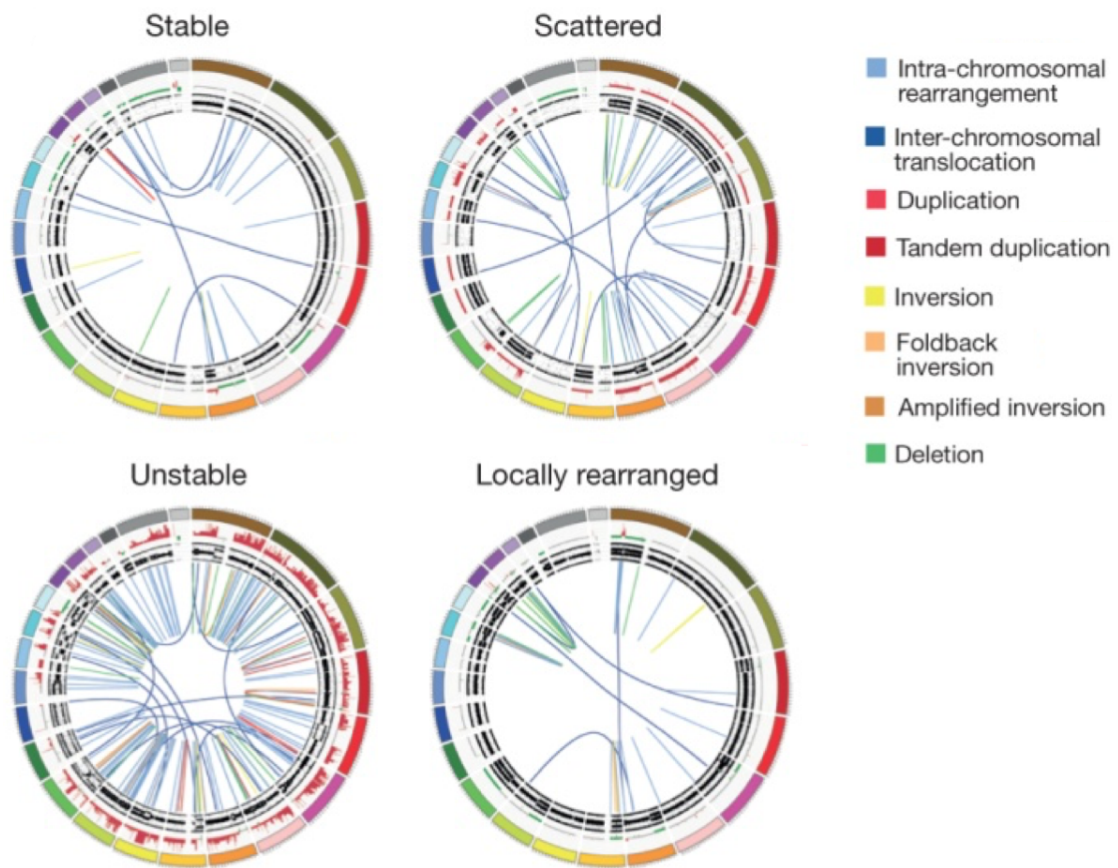


Figure 1.10: Whole genome characterisation of PDAC, based on the number of chromosomal variants. The outer ring represents chromosomes, followed by copy number changes (red = gain, green = loss), and then allele frequency. The lines represent structural rearrangements within the chromosome. Modified without permission from Waddell et al. [42].

1.5.2.2. Molecular subtyping

Moffitt *et al.* [65] classified PDAC based on mRNA microarray and RNA-seq data from 206 patients and identified the subtypes epithelial – further divided into

basal-like and classical, and stromal – further divided into activated and normal (Figure 1.11). The study also found that metastases had the same subtype as the original tumour. These subtypes were also used to establish the prognosis of the tumour, with poor survival in the basal-like and active stroma subtypes, and patients with basal-like tumours benefit from adjuvant chemotherapy.

Collisson *et al.* [66] classified PDAC using 85 primary, untreated tumours and identified the subtypes classical, quasi-mesenchymal and exocrine-like (Figure 1.11). The classic subtype expressed *GATA6* and had a *KRAS* addiction. *KRAS* addiction describes the process whereby tumours are dependent on *KRAS* activity, despite the development of multiple oncogenic mutations [67]. Patients with the quasi-mesenchymal subtype had the poorest prognosis, however, were found to be more sensitive to treatment with gemcitabine. Patients with a classical subtype had a better prognosis and were more sensitive to erlotinib.

Molecular subtyping was also performed by Bailey *et al.* [68] using 266 untreated PDAC tumours. The subtypes identified were squamous, immunogenic, pancreatic progenitor and aberrantly differentiated endocrine exocrine (ADEX) (Figure 1.11). An enrichment of immune signalling was observed in the squamous and immunogenic subtypes, and patients with the squamous subtype had the poorest prognosis.

The clinical trial: Changes and Characteristics of Genes in Patients with Pancreatic Cancer for Better Treatment Selection (COMPASS) (NCT02750657), aimed to find better ways of understanding and treating PDAC using cancer subtypes [69]. Patients were biopsied, and RNA-sequencing and whole genome sequencing were performed. The study found that when treated with modified-FOLFIRINOX (oxaliplatin 85 mg/m², leucovorin 400 mg/m², irinotecan 150 mg/m², and continuous fluorouracil 2400 mg/m²; without bolus fluorouracil; every 2 weeks for 12 cycles) patients with a classical subtype had the best progression free survival. This trial also validated a biomarker for subtyping the disease – *GATA6* expression measured using RNA *in situ* confirms the classical subtype of PDAC.

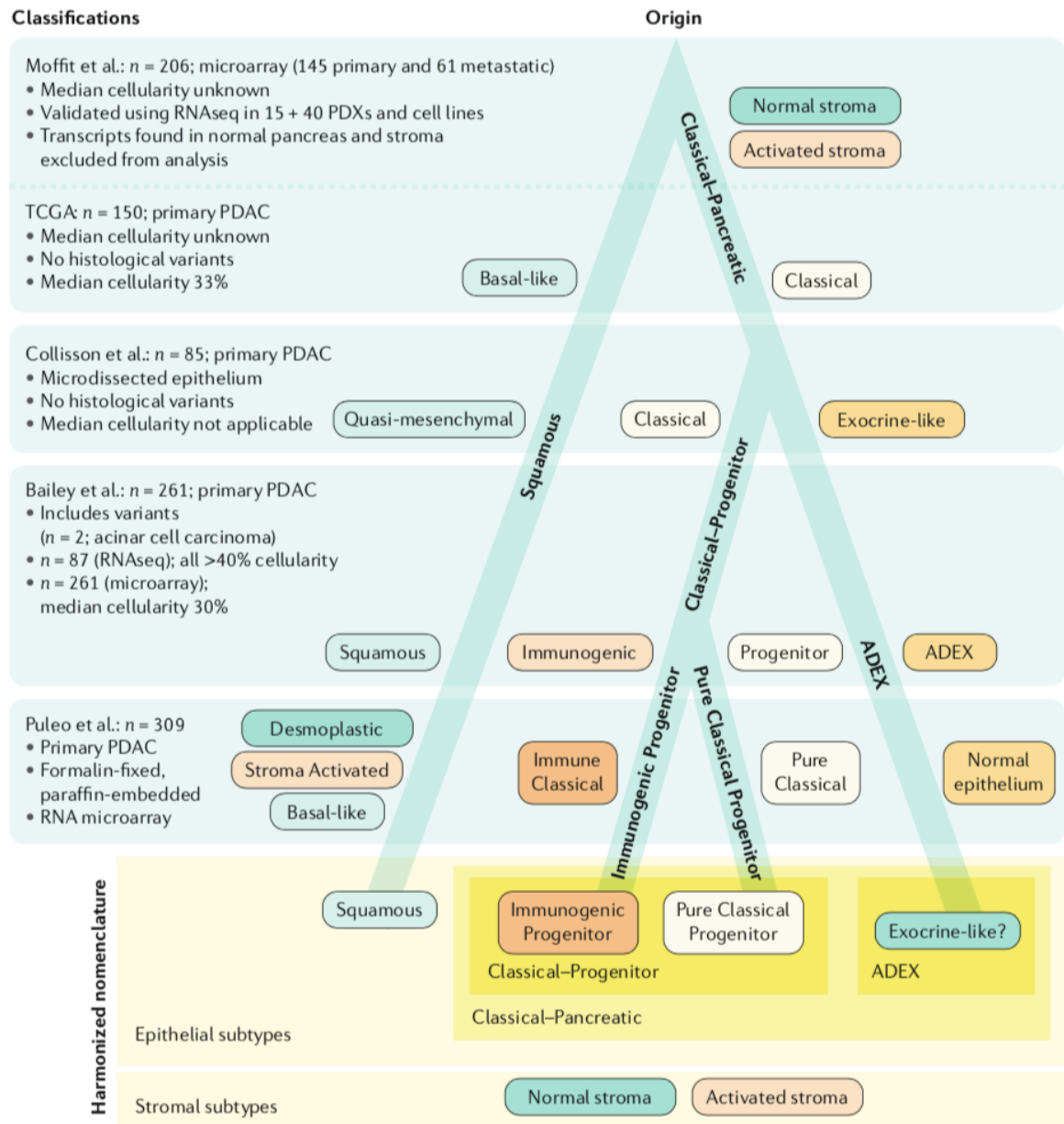


Figure 1.11: A phylo-transcriptomic tree of PDAC, illustrating the classifications of PDAC based on the studies by Moffitt et al. [65], Collisson et al. [66], Bailey et al. [68] and Puleo et al [70]. Taken from Collisson et al. [71].

1.5.3. Predisposition gene mutations

Familial cancer syndromes including Peutz-Jegher Syndrome (PJS), pancreatitis, familial atypical multiple mole and melanoma syndrome (FAMMM), Lynch Syndrome, Hereditary Breast and Ovarian Cancer (HBOC) syndrome and Familial adenomatous polyposis (FAP), account for approximately 5–10% of pancreatic cancers. Table 1.2 contains an outline of the diseases and syndromes associated with an increased risk of developing PDAC.

Table 1.2: Familial Cancer Syndromes associated with an increased risk of developing PDAC. The table includes increased risk, genes associated with syndrome/disease, pathways associated with syndrome/disease and pathway function. Taken without permission from Nelson and Walsh [72].

Title	PJS ¹	Pancreatitis	FAMMM ²	Lynch Syndrome	HBOC ³	FAP ⁴
Increased Risk	132-fold	69-fold	13–22-fold	8.6-fold	3.5–10-fold	4.5–6-fold
Genes	STK11/LKB1	PRSS1 SPINK1 CFTR	CDNK2A	MLH1 MSH2 MSH6 PMS2	BRCA1 BRCA2 PALB2	APC
Pathways	AMPK/mTOR	Trypsin	Retinoblastoma	Mismatch repair	Homologous recombination repair	Wnt signalling
Pathway Function	Cell growth Polarity Metabolism	Auto-activation of trypsin	G1 to S-phase checkpoint	Maintenance of genomic stability	Repair of double-stranded breaks in DNA	Regulation of gene transcription

¹Peutz-Jegher Syndrome; ²Familial atypical multiple mole and melanoma syndrome; ³Hereditary Breast and Ovarian Cancer syndrome; ⁴Familial adenomatous polyposis

1.5.3.1. Peutz-Jegher syndrome

PJS is a rare autosomal dominant disease, characterised by gastrointestinal polyposis, mucocutaneous pigmentation, and cancer predisposition [73]. PJS increases the risk of several malignancies, including breast, pancreatic and gynaecological cancers [74]. Individuals with PJS have a 132-fold increased risk of developing PDAC [75]. It is caused by a mutation in *STK11*, also known as liver kinase B1 (*LKB1*). *STK11/LKB1* is a serine/threonine protein kinase which drives many cell functions, including cell growth, regulation of metabolism and cell polarity, mainly through AMP-activated protein kinase/ mammalian target of rapamycin (AMPK/mTOR) signalling [76]. The most common *STK11/LKB1* mutations are deletions or inactivating mutations. Using genetically engineered mouse models (GEMM), a study by Helez *et al.* [77] showed *STK11/LKB1* deletion resulted in defective acinar cell polarity, abnormal cytoskeleton, loss of tight junctions, and progressive acinar degeneration. Deletion of *STK11/LKB1* in the pancreas also resulted in the development of serous cystadenomas. Morton *et al.* [78] showed that the *STK11/LKB1* deletion resulted in accelerated *KRAS*^{G12D} tumorigenesis, through decreased *TP53* and *p21* dependent growth arrest. These studies, along with others provide strong evidence for a tumour suppressor function for this gene [79].

1.5.3.2. Pancreatitis

Pancreatitis is the second most common hereditary cause of PDAC. Pancreatitis is an inflammatory disorder of the pancreas, caused by the premature activation or lack of inhibition of digestive enzymes. There are several forms of hereditary pancreatitis, including a gain of function mutation in serine-1 protease gene (*PRSS1*), which makes trypsinogen [80]. This gain of function mutation results in increased trypsinogen auto-activation, triggering pancreatic self-digestion. Other genes associated with hereditary pancreatitis include *SPINK1*, a pancreatic secretory trypsin inhibitor and *CFTR* (cystic fibrosis transmembrane regulator) [81]. The chronic inflammation of the pancreas, which characterises pancreatitis results in the presence of reactive oxygen species (ROS) in the pancreas [82]. These ROS, including nitric oxide and free radicals inhibit apoptosis, and can result in direct DNA damage, resulting in oncogenic mutations in genes such as *KRAS*,

CDKN2A and *TP53* [83,84]. Cytokines which are released in response to pancreatitis activate pancreatic stellate cell and result in the development of fibrosis, facilitating the development of PDAC [85–87]. People with chronic or hereditary pancreatitis have a 69-fold increased risk of developing pancreatic cancer [88].

1.5.3.3. Familial atypical multiple mole and melanoma syndrome

FAMMM is characterised by melanoma in more than one first- or second-degree relative, high total body mole count (often >50), and moles with certain histopathological features. The melanomas can arise from the atypical moles or *de novo*, superficially spreading melanoma and/or nodular melanoma [89]. Three original descriptions in different kindreds implicated germline mutations or microdeletions in *CDKN2A*, in particular the p16^{INK4a} isoform, as causative of FAMMM [90]. FAMMM is mainly associated with the autosomal dominant familial melanoma, but patients also have a 13 to 22-fold increased risk of PDAC [91]. A study by Vasen *et al.* [92] identified individuals with an increased risk of developing PDAC from a family history, or families with a gene defect which results in the disease. An annual magnetic resonance imaging (MRI) scan was provided, and PDAC was identified in 7.3% of the *CDKN2A* mutation carriers, resulting in a resection rate of 75% and an overall 5-year survival rate of 24%.

1.5.3.4. Lynch Syndrome

Lynch Syndrome is also associated with an increased risk of colorectal cancer and PDAC [93]. It is caused by mutations in the MMR genes mainly MutL homolog 1 (*MLH1*), MutS homolog 2/6 (*MSH2/MSH6*) and PMS1 Homolog 2 (*PMS2*). The MMR maintains the integrity of the genome by repairing DNA replication errors [94]. Biallelic loss of MMR genes results in genomic instability, and an increase of unrepaired replication errors, particularly affecting repeats, such as microsatellites, termed microsatellite instability-high (MSI-H). MSI-high results in genome hypermutability, with a 100 to 1000-fold increase in mutations [95,96]. Individuals with Lynch Syndrome have 8.6-fold increased risk of PDAC [97].

1.5.3.5. Hereditary breast and ovarian cancer syndrome

HBOC syndrome is caused by mutations in the tumour suppressor genes *BRCA1* and *BRCA2*. People with this syndrome have a 3.5 to 10-fold increased risk of developing PDAC [98]. Mutated variants of the Partner and Localiser of *BRCA2* (*PALB2*) gene are also associated with a familial risk of PDAC [99]. *PALB2* has a critical role in homologous recombination repair (HRR) and recruits *BRCA2* and *RAD51* to DNA breaks. Jones *et al.* [100] found that in 96 patients with PDAC, three had truncating mutations in the *PALB2* gene, producing a stop codon, which was not present in 1084 healthy controls. Slater *et al.* [101] observed a similar prevalence of *PALB2* mutations (3.7%) in a panel of 81 European patients with familial pancreatic cancer. The study of a 61-year-old patient with advanced localised PDAC who presented with a biallelic inactivation of *PALB2*, found treatment with mitomycin C – a chemotherapy and antibiotic which acts via crosslinking complementary strands of DNA – resulted in disease regression [102]. At the 3-year follow up, the patient remained asymptomatic. Studies have also shown that patients with wild type *PALB2* are mitomycin C resistant [103].

1.5.3.6. Familial adenomatous polyposis

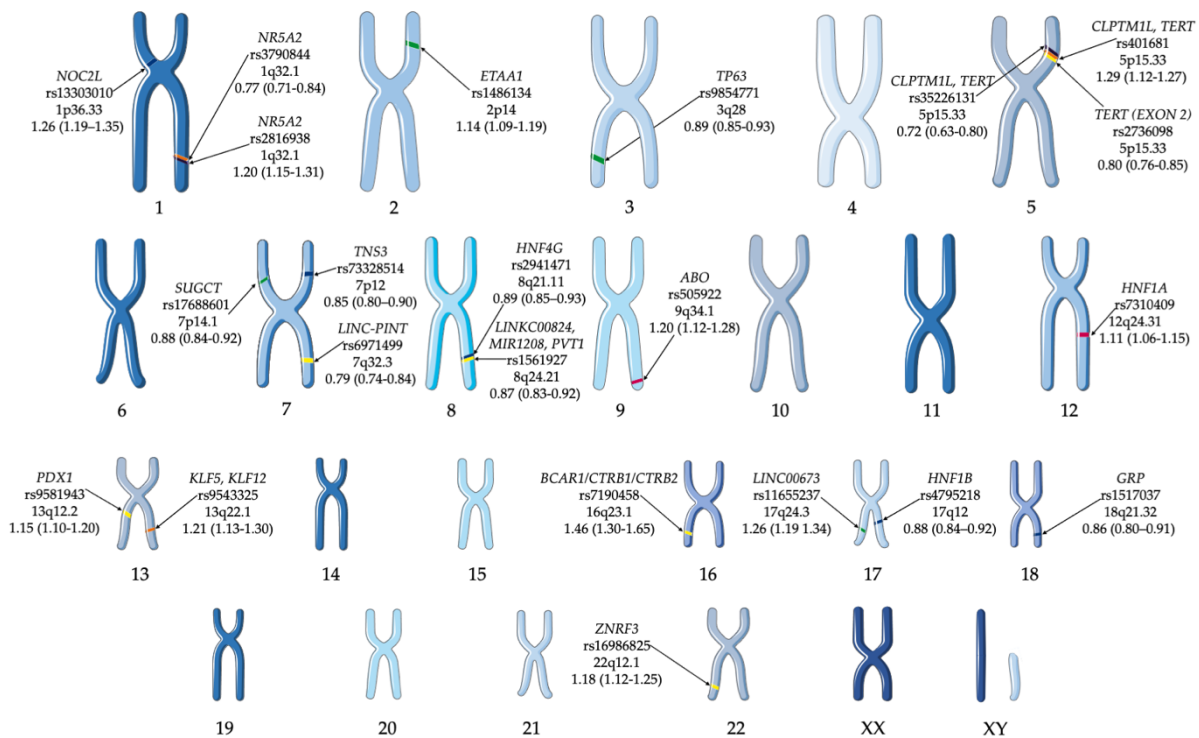
FAP is a familial cancer syndrome which results in an increased risk of colorectal cancer [104,105]. It is characterised by colorectal polyps, due to a mutation in the adenomatous polyposis coli (*APC*) gene. *APC* acts to negatively regulate the Wnt signalling pathway [106]. The Wnt proteins stabilise cytosolic β -catenin, which associates with the transcriptional regulators T cell factor/lymphoid enhancer factor-1 family (TCF), thereby allowing the expression of Wnt-regulated genes [107]. Murine studies of colorectal cancer have found that mutations in the *APC* gene result in hyperproliferation of cells [108]. Individuals with FAP have 4.5 to 6-fold increased risk of PDAC [105].

1.5.4. GWAS risk loci

Recently, the landscape of pancreatic cancer has been redefined through gene expression and genetic diversity signatures identified using next-generation sequencing (NGS) and genome wide association studies (GWAS). GWAS examines

hundreds of thousands of variants, in thousands of individuals, to identify genotype-phenotype associations, and helps to identify risk factors for multifactorial diseases [109]. Through this, GWAS can enable the identification of people at risk of developing a disease, and also can be used for the examination of the biological underpinnings of a disease. GWAS enables the use of potential preventative measures for those who are identified as at risk, and also for the development of treatments for the disease. GWAS use single nucleotide polymorphisms (SNPs) which are single base pair changes in the genome. SNPs can occur in the gene, in both introns and exons, which can result in amino acid changes, different mRNA splicing and reduce the mRNA transcript stability [110]. SNPs can also occur in the transcriptional regulatory elements such as transcription factor binding sites, enhancers and promoters, resulting in altered mRNA expression [111].

The Pancreatic Cancer Cohort Consortium (PanScan) and the Pancreatic Cancer Case-Control Consortium (PanC4) have performed GWAS of PDAC in populations with European ancestry, and have identified common variants associated with risk of pancreatic cancer in European populations [112–117]. This work has led to the identification of 20 genomic loci and 22 independent risk signals on chromosomes 1q32.1 (two independent signals in *NR5A2*), 1p36.33 (*NOC2L*), 2p14 (*ETAAI*), 3q28 (*TP63*), 5p15.33 (three independent risk loci in the *CLPTMIL-TERT* gene region), 7p12 (*TNS3*), 7p14.1 (*SUGCT*), 7q23.2 (*LINC-PINT*), 8q21.11 (*HNF4G*), 8q24.1 (two independent risk loci in the *MYC-PVT1* gene region), 9q34.2 (*ABO*), 13q12.2 (*PDXI*), 13q22.1 (*KLF5*, *KLF12*), 16q23.1 (*BCAR1*), 17q12 (*HNF1B*), 17q24.3 (*LINC00673*), 18q21.32 (*GRP*) and 22q12.1 (*ZNRFB3*) (Figure 1.12). GWAS of PDAC has also been performed in populations from China and Japan, which have resulted in the identification of eight additional PDAC risk loci on 5p13.1 (*DAB2*), 6p25.3 (*FOXQ1*), 7q36.2 (*DPP6*), 12p11.21 (*BICD1*), 10q26.11 (*PRLHR*), 21q21.3 (*BACH1*), 21q22.3 (*TFF1*), and 22q13.32 (*FAM19A5*) [118,119].



Study Name
Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer
A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33
Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer
Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer
Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer
Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21

Figure 1.12: Significant single nucleotide polymorphisms (SNPs) identified in pancreatic cancer cases of European ancestry. Highlighted GWAS SNP, closest gene, chromosome and odds ratio (95% confidence interval) [112–117]. This figure was taken without permission from Nelson and Walsh [72].

1.5.5. GWAS pathway analysis

While the results of the GWAS have informed the genetic component of predisposition loci, it does not give a clear indication of the cause of PDAC. The use of complementary GWAS pathway analysis – a method of analysing this genomic data through sets defined by functional pathways – offers the potential of greater power for discovery and natural connections to biological mechanisms. Pathway analysis allows for the identification of causative SNPs whose individual effects may not be significant enough to be detected in GWAS [120].

In 2012, Li *et al.* [121] performed a pathway analysis study using the adaptive rank truncated product (ARTP) method, this method combines SNP-association statistics across SNPs in a gene or pathway, and each gene includes up to two SNPs. This study used GWAS data from 3,851 PDAC cases, and 3,934 controls (PanScan), and compiled 23 biological pathways hypothesised to be relevant to PDAC. This study identified five pathways with an association with the development of PDAC (Table 1.3).

Following on from this study, Walsh *et al.* [122] performed a pathway analysis with 9040 PDAC cases and 12,496 controls from the PanScan and PanC4 GWAS. This study identified 14 pathways and gene sets associated with the development of PDAC. The strongest pathways associated with the development of the PDAC included genes involved in pancreas development and the regulation of beta-cell development, maturity onset diabetes of the young (MODY), the Breast Cancer 17Q11-Q21 amplicon, cardiac hypertrophy and the *ATM* Pearson Correlation Coefficient (PCC) Network (Table 1.3).

In this thesis, SNPs and genes from the MODY pathway and the *ATM* PCC Network which were identified as top gene sets associated with the development of PDAC will be analysed to identify if (1) SNPs have a functional effect and (2) genes within these pathways have a role in the development and progression of the disease.

Table 1.3: Gene sets/ pathways identified for risk of developing PDAC from pathway analysis studies by Li et al. [121] and Walsh et al. [122].

Pathway/Gene Set	Pathway Reference	Study	Pathway P-value
Maturity onset diabetes of the young	KEGG ³	Walsh et al.	5.10E-07
Regulation of Beta-cell development	REACTOME	Walsh et al.	1.92E-06
Breast Cancer 17Q11-Q21 amplicon ¹	NIKOLSKY	Walsh et al.	2.00E-06
Role of EGF Receptor Transactivation by GPCRs in Cardiac Hypertrophy	BIOCARTA	Walsh et al.	3.79E-06
ATM Pearson Correlation Coefficient (PCC) Network ²	PUJANA	Walsh et al.	1.25 E-05
Pancreas development		Li et al.	2.00E-06
<i>Helicobacter pylori</i> lacto/neolacto		Li et al.	1.60E-05
Hedgehog		Li et al.	1.90E-03
Th1/Th2 immune response		Li et al.	1.90E-02
Apoptosis		Li et al.	2.30E-02

P-values adjusted for multiple testing. ¹Genes within amplicon 17q11-q21 identified in a copy number alterations study of 191 breast tumour samples. ²Gene network transcripts whose expression positively correlated with ATM gene in normal tissues. ³KEGG: Kyoto Encyclopedia of Genes and Genomes.

1.6. Diabetes

Approximately 80% of PDAC patients have glucose intolerance, or diabetes at the time of their diagnosis [123]. The majority of diabetic cases associated with PDAC are diagnosed with, or within two years of, PDAC diagnosis [7]. Type 1 diabetes mellitus (T1DM) is a chronic autoimmune disease, which is characterised by insulin deficiency due to the loss of β -cells, and results in hyperglycaemia. T1DM onset usually occurs during childhood, or early adolescence [124]. Onset of Type 2 diabetes mellitus (T2DM) is usually in adults. It is associated with insulin resistance due to being overweight, a lack of physical activity, dietary, hereditary and stress response factors [125]. Diabetes is the third highest independent epidemiological risk factor for pancreatic cancer after smoking and obesity [8,126,127]. T2DM patients who had the disease for more than five years have a 1-1.5-fold increased risk of developing PDAC [128]. However, T2DM patients who have had the disease for less than one-year have a 5.4-fold increased risk of developing PDAC, with substantial research suggesting PDAC result induces hyperglycaemia and diabetes [129].

Use of anti-diabetic medication can modify PDAC risk [130]. Metformin, a biguanide oral hypoglycaemic drug is used in the treatment of T2DM and pre-diabetes. It works by increasing insulin sensitivity, and by decreasing liver glucose production [130]. A study has found that users of metformin have a decreased risk of developing PDAC, with an odds ratio (OR) of 0.38 (CI 0.22-0.69) [131]. However, use of sulfonylureas resulted in a 70% increased risk of PDAC [132]. Pancreatic cancer patients with diabetes treated with insulin had a worse outcome than those not treated with insulin, with a survival of 4.3 vs. 5.5 months [130].

1.6.1. MODY

MODY is a rare form of diabetes which accounts for 1-2% of cases [133]. MODY typically develops before the patient is 25 years old. Unlike T1DM and T2DM which are multifactorial disorders, which involve several genes and have a strong environmental component, MODY is a monogenic disease. MODY patients typically have a family history of diabetes in two or more generations [134].

Fourteen forms of MODY have been identified as mutations in autosomal dominant genes which disrupt insulin production or release, as described in Table 1.4. It is usually caused by mutations in genes which act as transcription factors.

The four most common forms of MODY – 1, 2, 3 and 5 account for 90% of the cases. Diagnosis of MODY, and the identification of the genetic subtype is important, as age of onset, response to treatment, and hyperglycaemia levels differ [135]. Diagnosis of MODY instead of T1DM or T2DM allows for the identification of the correct treatment for patients [136]. The genes involved in the MODY pathway are shown in Figure 1.13.

In a pathway analysis study previously performed in our group by Walsh *et al.* [122] the MODY pathway was identified as the most significant pathway associated with risk of developing PDAC. Regulation of beta-cell development was the next top pathway associated with risk of PDAC development. The genes *HNFI1A*, *HNFI1B*, *HNFI1G*, and *HNFI1A* are present in both of these pathways.

Table 1.4: Outline of types of maturity onside diabetes of the young (MODY), causative gene, percentage of cases and treatment. Modified without permission from Radha et al. [136] and Gardner et al. [137].

Type	Gene	% of Cases	Treatment
MODY1	<i>HNF4A</i>	5-10%	Sulfonylureas
MODY2	<i>GCK</i>	30-50%	Treatment usually not required
MODY3	<i>HNF1A</i>	30-60%	Low dose Sulfonylureas
MODY4	<i>PDX1</i>	1%	Treatment usually not required
MODY5	<i>HNF1B</i>	5-10%	Insulin
MODY6	<i>NEUROD1</i>	<1%	—
MODY7	<i>KLF11</i>	<1%	—
MODY8	<i>CEL</i>	<1%	—
MODY9	<i>PAX4</i>	<1%	—
MODY10	<i>INS</i>	<1%	Insulin
MODY11	<i>BLK</i>	<1%	—
MODY12	<i>ABCC8</i>	<1%	Sulfonylureas
MODY13	<i>KCNJ11</i>	<1%	Sulfonylureas
MODY14	<i>APPL1</i>	<1%	—

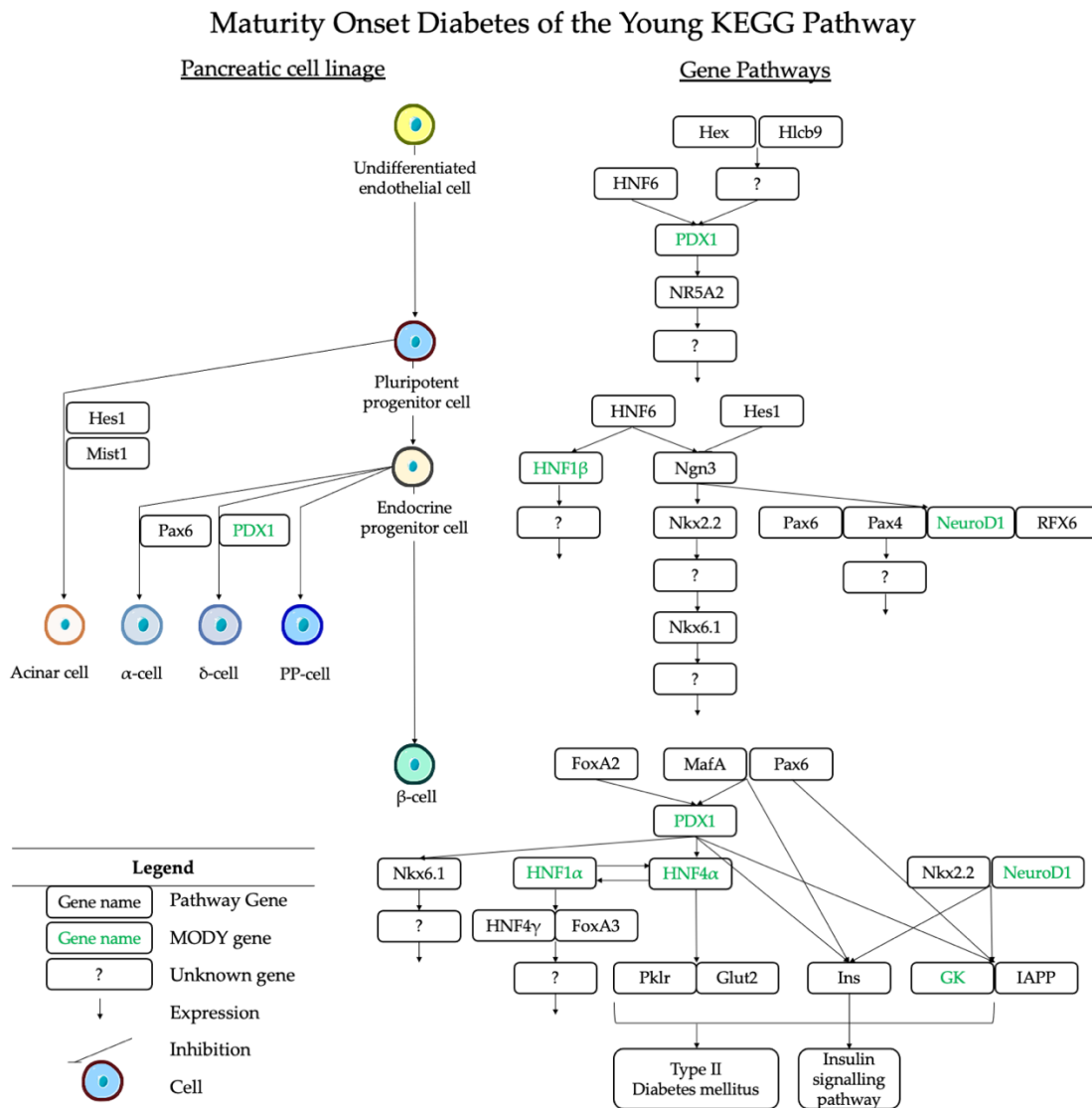


Figure 1.13: Maturity onset diabetes of the young (MODY) KEGG (Kyoto Encyclopaedia of Genes and Genomes) Pathway. The left of the image represents the pancreatic cell lineage and the right illustrates the gene pathways associated with cell differentiation. Gene names in black indicate pathway genes and the question marks (?) represent unknown genes. Gene names in green highlight MODY causative genes: *HNF4α* - MODY1; *GK* - MODY2; *HNF1α* - MODY3; *PDX* - MODY4; *HNF1β* - MODY5, and *NeuroD1* - MODY6. Figure legend is inset in the bottom left of the figure.

1.6.1.1. MODY1

HNF4A (hepatic nuclear factor-4 alpha) mutations account for 5-10% of the cases of MODY1 in the UK [138]. *HNF4A* is a master regulator of epithelial differentiation in multiple tissues and is vital for beta-cell function. It controls the expression of 11% of the pancreatic islet genes, and is involved in differentiation, homeostasis and metabolism [139,140]. The clinical features of MODY1-*HNF4A* are similar to those of MODY3-*HNFI*A, as *HNF4A* expression is regulated by the expression of *HNFI*A [141]. Little is known about the role of *HNF4A* in PDAC, however, its loss results in upregulated *GSK3 β* , and drives a squamous-like profile. The squamous classification is associated with metabolic reprogramming and a poor clinical outcome [142].

1.6.1.2. MODY2

The second most common form of MODY, MODY2, is caused by mutations in *GCK* (glucokinase). *GCK* is an enzyme that facilitates phosphorylation of glucose to glucose-6-phosphate. Glucose is transported into the cell by the GLUT-2 transporter and is then phosphorylated by *GCK* resulting in ATP generation [135]. Mutations in the *GCK* gene impair the proteins function through changes in enzymatic activity, protein stability and increased interaction with its receptor [143].

1.6.1.3. MODY3

The most common cause of MODY in the UK is MODY 3 caused by a mutation in the *HNFI*A (hepatocyte nuclear factor 1 homeobox A) gene. *HNFI*A is a transcription factor that regulates pancreatic differentiation and has a role in maintaining the homeostasis of the endocrine pancreas [144]. It is involved in the transcription of multiple genes, including the glucose transporters *GLUT1* and *GLUT2* [145]. *HNFI*A is a high penetrance mutation – 63% of carriers have MODY by the age of 25, and 96% have MODY by the age of 55 [135]. A truncation or a missense mutation in the binding domain results in a ten-year earlier onset compared to patients with mutations in the transactivation domain [136].

1.6.1.4. *MODY4*

MODY4 is a rare form of *MODY* caused by mutations in *PDX1* (Pancreatic and Duodenal Homeobox 1), a transcription factor required for the embryonic development of the pancreas. *PDX1* is required for multiple functions of the mature beta-cell, including insulin secretion, mitochondrial metabolism and cell survival [146–148]. A study by Park *et al.* [149] has identified increased *PDX1* expression in PDAC, and PDAC precursor lesions and neoplasms. Overexpression of *PDX1* in PDAC cell lines resulted in increased proliferation and invasion, suggesting it acts as an oncogene [150]. *PDX1* controls the expression of *NR5A2* (Nuclear Receptor Subfamily 5 Group A Member 2), an orphan nuclear receptor which acts with pancreas-specific transcription factor 1 (*PTF1*) complex to bind to acinar genes [151,152]. Loss of acinar expression is the first step in PDAC development, during this, *NR5A2* expression is decreased [153]. A proposed function of *NR5A2* is that it inhibits the ductal transformation of adult acinar cells [154].

1.6.1.5. *MODY5*

MODY5 is caused by mutations in the *HNF1B* (hepatocyte nuclear factor-1 beta) gene. *HNF1B* acts as a transcription factor involved in the expression of genes in the beta-cells, kidneys, kidney tracts and uterus. Patients with *MODY5* usually present with diabetes and kidney problems – including renal cysts and diabetes syndrome. *MODY5* patients are not sensitive to sulphonylurea drugs, and progress to insulin treatment [155,156]. Increased expression of *HNF1B* has been associated with increased risk of prostate and ovarian cancers [157,158].

1.6.2. The role of the *HNF1* family in PDAC

HNF1A and *HNF1B* are members of the hepatocyte nuclear factor family, which also includes *FOXA1/2/3*, *HNF4A/G* and *ONECUT1/2* [159]. These were first identified in the liver, but are expressed in a number of organs, including the pancreas and kidneys. These transcription factors act in cross-regulatory and auto-regulatory circuits, and play an essential role in regulating the development and in the functioning of multiple tissues [160]. The *HNF1* family is comprised of *HNF1A* and *HNF1B*, whose DNA-binding domain binds to the palindromic consensus

sequence GTTAATNATTANC. A dimerisation domain at the N-terminus of both proteins allows for the formation of homodimers and heterodimers (Figure 1.14) [161,162]. Each gene encodes three isoforms – A, B and C, which have tissue specific roles [163,164]. HNF1A(A) is highly expressed in the foetal pancreas, and HNF1A(B) is highly expressed in the adult pancreas [164]. In knockout studies in mice, a *HNF1A* knockout resulted in stunted growth and diabetes developed two weeks after birth [165]. *HNF1B* knockouts were lethal at the blastocyst stage (E3.5) because of abnormal or absent extraembryonic endoderm [166].

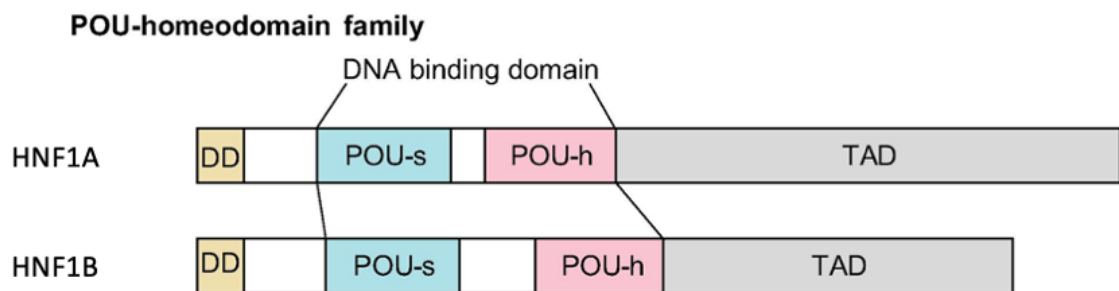


Figure 1.14: A schematic diagram of the structural domains of HNF1A and HNF1B.

The N-terminus consists of the DD (dimerisation domain). The (DNA-binding domain) DBD of the family is made up of a POU-s (POU-specific domain) and a POU-h (POU-homeodomain). The C-terminus contains the TAD (transactivation domain) which is less conserved between the two members of the HNF1 family.

Figure used without permission from Lau *et al.* [160].

1.6.2.1. HNF1A

A number of studies investigating the role of *HNF1A* in PDAC have been performed, with contradictory results. A study by Luo *et al.* [144] showed that there are significantly lower levels of HNF1A in pancreatic cancer tumours in comparison to normal adjacent tumours, suggesting that HNF1A plays a tumour suppressor role. A study by Kalisz *et al.* [167] showed that HNF1A recruits KDM6A (Lysine-specific demethylase 6A), which acts to catalyse the demethylation of histone H3K27me₃, to genomic binding sites in acinar cells. This results in the remodelling of the acinar enhancer landscape and activation of the differentiated acinar cell programs, thereby suppressing oncogenic and epithelial-mesenchymal transition

genes. Other studies, including transcriptome analysis of PDAC also suggest a tumour suppressor role for *HNFLA* in PDAC [168,169].

However, a study by Abel *et al.* [170] showed that *HNFLA* overexpression in PDAC increased cell growth, and resulted in increased pancreatic cancer stem cell marker expression. Increased expression of *HNFLA* has also been found to have a role in increasing PDAC sensitivity to gemcitabine treatment both *in vivo* and *ex vivo*, by regulating *ABCB1* expression through binding to its specific promoter region and suppressing its transcription levels [171]. An *in vitro* study by Subramani *et al.* [172] found expression of *HNFLA* is varied in PDAC cell lines, and *HNFLA* gene suppression through siRNA knockdown resulted in increased apoptosis associated proteins in cells, and resulted in decreased migrations and invasion of cells. It is evident that the role of *HNFLA* in the development and progression of PDAC has been largely debated, with conflicting results identifying *HNFLA* as both a tumour suppressor and an oncogene.

1.6.2.2. *HNFB*

Through an integrative analysis of online databases, Nie *et al.* [173] performed a study looking at the role of *HNFB* in human cancers. This study found that *HNFB* is dysregulated in several cancers, and that its expression is associated with immune cell infiltration. The study also found that patients who harboured *HNFB* mutations were more susceptible to mutations in other genes. In PDAC, *HNFB* was found to be frequently mutated, and its expression was also associated with increased immune cell infiltration.

An *in vivo* study by Quilichini *et al.* [174] found that *HNFB* deletion in pancreatic duct cells resulted in dilatation of ducts, loss of acinar cells and ADM. The study also found that the *HNFB* deletion resulted in PanIN and in combination with *KRAS* mutations, resulted in PanIN progression.

1.7. Role of Apoptosis in DNA repair pathway

During the cell cycle, cells go through a series of cell cycle check points before cell division [175]. In response to stress, the checkpoints are activated to prevent progression through the cell cycle and initiate DNA repair [176] (Figure 1.15). To ensure the elimination of potentially damaged cells if DNA damage is beyond repair, apoptosis is induced [177]. This process is called the DNA damage response (DDR), DDR proteins include some of the most frequently mutated genes in cancer, as they are able to induce the mutator phenotype and act as a driver of more mutations [177].

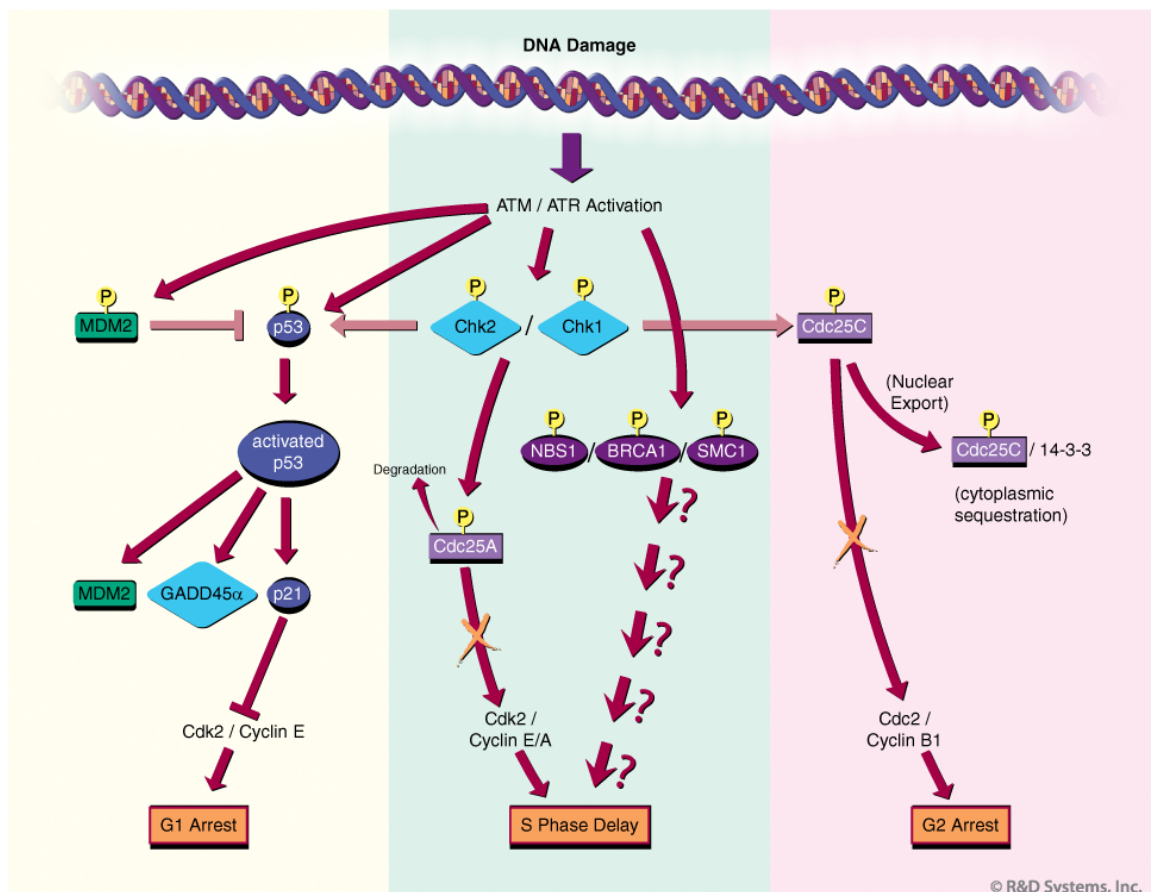


Figure 1.15: DNA repair pathway for DNA damage, ATM and/or (ataxia telangiectasia and Rad3-related) ATR trigger the activation of a checkpoint that leads to cell cycle arrest or delay. Checkpoint pathways are characterised by cascades of protein phosphorylation events (indicated with a "P") that alter the activity, stability, or localisation of the modified proteins. Figure used without permission from R&D Systems [178].

A recent study by Peterson *et al.* [179] investigated 303 patients with familial PDAC, and found germline mutations from ten genes, most of which are involved in DNA repair (*ATM*, *BRCA1*, *BRCA2*, *CDKN2A*, *PALB2*, *PMS2*, *BARD1*, *CHEK2*, *MUTYH*, and *NBN*), are present in 12% of the patients. Another study performed by Golan *et al.* [180] found that the prevalence of *BRCA* mutations in PDAC is 6-7%, and up to 20% in Ashkenazi Jews. In the phase III clinical trial – POLO (Pancreas cancer OLaparib Ongoing), platinum-sensitive patients with germline mutations in the *BRCA1/BRCA2* genes were found to respond to treatment with olaparib, a (poly ADP ribose polymerase) PARP inhibitor [181]. This trial used a biomarker for the selection and stratification of patients was the first of its kind in PDAC.

One of the hallmarks of cancer is the ability to evade apoptosis, as defects in damage induced apoptosis contribute to tumorigenesis and resistance in cancer [182]. Caspases (cystine-aspartate proteases) play a central role in apoptosis. Activation of caspases result in chromatin condensation, loss of cell adhesion, cell shrinkage, production of apoptotic bodies and engulfment by phagocytes [183]. Caspases are activated in a sequential manner through initiator and executioner caspases. Activation of caspases requires cleavage into two subunits, a small subunit and a large subunit. The active enzyme is a heterotetramer of two large and two small subunits. Inflammatory caspases (4, 5, 11) and, initiator caspases (2, 8, 9, 10) initiate the apoptosis process, while executioner caspases (3, 6, 7) perform the mass proteolysis that leads to apoptosis [183,184] CASP7 has a role in the execution phase of cell apoptosis [185]. It is cleaved by the initiator caspases CASP8 and CASP10, and its cleavage results in the induction of apoptosis [183].

The GWAS pathway analysis performed by Walsh *et al.* [122] identified the Pujana ATM PCC Network as one of the top five gene sets for risk of developing PDAC [186]. Within this gene set, a SNP rs3124737 (*CASP7*) was identified as having a positive influence on *CASP7* expression in normal pancreas using Genotype-Tissue Expression (GTEx) which provides data on tissue-specific gene expression and regulation, and Laboratory of Translational Genomics (LTG), which investigates the biological basis of common inherited genetic variants associated with cancer

susceptibility. Patients who are homozygous for this allele have a lower risk of developing PDAC (odds ratio=0.91; 95% CI 0.88-0.95) risk of developing PDAC. It was also observed to be an expression quantitative trait loci (eQTL), a locus that explains a fraction of the genetic variance of a gene expression phenotype effect, on *CASP7* mRNA in other tissues from GTEx such as whole blood, testis, and thyroid.

1.8.Models of PDAC research

1.8.1. Established PDAC cell line cultures

Two-dimensional (2D), cell-based assays are an important tool, and have been the mainstay of cancer research for over 50 years. The first cell line (HeLa) was developed in 1951 from cervical carcinoma [187]. Cell lines are able to grow indefinitely, making them an easy to use, low cost, repeatable models, and thus are important for both drug discovery, and proof-of-concept studies. Cell lines often undergo genetic modifications, including copy number variation and point mutations during passaging [188]. Cell lines tend to be homogeneous, which does not represent the heterogeneous nature of PDAC tumours. Also, cell lines are often developed from late stage, aggressive tumours, so they cannot be used to model tumour progression [189].

PDAC cell lines recapitulate the genomic changes which lead to the development of the disease. The four most common mutations (*KRAS*, *TP53*, *CDKN2A*, *SMAD4*) occurring in PDAC tumours are found in cell lines at similar percentages and PDAC cell lines also demonstrate the different phenotypes and genotypes which are found in PDAC subclasses. A commonly used PDAC cell line is BxPC3, developed from pancreatic adenocarcinoma of a 61-year-old female. BxPC3 has *TP53* mutations, a homozygous deletion in *SMAD4*, but is *CDKN2A* wild type and is the only *KRAS* wild type PDAC cell line available for research. Other common PDAC cell lines include, PANC-1, developed from PDAC of a 56-year-old male and MIAPaCa-2, developed from a PDAC of a 65-year-old male, both of which harbour mutations in *KRAS* and *TP53*, with homozygous deletion in *CDKN2A* and wild type *SMAD4*. Capan-1 was developed from a liver metastasis of a 40-year-old male with

PDAC, and harbours mutations in *KRAS*, *TP53*, *CDKN2A* and *SMAD4*, and is the only PDAC cell line with a *BRCA2* mutation currently available for research. A detailed review from Deer *et al.* [190] provides an overview of the available information on the most commonly used PDAC cell lines.

Due to genomic drift, differences in cell culture procedures and media in different labs may result in genotypic and phenotypic differences in the same cell line. Ben-David *et al.* [191] performed a full genomic characterisation of 27 different strains of the ER-negative breast cancer cell line MCF7. Changes observed included differential activation of gene expression programs, morphology and proliferation. Drug sensitivity was shown to vary in the cell lines, whereby 75% of the drugs tested which showed strong cytotoxicity in some of the MCF7 cell lines, were completely inactive in others.

Another issue with the use of established cell lines is cross contamination. Boonsta *et al.* [192] identified two oesophageal adenocarcinoma cell lines which were contaminated, and have been used in 11 patents, and more than 100 published studies, leading to clinical trials. Horbach and Halffman [193] identified 32,755 articles reporting on research with misidentified cells, which in turn have been cited by over half a million papers. To overcome these issues, a number of journals require cell lines to be verified before publishing a research paper. A method used to validate the cell lines is short tandem repeat (STR) profiling. This technique was initially developed for forensic applications [194]. STR profiling compares microsatellite (2 to 7 base pairs) repeats at specific loci which collectively are unique to each individual [195]. It is carried out by using commercially available PCR primers and amplicons are compared to size markers, allowing for a comparison of the lengths of the PCR products at each locus to the STR profile made from the original donor material [196].

Two-dimensional established cell line models have been the standard method for cancer drug testing for many years. However, of late, the limitations of using established cell lines in 2D are being increasingly recognised. In actuality, 2D cell

culture platforms often fail to recapitulate the physiology of tumours *in vivo* due to different cellular architecture, adherence structures and biochemical gradients.

1.8.2. Primary tumour cultures

Primary cell lines are an emerging tool for cancer research. These cell lines are derived from a patient tumour or biopsy, dissociated, and grown *in vitro* [197]. Primary cell lines are heterogeneous, and are at an early passage number, so are more representative of the original tumour [198]. Primary cell lines may allow for the development of personalised cancer therapy through the development of primary cell lines from patient tumour, and the function testing of chemotherapeutic drugs on the living cancer cells [199]. While primary cultures are more representative of the original patient tumour, there are several issues with PDAC primary cell lines – they are often difficult to establish, only grow for a limited number of passages, and often tumour cells are overgrown by stromal cells such as fibroblasts [198,199].

1.8.3. Patient derived xenografts

Patient derived xenografts (PDX) are another commonly used model of PDAC. A patient tumour is implanted subcutaneously or orthotopically in severe combined immune deficiency (SCID) mice until the tumour has grown to a sufficient size to be sub-cultured in new mice. These models allow for tumours to have the original cell-to-cell interactions [200]. The original tumour microenvironment can also be recapitulated using orthotopic implants. A study by Garrido-Laguna *et al.* [201] showed that orthotopically implanted tumours treated with gemcitabine had a similar response to that in patients, which was not observed in the subcutaneous implanted tumours.

PDX studies have been used for the identification of biomarkers of PDAC. Jimeno *et al.* [202] used 11 PDX tumour samples, with known gemcitabine sensitivity to identify biomarkers for gemcitabine response in patients. This group exposed fine-needle biopsies of the PDX tumour to gemcitabine or vehicle control for 6 hours and compared gene expression of the treated and untreated samples using RT-

qPCR 45-gene array. This assay identified that Polo-Like Kinase 1 (*Plkl*), a serine/threonine protein kinase had differential expression of >50% in the sensitive samples compared to resistant tumours. To further validate this biomarker, siRNA knockdown was performed and the *Plkl* pathway inhibitor rigosertib, a *Plkl* pathway modulator. This resulted in a synergistic effect with gemcitabine in gemcitabine-resistant *in vitro* models. The study illustrates the ability to use PDX models to identify and validate a biomarker of PDAC.

There are many advantages to using PDX tumours for the study of pancreatic cancer. Tumours can be established in mice using a small amount of patient material; tumours retain the heterogeneity, as well as the genetics, and histological characteristics of the original tumour during passaging. PDX tumours also provide an unlimited source of tumour, which can be used for *in vivo* and *ex vivo* drug testing. Nevertheless, there are several disadvantages to the use of PDX models – they are expensive, time consuming, require the use of animals, and their use is subject to strict regulations [203]. PDX models take up to four months to develop tumours. Subcutaneously implanted tumours are not grown in the same microenvironment as PDAC tumours, and rarely form metastases [204]. As the tumour is grown, and sub-cultured in mice, the human stromal cells, such as fibroblasts and blood vessels are replaced by murine cells [205]. Finally, as SCID mice do not have an immune system, the PDX tumours cannot recapitulate the complex interactions between the PDAC tumour and the immune system, which is critical in resistance mechanisms of PDAC, and it also prevents the use of PDX models in testing of immune modulating drugs, which are increasingly being used in cancer treatment.

1.8.4. Organoids

The use of three-dimensional (3D) *in vitro* PDAC models can overcome many of the limitations of traditional cancer research models. As they are not attached to plastic, 3D models have more appropriate physiological morphology and signalling pathways compared to cells grown in 2D [206]. Similar to *in vivo* conditions, 3D cultures are exposed to complex environments, with varied exposure to oxygen,

nutrients, stress and waste. The use of 3D cultures also allows for the study of cell-to-cell interaction, drug penetration, response and resistance [207,208]. Another advantage of using 3D models, is the cultures contain cells in multiple growth phases, with cells which are proliferating, quiescent, hypoxic and necrotic, whereas cells grown in 2D tend to be in the same growth phase [209]. Thus, 3D models have the same gene and protein expression profiles as the original tumour whereas differential expression is present in 2D cell models [210–212]. Cells which are grown in 3D can also be cultured and tested for longer, as 2D cells require regular trypsinisation as the cells reach confluency faster [213]. Previous studies have shown that cells grown in 2D and 3D have different sensitivity levels to chemotherapeutic drugs, with 3D models showing increased levels of drug resistance, which is more representative of the *in vivo* drug response [214,215]. However, in a study assessing the effect of 129 drugs in 27 liver cancer organoid lines by Li *et al.* [216], the group found that some drugs were highly heterogeneous in the responses among organoid lines, while others were uniformly toxic in all cases.

The first publication describing intestinal organoids was published in 2009, and since then, the methods have been used to create organoids in a large range of tissues and to study a wide range of diseases [217]. Organoids are 3D spheroid cultures which represent the *in vivo* architecture of the organ or tumour. Organoids are developed from stem cells, which self-organise to resemble tissues from within the organ. They can be derived from multiple types of stem cells, including embryonic, induced pluripotent, tumour and normal adult stem cells [218]. Organoids produce a relevant and highly adaptable model cancer system [219]. They are grown within a 3D matrix system, such as hydrogels, basement membrane extract and Matrigel®, which have been supplemented with growth factors allowing them to mimic the organ's microenvironment [220]. A specialised media is required, with multiple growth factors, which mirror the niche organ microenvironments. As organoids are derived from stem cells, they display cell heterogeneity after several passages [219]. Like cell lines, organoids can grow indefinitely and can be cryopreserved, so are a useful PDAC model. Growing organoids *in vitro* also allows for the observation of disease progression, which is

not possible in other models as it requires complex imaging systems in PDX models, or in established cell lines, as these are usually produced from late stage tumours [221–223]. Organoids can also be developed from tiny volumes of patient tumour, such as fine-needle biopsy, and grown to allow the high throughput screening of drugs and drug combinations [224]. The Tuveson Group developed a method that allows for the development of organoids from fine-needle biopsies guided by endoscopic ultrasounds [224]. Organoids can also be developed using tissue biobanks. Walsh *et al.* [225] investigated the morphology, viability and drug response of frozen organoids with both flash freezing, and DMSO frozen organoids. Expression of Ki67 and cleaved caspase3 were assessed to determine viability. Both samples that were flash-frozen and frozen slowly with DMSO were viable, indicating that biobanks of tumour samples could potentially be used for the establishment of organoid cultures.

Organoids can also be orthotopically transplanted into mice, this results in the organoid progressing through all stages of tumour development from PanIN to a PDAC tumour, which represents the tumour of origin [226]. In comparison, when a monolayer of cells is transplanted, the cells rapidly become an aggressive carcinoma [227].

1.8.4.1. Development of PDAC organoids

Studies outlining the methods for the production of human and mouse pancreatic organoids have been published by multiple groups [228–230]. Both Grapin-Botton and Clevers groups developed methods for the production of murine pancreatic organoids in Matrigel®. Grapin-Botton developed pancreatic organoids for use as a model for diabetes [228]. This group used dissociated mouse embryonic pancreatic progenitor cells for the development of organoids. These organoids showed both pancreas morphology and differentiation, and the *in vitro* maintenance of these pancreatic organoids required the activation of both Notch and fibroblast growth factor (FGF) signalling pathways, which recapitulated the *in vivo* niche signalling pathway within the pancreas. Clevers [229] developed a method which allowed for the propagation of adult murine pancreatic duct cells as organoids. These

organoids were embedded in Matrigel®, and a cocktail of growth factors, including R-spondin1 and Wnt3a which stimulate the Wnt signalling pathway.

Clevers, in collaboration with the Tuveson lab, developed a method for the establishment, and growth of normal and cancerous pancreatic organoids from human and mouse tissues [226]. This study showed that both tissues could be established using the same conditions, however, human organoids required additional growth factors, such as Wnt3a. Orthotopic implantation of the tumour and normal organoids resulted in full tumour and ductal development. The methods have been used to further PDAC research, by researching the tumour microenvironment, personalised treatment, genetics and testing of novel therapeutics for PDAC.

To prevent differences in drug response due to batch-to-batch variation of media and extracellular matrix (ECM), Georgakopoulos *et al.* [231] developed a chemically defined, serum free media, and a chemically defined hydrogel, which allows for the development and propagation of human pancreas tissue. Their study found that the organoids retained their ductal morphology, biomarker expression, and genomic integrity after growth for several months.

1.8.4.2. Organoids as PDAC tumour-drug response predictors

Using organoids as predictors of drug response can facilitate advanced pre-clinical drug discovery and the personalised treatment of PDAC. Hou *et al.* [232] used four patient derived organoid lines including two PDAC derived organoids and two cancer-associated fibroblast (CAF) organoids for high throughput screening (HTS) of the National Cancer Institute (NCI) approved oncology set, and 3300 clinically approved drugs.

Huang *et al.* [233] identified a genotype-phenotype relationship where *TP53*^{R175H} induces cytosolic *SOX9* localisation, whereas in normal pancreas *SOX9* is localised in the nucleus. This finding was verified in two independent PDAC cohorts, and cytoplasmic *SOX9* was associated with a higher tumour grade, poor-disease free survival, and poor overall survival. In this study, organoids also showed a poor

response to gemcitabine, the mainstay for PDAC treatment with only 30% growth inhibition.

Studies by Romero-Calvo *et al.* [234] compared the structural and genetic features of organoids with the primary PDAC tumours, and found the organoids had similar morphologic features with the same glandular architecture. The organoids and primary tumours also had the same protein expression, molecular, genomic and transcriptomic profiles. Response to FOLFIRINOX treatment in organoids recapitulated the matched PDX models. Frappart *et al.* [235] also found that organoids derived from PDX tumours faithfully recapitulated the PDX morphology and protein expression, and predicted PDX drug response.

A number of proof-of-concept clinical trials using PDAC organoids are underway using fine-needle biopsies for the development of organoids (ClinicalTrials.gov: NCT03896958, NCT03544255, NCT03990675 and NCT03140592). Tiriach *et al.* [236] created a patient derived organoid library from primary PDAC tumours, and metastases, with 75% success. This organoid library was screened using gemcitabine, paclitaxel, SN38, 5-FU and oxaliplatin at clinically relevant concentrations, and the organoid library showed a heterogeneous response to chemotherapies. The outcomes of these assays paralleled the patient outcomes in the clinic. On publishing, Tiriach *et al.* [236] had also performed whole exome sequencing and RNA-seq and developed gene expression signatures to determine improved response to therapies. In addition to establishing organoid cultures from patient biopsies in these clinical trials, the organoids will be used for drug screening as an indicator of response to therapies. These clinical trials are setting the foundation for the use of organoids in the personalised treatment of pancreatic cancer.

Clinical trials that aim to establish and test organoids to determine the best treatment for patients are currently ongoing. This includes a proof-of-concept clinical trial, the PIONEER Initiative – Precision Insights On N-of-1 Ex Vivo Effectiveness Research (NCT03896958), which aims to provide precision medicine to any patient with any cancer in any medical facility, and to enable functional

testing of the patient's tumour using organoids to identify the best treatment for the patient.

A number of proof-of-concept clinical trials are underway to assess if patient derived PDAC organoids can be developed from fine-needle biopsies of the patient's tumour (Clinicaltrials.gov: NCT03544255; NCT03990675; NCT03140592). One of these clinical trials (NCT03544255) will include drug screening of the organoids established in the study. These clinical trials will set the foundation for the use of organoids in the personalised treatment of pancreatic cancer.

1.8.4.3. Organoids as models of tumour microenvironment

The tumour microenvironment is known to play an important role in PDAC. By nature, PDAC tumours are dense, fibrotic and hypoxic, and combined with the suppression of tumour infiltrating lymphocytes (TILs) by cytokines such as TGF β and interleukin-10 (IL-10), PDAC is a non-immunogenic tumour [237].

To identify the role of the tumour microenvironment in PDAC, Tsai *et al.* [238] produced a new, patient matched organoid model containing primary PDAC organoids, stromal and immune components. The co-culture of the organoids with cancer-associated fibroblasts resulted in an increased IC₅₀ of 3.8 μ M compared to 1.8 μ M for organoids alone in response to treatment with gemcitabine. This group also described a method for the introduction of lymphocytes into the organoid culture by adding 500,000 CD3⁺ T lymphocytes per well suspended in 500 μ L organoid growth medium. They demonstrated that lymphocytes only infiltrated into Matrigel[®] domes containing organoids. The incorporation of lymphocytes into an organoid co-culture would allow for the use of these models in the study of immunotherapies. The development of methods to study the immune system in PDAC may help overcome the disappointing attempts to use immunotherapy in the treatment of this devastating disease. The use of immunotherapy has resulted in increased survival rates in solid tumour cancers such as melanoma, non-small cell lung cancer, and gastric cancers [239]. However, a successful immunotherapy for PDAC has yet to be developed.

A study by Öhlund *et al.* [240] used organoids to identify the role of pancreatic stellate cells and CAFs in PDAC tumour microenvironment and tumour progression. CAFs are derived from activated stellate cells and produce desmoplastic stroma, resulting in differences in disease progression and response to therapies. Co-cultures of organoids and CAFs were established, which resulted in activation of the CAFs to make desmoplastic stroma. These findings were also validated in human and mouse tissues.

1.8.4.4. Organoids for biomarker discovery

As well as using PDAC organoids for modelling *in vivo* drug response, they have also been used for discovering clinically actionable biomarkers. Huang *et al.* [241] demonstrated that organoid models recapitulate the glycomics and drug responses observed in PDX models. The basis of this study was for the identification of N-glycans enriched in the glycome of PDAC using organoids. They identified a set of 57 N-glycans represent 50–94% of the relative abundance of all N-glycans detected. They have also developed a method to use organoids as a discovery platform for blood-based biomarkers in PDAC patients, which can be used for the identification of secreted extracellular vesicles in the blood of patients. This method used 4.0 mL of organoid media supernatant and subjected to liquid chromatography–mass spectrometry/mass spectrometry (LC-MS/MS) and identified 241 proteins that were at least two-fold higher in tumour organoid extracellular vesicles compared to exocrine organoids and expressed in at least four out of the six tumour organoid lines.

1.8.4.5. Modelling human diseases with CRISPR-Cas9-modified organoids

The discovery of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated (Cas) proteins has revolutionised gene editing. These techniques are readily used in germline gene editing for research *in vitro* in cell lines, and *in vivo* in zebrafish, mice, pigs and primates, and recently have been used in organoid technology [242–245]. For a more detailed review of CRISPR/Cas9 genome editing in organoids, refer to the review by Driehuis and Clevers [246].

Lee *et al.* [247] edited *KRAS*^{G12V} and *ERBB2*, and inactivated *TP53*, *CDKN2A*, and *SMAD4* by lentiviral delivery of CRISPR-Cas9 into pancreatic ductal organoids. Non-transformed or *KRAS*-mutated ductal organoids ceased proliferation after several passages. *KRAS*, *TP53*, *CDKN2A* and *SMAD4*, and *KRAS*, *TP53*, *CDKN2A*, *SMAD4*, and *ERBB2* CRISPR edited ductal organoids propagated exponentially at least for 4 months. Upon orthotopic xenotransplantation to immunodeficient mice, these ductal organoids developed lesions resembling PanINs, but not PDAC. Seino *et al.* [248] used a similar CRISPR-Cas9 based method to show the stepwise tumorigenesis of PDAC, and a loss of niche stem cell factor dependence during tumour progression.

1.8.4.6. Advantages/limitations of organoids

The future of organoids in the treatment of PDAC includes their use in personalised medicine such as: next-generation sequencing of the tumour and using organoids for the screening of therapeutics for the identification of the best therapy for patients [249]. Organoids produce an unlimited supply of material for study, thus reducing the need for animal studies, and helping with the implementation of Article 4 of EU Directive 2010/63/EU, which describes the requirements of the 3Rs (Replacement, Reduction and Refinement) that aims to improve the welfare of animals in research [250]. Organoids are derived from stem cells and can form many different cell types and contain a much more realistic mixture of cells for *in vitro* testing. However, there are several limitations to organoids, including the difficulty in obtaining patient samples. Culturing organoids requires specialist training, and is laborious, and expensive. The use of a 3D matrix environment requires specialist approaches for sample handling, manipulation and functional assays. Finally, in order to analyse in organoid structures, novel imaging and quantitative analysis techniques must be implemented. A comparison of the advantages and disadvantages of organoids, and other PDAC models are outlined in Table 1.5.

Table 1.5: An overview of the advantages and disadvantages of PDAC models. Taken without permission from Nelson and Walsh [72].

Model	Represents Patient Sample?	Usage	Maintenance	Success Rates/ Growth Rate	Cost
Established Cell Lines	Homogenic [190]; Undergo genetic modifications[188]; Fail to recapitulate the physiology of tumours	High throughput testing	Low maintenance	Fast-growing; Commercially available	Low cost
Primary Cell Cultures	Heterogenous; Early passage number; Representative of original tumour [198]	High throughput testing	Low maintenance; Only grow for a limited number of passages [198]	Some commercially available lines; Difficult to establish [199]	Low cost
Organ-on-chip	Heterogenous [251]; Allows for the study of the interactions of multiple cell/organ types [252]	Low throughput testing	Medium maintenance	High success rates	Chips are expensive; High usage of media and drugs
Organoids	Heterogenous [219]; Tumours retain heterogeneity, genetics, and histological characteristics [234]	High throughput testing	Medium maintenance	Medium growing rate; High success rates; Established from small volumes of tumour [224]	Expensive ECM and media
PDX ¹	Tumours retain heterogeneity, genetics, and histological characteristics [200]; Replacement of human stroma with murine stroma [205]; Orthotopic tumours in correct microenvironment [201]	<i>In vivo</i> and <i>ex vivo</i> drug testing	High maintenance Requires specialist training, and multiple licences [203]	Slow-growing (up to 16 weeks); Medium success rates	Expensive to maintain
GEMM ²	Tumours in correct microenvironment [253]; Immune cells present [254]	<i>In vivo</i> and <i>ex vivo</i> drug testing; Testing of immune targeted therapies	High maintenance Requires specialist training, and multiple licences	High success rates; Slow-growing (up to 16 weeks)	Commercially available

¹PDX – patient derived xenograft; ²GEMM – genetically engineered mouse model.

1.9. Methods used in assessing functionality of genes in PDAC

A number of methods were used to assess the functionality of genes and SNPs in PDAC, including CRISPR, CUT&RUN (Cleavage Under Targets and Release Using Nuclease) and dual luciferase reporter assays. The background and use of these methods are outlined below.

1.9.1. CRISPR

CRISPR/Cas are RNA-mediated adaptive immune systems which have been evolved by bacteria and archaea to protect the organism against invading plasmids and phages [255]. In 2012, Jinek *et al.* [256] discovered that type II CRISPR systems could be used to induce targeted double-stranded breaks (DSB) in DNA. This discovery marked the beginning of the “CRISPR craze” [257]. Genomic engineering of cell lines and animal models has traditionally been accomplished through random mutagenesis or low-efficiency gene targeting. The ability to edit the genome in a precise and targeted manner is critical for understanding genetic contributions to disease. While other systems such as zinc finger nucleases (ZFNs) and transcription activator like effector nucleases (TALENs) have the ability to edit targeted sequences, the process that is required to make these nucleases is both time consuming and expensive [258].

CRISPR sequences are found in approximately 40% of bacterial and 90% of archaeal genomes. CRISPR was originally discovered in *Escherichia coli* in 1987 [259]. However, it was not until 2002 when Schouls *et al.* [260], who were decoding bacterial genomes found patterns in which a palindromic sequence was followed by approximately 30 bases of spacer DNA, then a repeat of the palindromic sequence, followed by a different sequence of spacer DNA. These spacers often originated from viral phage or plasmid DNA [261]. The CRISPR/Cas system works by incorporating fragments of the invading nucleic acid as spacers into the host genome. These spacers are later used as templates to transcribe small CRISPR-RNA (crRNA) molecules that are combined with the Cas enzymes into an effector complex which silences foreign nucleic acids during subsequent rounds of

infection [262]. CRISPR/Cas systems are distinct and have been classed into three main categories; the type I and II systems both have specialised Cas endonucleases which cleave the pre-crRNA. Once processed, each crRNA forms to make a large, multi-Cas protein complex. This complex is able to recognise and cleave nucleic acids complementary to the crRNA [256].

This system has been modified to allow for CRISPR-Cas9 mediated editing *in vivo* and *in vitro* to make knockout or specific gene edits in laboratory models. A single guide RNA (sgRNA) is engineered to specifically target a genomic region of interest. This sgRNA is 17-22 nucleotides in length and must be followed by a protospacer adjacent motif (PAM) sequence 5'-NGG-3'. The Cas9 endonuclease cleaves the DNA to induce a DSB, which is repaired through non-homologous end joining (NHEJ) resulting in an insertion or deletion (indel) or homology directed repair (HDR) in the presence of a donor sequence (Figure 1.16) [263].

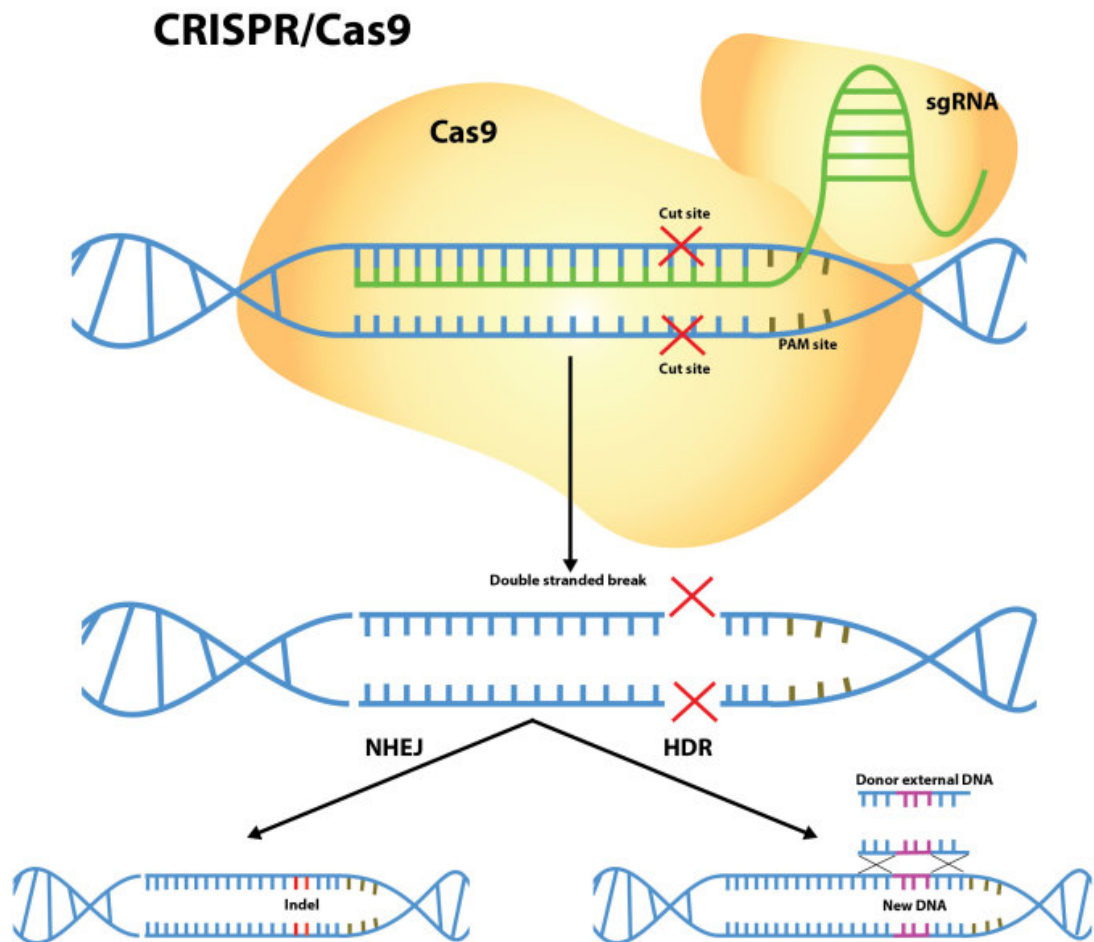


Figure 1.16: CRISPR-Cas9 mediated gene editing mechanisms. A single guide RNA (sgRNA) recruits the Cas9 DNA endonuclease to its complementary DNA sequence. The endonuclease induces a double-stranded break that is repaired by non-homologous end joining (NHEJ) resulting in (insertions/deletions) indels, or homology directed repair (HDR) in the presence of a donor construct, resulting in gene editing. This figure was taken without permission from Cribbs et al. [263].

1.9.2. CUT&RUN – Cleavage Under Targets and Release Using Nuclease

Transcription factors (TFs) are proteins that decode the information in the genome to express a precise and unique group of mRNAs in each different cell type. Knowing where TFs bind to the genome is essential to understanding their function. There are a number of methods used for the mapping of protein-DNA

interactions, including ChIP-seq (Chromatin immunoprecipitation with massively parallel sequencing), CUT&RUN, ChIC (Chromatin immuno-cleavage), DamID (DNA adenine methyltransferase identification) and ChEC (Chromatin endogenous cleavage).

ChIP-seq has been used extensively for epigenome and TF profiling [264]. Crosslinking using formaldehyde is performed to fix the protein-DNA interactions which allows the capture of weaker/transient interactions. These are non-specifically sheared by sonication. The proteins of interest are immunoprecipitated with an antibody coupled to a magnetic or agarose bead [265]. The DNA is purified and sequenced. ChIP-seq requires large amounts of input material to produce strong signal over noise, is expensive, labour intensive, and prone to artefacts [266]. While ChIP-seq has been the mainstay of studying DNA-protein interactions for over thirty years, novel methods have been developed which increase specificity and decrease background (Figure 1.17).

DamID identifies DNA-binding sites by expressing the DNA-binding protein as a fusion with DNA methyltransferase [267]. The binding of this fusion protein to the region of interest results in a methyltransferase of a “GATC” sequence in the region. Adenosine methylation does not occur in eukaryotes which allows the mapping of the binding sites through digestion of unmethylated DNA, and sequencing. In ChEC, pA-MNase is fused to a protein of interest [268].

CUT&RUN uses antibody-targeted controlled cleavage by protein-A micrococcal nuclease (pA-MNase), which is recruited by the antibody targeting protein of interest and activated by adding calcium (Figure 1.18). It introduces double-stranded breaks into the DNA and the DNA fragments are released into the supernatant, and sequenced [269]. It is performed *in situ*, allowing for high resolution mapping, and does not require crosslinking, which results in a low background and reduced sequencing costs due to a lower sequencing depth. The CUT&RUN method was developed from the ChIC method, which had a resolution of 100-200 base pairs [268].

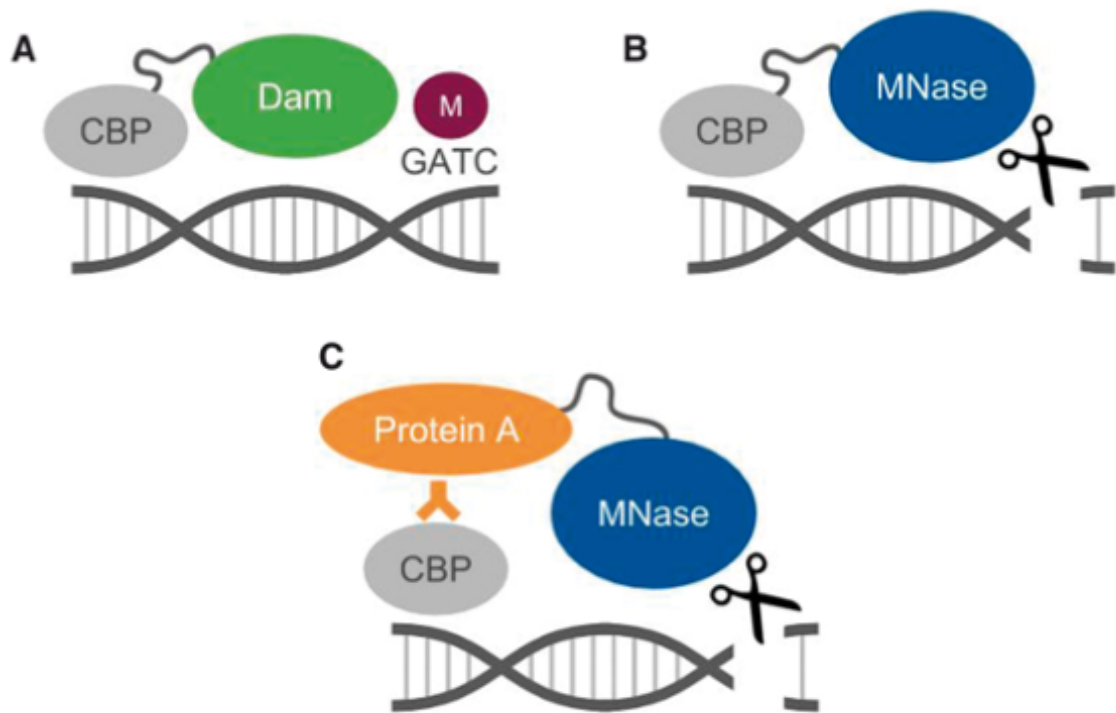


Figure 1.17: The principles of mapping chromatin-binding protein (CBP)-DNA interactions. (A) DamID: CBP fused to Dam resulting in adenine methylation (M) within GATC sequences (B) ChEC: CBP fused to micrococcal nuclease (MNase), and is activated by the addition of calcium, resulting in DNA cleavage (C) CUT&RUN/ChIC: cells are incubated with antibody against CBP of interest. Protein-A is fused to MNase, which binds the antibody and the addition of calcium results in cleavage of DNA. Figure modified without permission from Policastro et al. [270].

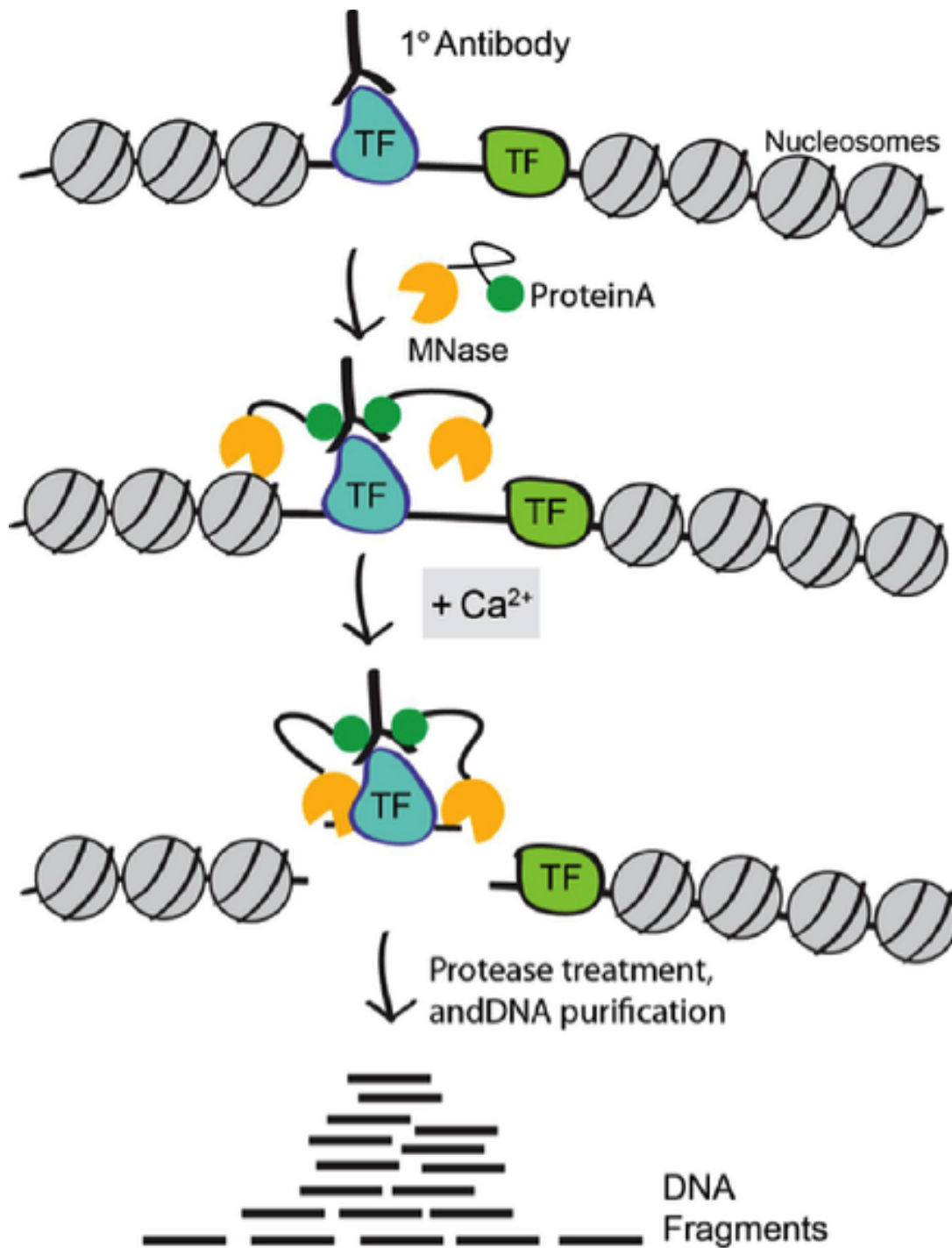


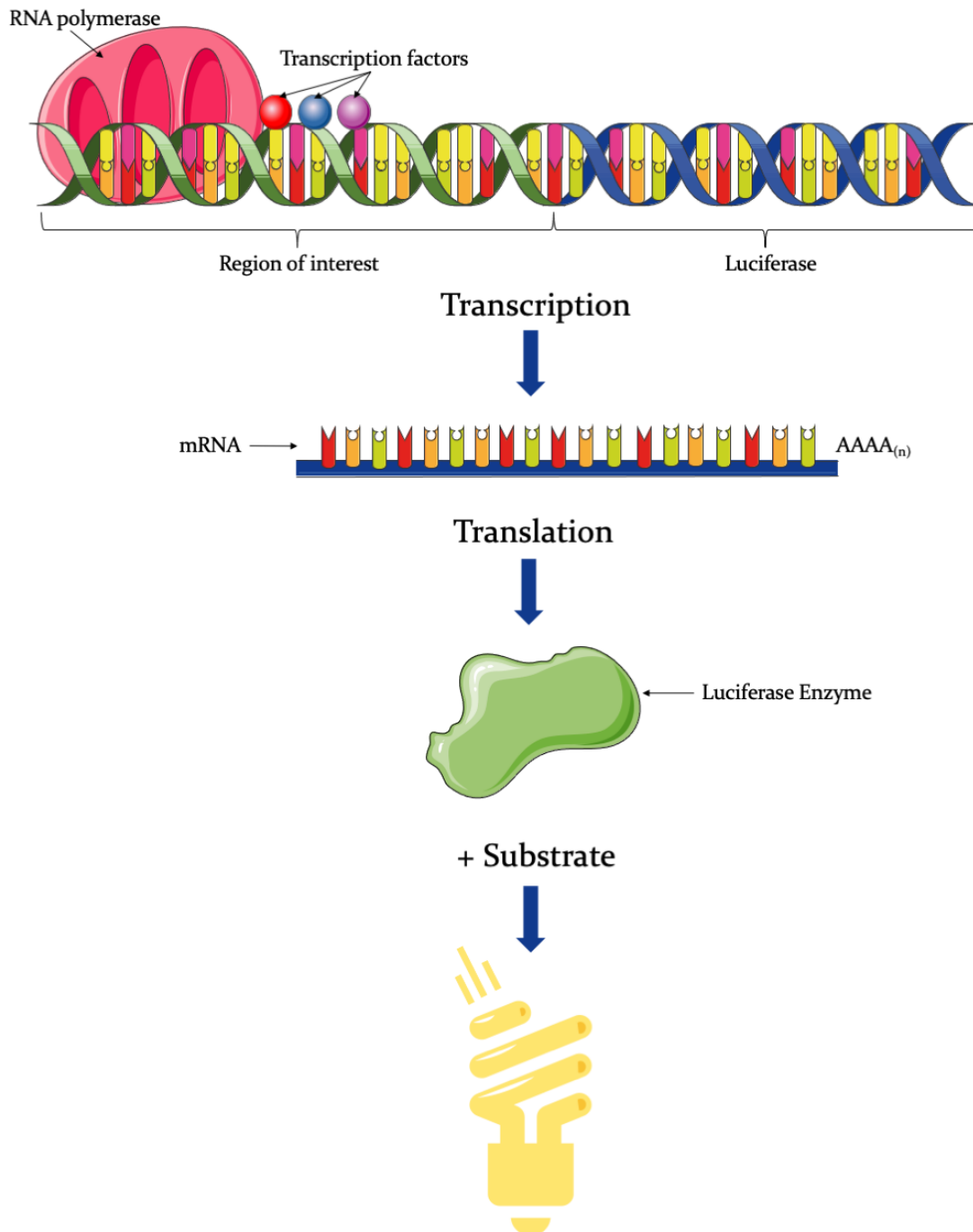
Figure 1.18: CUT&RUN schematic illustrating antibody binding to transcription factors (TF), Protein-A and Protein G IgG binding domains fused to micrococcal nuclease (pAG-MNase) activation by calcium chloride (Ca^{2+}) and cleavage of DNA.

Taken without permission from Addgene.com [271].

1.9.3. Dual luciferase reporter assays

Dual luciferase reporter assays are a system used to assess gene expression. This assay depends on two different reporter genes – Renilla (*Renilla reniformis*) luciferase and Firefly (*Photinus pyralis*) luciferase. Firefly luciferase is fused to the test promoter or region of interest to assess if the gene is translated, or if mutations result in differential gene expression (Figure 1.19). Renilla luciferase is fused to a constitutive promoter and acts as an internal control for transfection efficiency and transcriptional activity is used for normalising the assay [272]. Firefly luciferase and Renilla luciferase have distinct enzyme structures and substrate requirements, so the dual luciferase reporter assay is able to selectively discriminate the two. Firefly luciferase is a 61 kDa protein which does not require post translational processing for enzymatic activity [273,274]. Photon emission occurs via oxidation of beetle luciferin and requires the addition of ATP, Mg^{2+} and O_2 [272]. Renilla luciferase is a 36 kDa monomeric protein, and also does not require post translational modifications, the luminescence generated by Renilla luciferase utilises O_2 and coelenterazine [275]. In a dual luciferase reporter assay, a p2Luc (plasmid containing Firefly luciferase) is produced to contain the region of interest upstream of the Firefly luciferase gene. p2Luc and Renilla plasmids are transfected in to the HEK293 cell line, and after 24 hours substrates for Firefly luciferase are added, Firefly luminescence is quantified and quenched, prior to the addition of substrates for Renilla, and quantification of Renilla luminescence.

Potential issues with this assay include the loss of Firefly luciferase activity. This can occur if a start codon is present in upstream open reading frame, resulting in a 5' extension of the protein, or out of frame translation of the RNA. This can be overcome by the use of a modified variant of the p2Luc plasmid encoding the porcine enterovirus StopGo sequence to the start of the luciferase resulting in skipping of a peptide bond so the luciferase amino acid sequence will remain the same [276].



Luminesce = Luciferase Expression = Promoter Activity

Figure 1.19: Schematic of a luciferase reporter assay, the region of interest is fused to the 5'UTR of the luciferase gene and transfected in to HEK293 cells. If the region of interest allows for translation, the luciferase enzyme is produced by the cells and the addition of the substrate results in luminesce.

1.10. Aims

The overall aim of this thesis was to develop novel biologically relevant models of pancreatic cancer and to functionally validate the role of genetic variants and genes identified from a pathway based GWAS analysis in pancreatic cancer development.

The first aim of this thesis was to develop a more representative model of PDAC for pre-clinical research. This aim was achieved by the development and validation of a novel method for the development of organoids from a PDX PDAC sample. In addition to this, a novel method for the simultaneous generation of isogenic matched cell lines was developed, as well as a method for the development of organotypic models generated from primary cell lines, which was termed cell line organoids – CLOs. The second aim of this thesis was to validate these novel CLO models as representative models of PDAC, this was achieved by comparing phenotypes, drug response, morphology, tumorigenicity and transcriptomics with matched organoid and cell line models

The third aim of this thesis was to functionally validate genomic variants of PDAC identified in pathway analysis studies. This was achieved by prioritisation of SNPs from the MODY and Pujana ATM PCC Network pathways through *in silico* assessment. The effects of the selected SNPs on gene expression were assessed through dual luciferase reporter assays.

The fourth aim of this thesis was to assess the role of genes from the MODY in PDAC. The role of the genes was analysed through publicly available datasets and RNA-seq data. A CUT&RUN assay was performed to identify the downstream targets of MODY pathway transcription factors *HNFI*A and *HNFI*B.

The final aim of this thesis was to assess the effectiveness of targeting the DDR pathway in non-*BRCA* mutated PDAC for the treatment of the disease. This was performed by testing the potential of a novel the RAD51-*BRCA*2 small molecule disruptor (ARN24089) in organoid models using drug toxicity assays and evaluating the mechanism of action by IHC of RAD51 and assessing levels caspase3/7 cleavage.

[This page is intentionally left blank]

Chapter 2. Methods

2.1. Cell culture

2.1.1. Cell lines

All cell culture was carried out with strict adherence to the National Institute of Cellular Biotechnology (NICB) cell culture standard operating procedures (SOPs). All lines are mycoplasma free – mycoplasma testing was carried out in-house every 3 months by Justine Meiller. STR fingerprinting was carried out by IDEXX BioResearch.

2.1.2. Cell lines and culture conditions

All cell lines were incubated in 5% CO₂ at 37°C. Cell lines used are represented in Table 2.1. Cell culture media used were Dulbecco's Modified Eagle Medium (DMEM) (Sigma, D8437), Roswell Park Memorial Institute-1640 medium (RPMI-1640) (Sigma, R8758), keratinocyte-serum free medium (SFM) (kit) with L-glutamine, epidermal growth factor, and bovine pituitary extract (Gibco, 17005075), DMEM, high glucose with GlutaMax (Gibco, 31966021), foetal bovine serum (FBS) (Gibco, 10270106), L-glutamine (Gibco, 35050038) and sodium pyruvate (Gibco, 11360070). Cell lines purchased include L-WRN (ATCC, 3276), and H6c7 (KeraFast, ECA001). Specific media requirements for each cell line are outlined in Table 2.2.

Table 2.1: Cell lines used in this project, the histology, and source.

Cell Line	Histology	Source
ASPC1	PDAC ¹ Metastasis in Ascites	ATCC ²
BxPC3	PDAC	ATCC
CAPAN1	PDAC Metastasis in liver	ATCC
CAPAN2	PDAC	DSMZ ³
HEK293	Transformed Human Embryonic Kidney cells	ATCC
HPAC	PDAC	DSMZ
HuPT4	PDAC Metastasis in Ascites	DSMZ
H6c7	Human Pancreatic Duct Epithelial Cell Line	KeraFast
L-WRN	L-Wnt3A cell line transfected with an R-spondin3 and noggin co-expressing vector	ATCC
MiaPaCa-2	PDAC	EACACC
PANCI	PDAC	ATCC
PDM37-CL	Cell line derived from PDM37 organoid	NICB
PDM41-CL	Cell line derived from PDM41 organoid	NICB
PDM106-CL	Cell line derived from PDM106 organoid	NICB
PT127-CL	Invasive moderately differentiated adenocarcinoma	NICB
PT291-CL	Cell line derived from PT291 organoid	NICB
SW1990	Spleen metastasis of a grade II PDAC	ATCC

¹PDAC: Pancreatic ductal adenocarcinoma; ²ATCC: American Type Culture Collection; ³DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (German Collection of Microorganisms and Cell Cultures GmbH).

Table 2.2: Media requirements for different cell lines used in this project.

Cell Line	Media
ASPC1	RPMI-1640 ¹ with 5% FBS ²
BxPC3	RPMI-1640 with 5% FBS
CAPAN1	RPMI-1640 with 10% FBS
CAPAN2	RPMI-1640 with 10% FBS
HEK293T	DMEM ³ with 5% FBS and 2 mM L-GLUT ⁴
HPAC	RPMI-1640 with 5% FBS
HuPT4	DMEM with 10% FBS and 2 mM L-GLUT
H6c7	Keratinocyte-SFM ⁵ medium (Kit) with L-GLUT, ⁶ EGF, and ⁷ BPE
L-WRN	DMEM-GlutaMax with 10% FBS and 1 mM sodium pyruvate
MiaPaCa2	DMEM with 5% FBS and 2 mM L-GLUT
PANCI	DMEM with 5% FBS and 2 mM L-GLUT
PDM37-CL	DMEM with 10% FBS and 2 mM L-GLUT
PDM41-CL	DMEM with 10% FBS and 2 mM L-GLUT
PDM106-CL	DMEM with 10% FBS and 2 mM L-GLUT
PT127-CL	DMEM with 10% FBS and 2 mM L-GLUT
PT291-CL	DMEM with 10% FBS and 2 mM L-GLUT
SWI990	RPMI-1640 with 10% FBS

¹RPMI-1640: Roswell Park Memorial Institute; ²FBS: Foetal bovine serum; ³DMEM: Dulbecco's Modified Eagle's Medium; ⁴L-GLUT: L-glutamine; ⁵SFM: Serum free media; ⁶EGF: Epidermal growth factor, and ⁷BPE: Bovine pituitary extract.

2.2. Wnt3a conditioned media

2.2.1. Conditioned media collection

L-WRN cells (ATCC, 3276) were thawed without centrifugation into a T75 flask. The following day, cells were treated with 0.5 mg/mL G-418 (Gibco, 10131035) and 0.5 mg/mL hygromycin B (Merck, 10843555001) for 3 days. When cells reached 80% confluency, cells were split into ten T175 flasks, and media was added as required. Cells were grown to over-confluency before conditioned media was collected. Media was removed, and cells were washed with 15 mL warm phosphate buffered saline (PBS) and PBS was discarded. 25 mL of complete media was added, and cells were incubated at 37°C for 24 hours. Media was collected, centrifuged at 1000 RPM for 5 minutes and stored at 4°C; wash and media addition steps were repeated. This process occurred over 4 days. The collected media was mixed and filtered, and frozen in 25 mL aliquots. A TOPFlash assay (dual luciferase reporter assay) was carried out to assess the concentration of Wnt3a and Rspondin3.

2.2.2. TOPFlash assay

2.2.2.1. Plasmid growth and DNA extraction

LB agar plates with 100 µg/mL ampicillin were prepared. Plasmids TOPFlash (Addgene, 12456) and Renilla (Addgene, 87068) were streaked onto plates. Plates were incubated, upside-down at 37°C for 18 hours, then inspected for single colonies. A single colony was picked and placed into 5 mL LB broth with 100 µg/mL ampicillin in a 50 mL tube. Tubes were placed in a shaking incubator at 37°C overnight, with the lids slightly open. DNA extraction was carried out using a PureLink Quick plasmid Miniprep Kit (Invitrogen, K210011) and DNA was quantified using a Nanodrop, diluted to 100 ng/µL for TOPFlash, and 10 ng/µL for Renilla and aliquoted.

2.2.2.2. Dual luciferase reporter assay

Day 1: Seeding - HEK293T cells were grown to 70% confluency, and a 12 mL cell suspension of 2.5×10^4 cells/mL was made. Cells were seeded at 100 µL per well in a black-sided 96-well plate (Corning, CLS3603). Plate was placed overnight in a 37°C, 5% CO₂ incubator.

Day 2: Transfection – 100 ng DNA (10:1 TOPFlash:Renilla) was diluted in 25 μ L Opti-MEM (Gibco, 11058021) per well of cells to be transfected. 0.5 μ L Lipofectamine 2000 (Invitrogen, 11668027) was diluted in 25 μ L Opti-MEM (Gibco, 11058021) per well to be transfected. These were incubated at room temperature for 5 minutes. DNA and Lipofectamine 2000 mixtures were combined, mixed, and incubated at room temperature for 20 minutes. 50 μ L of the transfection mixture was added to each well to be transfected with Lipofectamine 2000 only, and DNA only used as negative controls. The plate was incubated at 37°C, 5% CO₂ overnight.

Day 3: Feeding - conditioned medium was diluted 1:1 with Complete GlutaMAX DMEM (Gibco, 61965026). Media was removed from transfected cells and 100 μ L of the diluted conditioned media was added. Complete GlutaMAX DMEM with 250 ng/mL Wnt3a and 500 ng/mL Rspondin3 was used as positive control.

Day 4: Dual luciferase reporter assay – the Stop and Glow (Promega, E2920) Dual Luciferase Reporter Assay System was used. Reagents were prepared according to protocols provided. The 96 well plate was removed from the incubator and equilibrated to room temperature. Dual Glo Substrate (100 μ L) was added to each and the plate was incubated at room temperature for 10 minutes. The plate was read using a plate reader (Biotek with Gen4 software) with a gain of 100, and integration time of 1 second to determine Firefly luciferase (TOPFlash) luminescence. Stop and Glo reagent (100 μ L) was added, and the plate was incubated at room temperature for 10 minutes. The plate was read using the same settings to determine Renilla luciferase luminescence. Relative luminescence was calculated by normalising Firefly to Renilla for each well.

2.3. PDAC organoid culture

2.3.1. Organoids & CLOs

All organoids and CLOs were incubated in 5% CO₂ at 37°C. Cell lines used are represented in Table 2.3. Organoids were established directly from tumour samples, whereas CLOs were established from primary cell lines.

Table 2.3. Organoids and CLOs used in this project, the histology, media required and source.

Organoid Line	Histology	Media	Source
PDM37	Squamous cell carcinoma	CHFM ¹	ATCC ²
PDM37-CLO ³	CLO derived from PDM37-CL ⁴	CHFM	NICB ⁵
PDM41	Adenocarcinoma ductal type	CHFM	ATCC ⁶
PDM41-CLO	CLO derived from PDM41-CL	CHFM	NICB
PDM106	PDAC ⁷ metastasis in liver	CHFM	ATCC
PDM106-CLO	CLO derived from PDM106-CL	CHFM	NICB
PT127-CLO	Invasive moderately differentiated adenocarcinoma	CHFM	NICB
PT291	Invasive adenocarcinoma / cholangiocarcinoma	CHFM	NICB
PT291-CLO	CLO derived from PT127-CL	CHFM	NICB

¹CHFM: Complete human feeding media; ²ATCC: American Type Culture Collection; ³CLO: Cell line organoid; ⁴CL: cell line; ⁵NICB: National Institute for Cellular Biotechnology; ⁶ATCC: American Type Culture Collection; ⁷PDAC: Pancreatic ductal adenocarcinoma.

2.3.2. Media and plate preparation

A number of different media are required for organoid culture, including basic media (Table 2.4), wash media (Table 2.5), and complete human feeding media (Table 2.6).

Table 2.4: Composition of basic media, including reagents, source, stock concentration, volume and final concentration.

Reagent	Source, Product Code	Stock Concentration	Volume	Final Concentration
DMEM-F12	Merck, D6421	—	98.75 mL	—
HEPES Buffer	Merck, H3375	1 M	1 mL	10 mM
Antibiotic-Antimycotic	ThermoFisher, I5240062	100X	0.25 mL	1X

Table 2.5: Composition of wash media, including reagents, source, stock concentration, volume and final concentration.

Reagent	Source, Product Code	Stock Concentration	Volume	Final Concentration
DMEM-F12	Merck, D6421	—	94.5 mL	—
HEPES Buffer	Merck, H3375	100X (1 M)	1 mL	1X
GlutaMAX	Gibco, 35050038	100X	1 mL	1X
Antibiotic-Antimycotic	ThermoFisher, I5240062	100X	1 mL	1X
FBS	(Gibco, 10270106)	100%	2.5 mL	2.5% (v/v)

Table 2.6: Composition of Complete Human Feeding Media (CHFM), including reagents, source, stock concentration, volume and final concentration.

Reagent	Source, Product Code	Stock Concentration	Volume	Final Concentration
Basic Media	—	—	23.125 mL	—
Conditioned Media	—	—	25 mL	50% v/v
A83-01	Sigma, SML0788	1000X (0.5 mM)	50 µL	500 nM
hFGF10 ¹	Biolegend, 559308	1000X (100 µg/mL)	50 µL	100 ng/mL
EGF ²	Gibco, PHG0311	1000X (50 µg/mL)	50 µL	50 ng/mL
Gastrin I	TOCRIS, 3006	1000X (10 µM)	50 µL	0.01 µM
N-acetylcysteine	R&D, 5619	400X (500 mM)	125 µL	1.25 mM
Nicotinamide	Sigma, N0636	100X (1M)	500 µL	10 mM
B-27 Supplement	Life Technologies, 17504-044	50X	1000 µL	1X
Y-27632 ³	Sigma, Y0503	1000X (10.5 mM)	50 µL	10.5 µM

¹hFGF10: Human fibroblast growth factor 10; ²EGF: Epidermal growth factor; ³Y-27632 (Rho Kinase Inhibitor) is only necessary when organoids are first prepared, when organoids are thawed, or when organoids are dissociated to single cell.

2.3.2.1. PolyHEMA-coated plates

Poly-2-hydroxyethyl methacrylate (polyHEMA) is a soft, flexible, water-absorbing plastic, which is used for coating cell culture plates to prevent cells attaching. A 1% solution of polyHEMA (Merck, P3932) was prepared by adding 12 g of polyHEMA to 1 L of 95% ethanol, overnight at 60°C with a stirrer in agitation. The plates were prepared in a sterile hood by adding 300 µL of the polyHEMA solution to each well of a 24-well plate. Plates were placed in a 50°C oven until the ethanol completely evaporated. The procedure was repeated, and the plates were stored at room temperature. The plates were preheated overnight to 37°C prior to use in culture.

2.3.3. PDAC PDX organoid culture establishment

PDX work was carried out by Sandra Roche, Fiona O'Neill, Neil Conlon and Justine Meiller.

2.3.3.1. Patient demographics and PDX information

PDAC primary tumour tissue specimens were obtained from Saint Vincent's University Hospital Ethics and Medical Research Committee, Dublin, Ireland between 2013 and 2017. All patients underwent PDAC resection. All donors agreed by written informed consent to donate tissue and for study participation. The specimen donors' personal information was confidential and protected except the date of birth. All samples and methods used in this study were approved by Saint Vincent's University Hospital (SVUH) Ethics and Medical Research Committee and conducted in accordance with the relevant guidelines and regulations in compliance with the SVUH Ethics and Medical Research Committee. PDAC PDX tissues were generated by subcutaneous seeding in SCID mice. Primary patient samples were confirmed as PDAC by a pathologist in SVUH, and PDX samples were also confirmed by pathology examination to maintain the human tumour content and morphology of the original tumour (Table 2.7).

Table 2.7: Patient information of cell lines and organoids established from PDAC samples.

<i>Patient</i>	<i>Gender</i>	<i>Age</i>	<i>Origin</i>
PT127	Male	73	Invasive moderately differentiated adenocarcinoma
PT291	Female	56	Invasive adenocarcinoma / cholangiocarcinoma

All animal work has received ethical approval from the DCU Research Ethics Committee (DCUREC/2012/202) and was licensed by the Department of Health (B100-4501). All methods were performed in accordance with the relevant guidelines and regulations in compliance with the DCU Research Ethics Committee and the Department of Health.

2.3.3.2. PDX tumour development

After initial macroscopic pathological confirmation, material remaining after diagnostic sampling was cold transferred in RPMI-1640 media containing 1% Penicillin-Streptomycin (Gibco, 15140148) and 1% Amphotericin B (Gibco, 15290018) to DCU.

The tumour was cut into implant sized pieces (<2 mm) and rinsed with fresh serum free RPMI-1640 media following transport. SCID – CBl7/lcr-Prkdc^{scid}/lcrCrl – mice (Charles River, UK) were implanted subcutaneously with fresh patient tumour material. Under anaesthesia (isoflourane, O₂ carrier gas) a small incision was made in the skin of the left flank of the animal. The tumour piece was placed in the pocket under the skin and the wound sealed with a single staple. The animals were monitored post-surgery, and staple removal was within 10 days. Animals were monitored weekly for body weight and tumour development. Mice were monitored for tumours development for up to 1 year post implantation. Animal welfare monitoring criteria included tumour volume, tumour axis, body weight and condition, (where tumour volume <2000 mm³ and tumour axis <20 mm). A decrease in body weight of >10% resulted in increased monitoring, with body weight decrease of 20% resulting in humane euthanasia. Tumour measurements

were by calliper measurement and tumour volume was calculated as outlined in Equation 1 below.

Equation 1. Calculation of tumour volume using tumour height, depth and width measurements.

$$\frac{\text{Height} \times \text{Depth} \times \text{Width}}{2}$$

At the end of the experiment, the animals were humanely euthanised, and the tumours were harvested. PDX samples were confirmed by pathology examination to maintain the human tumour content and morphology of the original tumour.

2.3.3.3. PDX tumour preparation

The media used for organoid establishment and culture are represented in Table 2.4 to Table 2.6. First, a 21 mL solution of 20:1 PBS to PB Buffer (Miltenyi Biotec, 130-091-376), and a 35 mL 2% antibiotic-antimycotic in PBS solution were prepared. The PDAC PDX tumour sample was received on ice, in basic media (Table 2.4). Using a Pasteur pipette, all media was removed and 10 mL of the PBS-Antibiotic-Antimycotic solution was added to the sample, and the tube was shaken for 20 seconds. The tumour sample was allowed to settle, and the PBS was aspirated. This process was repeated with 10 mL and 15 mL washes. In a petri-dish, the tumour was chopped finely using scalpel and forceps. The sample was placed in a 50 mL Falcon tube, and 10 mL TrypLE (Gibco, 12604013) was added. The tube was placed on a rocking table at room temperature for 15 minutes. The tumour sample was removed to a laminar flow hood. The digested tumour sample was put through a MACS SmartStrainers (70 µm) (Miltenyi Biotec, 130-098-462) using a cell scraper into a fresh 50 mL falcon tube. The cell strainer was rinsed with 20mL wash media (Table 2.5). The cells were centrifuged at 1500 RPM for 7 minutes at room temperature.

2.3.3.4. Mouse cell depletion kit

Following centrifugation, the media was aspirated, and the pellet was resuspended in 5 mL PB buffer, making a single cell suspension. Using trypan blue (Sigma, T8154O), the cells were counted and a cell suspension of 2×10^6 cells was made and centrifuged at 1500 RPM for 10 minutes. The cell pellet was resuspended in 80 μ L PB buffer, and 20 μ L of the mouse cell depletion kit antibody cocktail (Miltenyi Biotec, 130-104-694). This suspension was incubated at 4°C for 15 minutes. Following incubation, 400 μ L of PB buffer was added to the cell suspension and mixed thoroughly. An LS column (Miltenyi Biotec, 130-042-401) was placed in the MidiMacs Separator (Miltenyi Biotec, 130-042-302) and rinsed with 3 mL PB buffer. The cell suspension was pipetted into the column, and the flow-through was collected in a 15 mL tube. The column was washed twice with 1 mL PB buffer, and the flow-through was collected in the same tube. The cells were centrifuged at 1000 RPM for 5 minutes at 4°C. All media was removed, using a P200 to avoid disrupting the pellet.

2.3.3.5. Organoid establishment

The cells were resuspended in 50 μ L extracellular matrix (ECM) (Sigma, E1270, 8-12 mg/mL) per 1.5×10^5 cells. 20 μ L of organoid:12 mg/mL ECM solution was pipetted on to a poly-HEMA coated 24-well plate. The plate was placed in an incubator upside-down at 37°C for 20 minutes to allow the ECM to polymerise. 500 μ L of the Complete Human Feeding Media (CHFM) (Table 2.6) with Rho kinase inhibitor (ROCKi) was added to each well. Two days after establishing the organoid sample, cells were fed using CHFM with ROCKi, then subsequently fed every two days using CHFM without ROCKi.

2.3.3.6. Passaging organoids

The media and ECM from each well was removed to a 30 mL tube and centrifuged at 1500 RPM for 10 minutes at 4°C. Pellets were washed with 500 μ L of ice-cold PBS to remove remaining ECM and centrifuged at 1500 RPM for 10 minutes at 4°C. 3mL TrypLE (Gibco, 12605010) was added to the pellet and placed at 37°C for 15 minutes. The pellet disrupted using “hard pipetting” (using a P1000 pipette, the

organoids were pipetted quickly and vigorously). DMEM with 10% FBS was added to stop trypsinisation. The organoids were plated as outlined in Section 2.3.3.5.

2.3.3.7. Freezing organoids

Organoids were passaged as described in Section 2.3.3.6. and allowed to grow for 3 days. Four wells of organoids per vial of cells to be frozen were removed from the plate and centrifuged at 1500 RPM for 10 minutes. All media was removed from the cells, and 1 mL of Recovery Cell Culture Freezing media (ThermoFisher, 12648010) per vial of cells, was added in a dropwise manner, removed to an internal-thread cryovial. Cryovials were placed into a Mr Frosty Freezing Container (ThermoFisher, 5100-0001) and placed in a -80°C freezer for 48 hours, then moved to liquid nitrogen for long term storage.

2.3.4. Establishment of isogenic matched 2D primary cell line

Protocol was followed as outlined in Sections 2.3.3.3 and 2.3.3.4. Following the removal of infiltrating mouse cells, 1.5×10^5 cells were mixed in 100 μ L of ECM diluted to 1 mg/mL and 50 μ L was plated per well in a non-polyHEMA coated 24 well plate. Cells were fed 500 μ L of the CHFM with ROCKi. Then, cells were grown until the cells began to adhere to the bottom of the plate, feeding with CHFM. To passage cells, 1 mL TrypLE (Gibco, 12605010) was added to each well until the cells detached, and 1 mL of DMEM High Glucose GlutaMAX (Gibco, 10566-016) with 10% FBS was added to stop trypsinisation. Cells were transferred to 6-well plate, then up scaled to a T25 cm³ flask. After this, the media was changed from CHFM to 50:50 GlutaMax DMEM and L-WRN conditioned media with 10% FBS and 1% antibiotic-antimycotic.

2.3.5. CLO establishment

Using early passage 2D primary cell lines, cells were trypsinised, and a cell suspension of 5×10^4 cells per well to be seeded was made and centrifuged at 1500 RPM for 10 minutes at 4°C. Cells were resuspended in 50 μ L ECM per well and plated on to a 24-well polyHEMA coated plate and fed using CHFM with ROCKi, then subsequently fed every 2 days using CHFM without ROCKi. Two weeks after

establishment, cells were passaged as organoids, as outlined in Section 2.3.3.6. Cells were grown as organoids for a minimum of three passages before use in experiments.

2.3.6. Species confirmation PCR

DNA was extracted from organoids and primary cell lines using DNeasy Blood & Tissue Kit (Qiagen, 69504). PCRs were performed using MyTaq™ PCR Kit (Meridian Bioscience, BIO-21126). The first PCR is performed using a universal primer pair complementary to conserved sequences in cytochrome B and 16S rRNA genes. Each PCR reaction contained 100 ng DNA, 0.5 µL 10 pmol Universal Primers (Table 2.9), 12.5 µL MyTaq HS Red Mix 2x and water (ddH₂O) up to 25 µL. Amplification was carried out in G-Storm, at 94°C for 5 minutes, 59°C for 5 minutes, followed by 35 cycles of 72°C for 2.5 minutes, 94°C for 30 seconds, 59°C for 45 seconds, and followed by a final extension step of 72°C for 10 minutes. Samples were stored at 4°C.

Table 2.8: Primers for species confirmation PCR.

	Forward 5'-3'	Reverse 5'-3'
Universal	THGTHSAATGAATCTGAGGVGGVT	CGATGTTGGATCAGGACATC
Mouse	GCACTGAAAATGCTTAGATGGATAATTG	CCTCTCATAAACGGATGTCTAG

The amplified product from the first PCR was diluted 1:10 with ddH₂O. 1 µL of diluted PCR product, 0.5 µL 10 pmol mouse primers (Table 2.8), 12.5µL MyTaq HS Red Mix 2X and ddH₂O up to 25µL were mixed. Amplification was carried out in G-Storm, 94°C for 5 minutes, 60°C for 5 minutes, followed by 30 cycles of 72°C for 1.5 minutes, 94°C for 30 seconds, 60°C for 30 seconds, followed by a final extension step of 72°C for 5 minutes. Samples were stored at 4°C. 5 µL of each PCR product was run on a 1% agarose (Sigma, A9539) gel, stained with SafeView (NBS Biologicals, NBS-SV1), visualised under UV light, and imaged. A low molecular weight DNA Ladder (ThermoFisher, 15628019) was applied as a size marker.

2.4. Mutation analysis

DNA was extracted from organoids using DNeasy Blood & Tissue Kit (Qiagen, 69504). PCRs were performed using the ThermoFisher DreamTaq PCR Master Mix (2X). Primers complementary to regions surrounding the mutations of interest were designed for these PCRs.

The reaction was set up by mixing 25 ng DNA, 1 μ L 10 pmol Primers (Table 2.9), 12.5 μ L PCR Master Mix (Thermo Fisher, EP0712), nuclease free water up to 25 μ L. Amplification was carried out in G-Storm, at 95°C for 3 minutes, followed by 35 cycles of 95°C for 30 seconds, 58-63°C for 30 seconds, 72°C for 1 minute and followed by a final extension step of 72°C for 10 minutes. Samples were stored at 4°C. 5 μ L of each sample was run on a 1% agarose gel with a low molecular weight gene ladder (ThermoFisher, SM1193) for 30 minutes to confirm the presence of PCR product. A PCR product clean-up was performed using GenElute™ PCR Clean-Up Kit (Sigma, NA1020). Products were sent for Sanger sequencing with Eurofins, with 2 μ L 10 pmol forward primer and 15 μ L of 5 ng/ μ L PCR product. Representative chromatogram shown in Appendix A (Supplementary Figure 8.1). Mutations were identified by comparison of DNA sequence to wildtype sequence, and annotated using Variobox [277].

Table 2.9: Primers for mutation analysis PCR.

Gene Region	Forward 5'-3'	Reverse 5'-3'	Temperature (°C)
CDKN2A Exon 1	GGGAGCAGCATGGAGCCG	AGTCGCCCGCCATCCCCCT	63°C
CDKN2A Exon 2A	CCTGGCTCTGACCAATTCTGT	GGGCAGCGTCGTGCACGGGT	54°C
CDKN2A Exon 2B	AACTGCGCCGACCCCGCCACTCTCA	CCAGCTCCTCAGCCAGGTCCACG	62°C
CDKN2A Exon 2C	CGATGCCCTGGGGCCGTCTGC	GTACAAAATTCTCAGATCATC	55°C
CDKN2A Exon 3	CGGTAGGGACGGCAAGAGAG	CCTGTAGGACCTTCGGTGACTGA	63°C
KRAS Codon 11-13	CGATACACGTCTGCAGTCAACT	CCTGACATACTCCCAAGGAAAG	58°C
KRAS Codon 61	CTTAGTGGCCATTGTCCGTCAT	GATGCAGTCTGGAGCAAGTT	58°C
TP53 Exon 4	ATCTACAGTCCCCCCTTGCCG	GCAACTGACCGTGCAAGTCA	63°C
TP53 Exon 6	GCCTCTGATTTCCTCACTGAT	TTAACCCCTCCTCCCAGAGA	63°C
TP53 Exon 7	CTTGCCACAGGTCTCCCCAA	AGGGGTCAGAGGCAAGCAGA	63°C
SMAD4 Exon 1	CCTGATAGGCCCATGGGTGAGT	GCTTGAAAGGAAACGTAGCAAGTT	56°C
SMAD4 Exon 2	TGGTAGGATTGTGAGGATTAAATCAG	CGCGGGCTATCTTCCAAAT	56°C
SMAD4 Exon 8	CCTTAACCAAAGTGTGCAGCTT	TTGTAGTCCACCATCCTGATAAGGT	56°C

2.5. Proliferation assays

2.5.1. Cell line proliferation assay

Cell lines were trypsinised, and 100 μL of a 1×10^4 - 5×10^4 cells/mL cell suspension was seeded in a 96 well plate. An additional 100 μL of complete media was added to each well, and the plate was placed in the Incucyte Live Cell Imaging System (S3, Sartorius) and imaged every 6 hours over a 10-day period.

2.5.2. Organoid and CLO cell proliferation assay

Organoids and CLOs were passaged and seeded in polyHEMA-coated 96 well plates at a cell density of 2.5×10^3 cells/well in 10 μL of ECM. Plates were placed in an incubator to allow the ECM to solidify. After 20 minutes, 190 μL of CHFM with ROCKi was added to each well. The plate was placed in the Incucyte Live Cell Imaging System and imaged every 6 hours over a 10-day period.

2.6. Toxicity assays

2.6.1. Cell line drug toxicity assays

Cell lines were trypsinised, and 100 μL of a 5×10^4 cells/mL cell suspension was seeded in a 96 well plate. After 24 hours, cells were treated with 2X concentration of gemcitabine in 100 μL of medium. After five days' incubation, proliferation was measured using the acid phosphatase assay. Medium was removed and cells were washed three times with PBS. 100 μL of acid phosphatase substrate (10 mM p-nitrophenol phosphate (Sigma) in 0.1 M sodium acetate (Sigma), 0.1% Triton X-100 (BDH, pH 5.5) was added to each well and the plate was incubated at 37°C for one hour. The reaction was halted by the addition of 50 μL of 0.1 M sodium hydroxide. The absorbance was then read at 405 nm and at 620 nm as a reference on a plate reader (Biotek) using Gen4 software. Percentage growth was calculated relative to an untreated control. All assays were performed in triplicate.

2.6.2. Organoid and CLO drug toxicity assay

Organoids and CLOs were passaged and seeded in polyHEMA-coated black-sided, clear bottom 96-well plate (Corning, 353219) at a cell density of 5×10^3 cells/well in

10 μ L of ECM. The plate was placed in an incubator to allow the ECM to solidify. After 20 minutes, 90 μ L of CHFM with ROCKi was added to each well. After three days, cells were treated with 2X concentration of drug in 100 μ L CHFM. After five days' incubation, proliferation was measured using Cell Titre Glo (Promega, G9682), as per manufacturer's instructions. The luminescence was then read with an integration of 1.0 s and gain of 100 on a plate reader (Biotek) using Gen4 software. Percentage growth was calculated relative to an untreated control. All assays were performed in triplicate.

2.6.3. Apoptosis assay

Organoids were seeded as outlined above. On day three, when preparing 2X drug solution, Caspase-3/7 Green Dye (Sartorius, 4440) was added at a concentration of 1:2000 (final assay concentration is 1.25 μ M). Cells were treated with 100 μ L 2X drug-Caspase-3/7 green dye. The plate was placed in the Incucyte (S3, Sartorius) and imaged using the phase contrast and green fluorescence channels every 2 hours over a 5-day period.

2.7. Staining and imaging

2.7.1. Brightfield imaging

Brightfield imaging was performed on days 0, 3, 7, and 10 to determine the morphological changes of the cells over this period. Pictures were taken using a Canon EOS 550D camera (Canon GmbH, Krefeld, Germany) through a Leica DM IL LED inverted microscope (Leica Microsystems, Wetzlar, Germany). Scale bars were added using ImageJ.

2.7.2. Haematoxylin and eosin stain

The sections were de-paraffinized by two five-minute xylene clearings. The slides were then rehydrated by two five-minute clearings in 100% ethanol, two minutes in 90% ethanol and two minutes in 70% ethanol. Slides were stained in haematoxylin (Sigma, MHS1) for 2 minutes. Slides were rinsed in running water for 5 minutes. They were then placed for 30 seconds in 1% acid alcohol (2 mL 1M HCL in 198 mL 70%EtOH), followed by running water for 1 minute. Slides were

placed in sodium bicarbonate solution (1 g in 1L dH₂O) for one minute, followed by 1 minute under running water. Slides were dipped 10 times into 95% ethanol, and stained with eosin (Sigma, E4009) for 1 minute. Slides were dehydrated using 70%, 90% and 100% ethanol for 2 three-minute washes, followed by xylenes for 2 five-minute washes. The coverslips were mounted using DPX (Sigma, 44581) and the slides were allowed to air dry overnight in the fume hood.

2.7.3. Immunofluorescence

2.7.3.1. Slide preparation

Cell lines

A cell suspension of 5×10^3 in 500 μ L complete media was seeded in a glass bottomed 8-well plate (Ibidi, 80827), and grown overnight. Media was removed from the cells and washed three times in a PBS 0.1% Tween 20 (Sigma, P1379) and 2% BSA (Sigma, A9418) (PTB) solution. Cells were fixed in ice-cold methanol for 5 minutes at -20°C and washed in the PTB solution.

Organoids and CLOs

Organoids seeded in a 12-well plate and grown for one week. Four wells of organoids were combined and centrifuged at 1500 RPM for 10 minutes in a 1.5 mL Eppendorf tube. All media was removed from the cell pellet, and cells were fixed in formalin. In samples with low number of organoids, iPGell (Funakoshi, PG20-1) was used according to the manufacturer's protocol. Samples with higher numbers of organoids were dehydrated using increasing concentrations of sucrose (10%, 20% and 30%) overnight. Organoids or gel plugs were placed in optimal cutting temperature (OCT) embedding matrix gel (Tissue-Tek, KMA-0100-00A) in a disposable histology mould (Lecia, 3803025) and placed at -80°C overnight. Using a cryostat (Leica, CM 1900) that was cooled to -25°C prior to use, OCT embedded organoids were mounted onto the freezing block using OCT and faced off by cutting 60 μ m sections until the organoids were exposed. The organoids were cut into 8 μ m sections and placed onto SuperFrost Plus slides (ThermoFisher, 10149870), and stored at 4°C until further use.

2.7.3.2. Immunofluorescence staining

Cells were blocked for 1 hour 30 minutes in 10% Goat Serum (Gibco, 16210064), the primary antibody (Table 2.10) was made in the PTB solution, and was added to the cells overnight at 4°C. The cells were washed three times in the PTB solution, and the secondary antibody was made up in the PTB solution and added to the cells. The cells were incubated on a rocker for an hour at room temperature in the dark. The secondary antibody was removed, and the cells were washed three times in the PTB solution, DAPI (1:2500) was added to the cells for 3 minutes, and the cells were washed three times in PTB solution. The coverslips were mounted using ProLong Gold Antifade Mountant (Invitrogen, P36930), and allowed to dry for 24 hours.

Table 2.10: Antibody conditions, secondary antibodies, suppliers and catalogue numbers for antibodies.

<i>Antibody</i>	<i>Concentration</i>	<i>Secondary</i>	<i>Supplier</i>	<i>Catalogue number</i>
ALDH1A1	1:200	Rabbit	Abcam	ab52492
CXCR4	1:100	Mouse	SCBT ¹	sc-53534
EpCAM	1:50	Mouse	SCBT	sc-25308
HCAM	1:100	Mouse	SCBT	sc-7297
MASPIN	1:50	Rabbit	Invitrogen	PA5-35104
PANCK	1:200	Mouse	SCBT	sc-57012
PDX1	1:50	Rabbit	Invitrogen	PA5-14824
RAD51	1:500	Rabbit	Invitrogen	PA5-27195
Alexa Fluor 488 Mouse	1:1500	—	Invitrogen	A-11001
Alexa Fluor 488 Rabbit	1:2000	—	Invitrogen	A-11008

¹SCBT: Santa Cruz Biotechnology.

2.7.3.3. Confocal imaging

The immunofluorescence was observed using a Leica TCS SP8 STED super-resolution microscope equipped with a CCD camera and 100X oil immersion objective. DAPI was excited with a 405 nm PicoQuant laser unit and emission

captured between 387 and 474 nm. Alexa Fluor 488 was excited at 499 nm with emission captured between 490 and 566 nm. Images were acquired whereby combinations of excitation and emission wavelengths for specific dyes were applied sequentially. The fluorescence intensity was quantified using ImageJ software. The area integrated intensity and mean grey value around the cells was measured, and the average background was subtracted from the images.

2.7.4. Immunohistochemistry

Immunohistochemistry (IHC) was performed on a DAKO Autostainer. Deparaffinisation and antigen retrieval was performed on slides using the DAKO PT Link and pH 6 (DAKO, S169984) or pH 9 (DAKO, S236784-2) antigen retrieval solution. The slides were placed in the PT Link at 65°C, heated to 95°C and maintained at this temperature for 20 minutes. The slides were then cooled to 65°C and placed in the DAKO Autostainer, and reagents were added as listed in Table 2.11 and antibodies are listed in Table 2.12. Following the staining process, the slides were dehydrated in a fume hood using 70%, 90% and 100% ethanol for two 3-minute washes, followed by two 5-minute xylene washes. Slides were mounted using DPX (Sigma, 44581) and the slides were allowed to air dry overnight in the fume hood.

Table 2.11: Reagents and times for DAKO Autostainer.

<i>Reagent</i>	<i>Time (minutes)</i>
Real HP Blocking (DAKO, S202386)	10
Antibody CASP7 (SCBT, SC-365034)	30
Real EnVision (DAKO, K500711-2)	30
Real DAB (DAKO, K500711-2)	5
Real DAB (DAKO, K500711-2)	5
Haematoxylin (DAKO, S330130-2)	10

Table 2.12: Antibody conditions, suppliers and catalogue numbers for antibodies used in immunohistochemistry.

<i>Antibody</i>	<i>Concentration</i>	<i>Supplier</i>	<i>Catalogue number</i>
ALDH1A1	1:200	Abcam	ab52492
CASP7	1:250	SCBT ¹	sc-365034
CXCR4	1:100	SCBT	sc-53534
EpCAM	1:50	SCBT	sc-25308
HCAM	1:100	SCBT	sc-7297
Ki67	1:200	DAKO	M724029-2
MASPIN	1:50	Invitrogen	PA5-35104
PANCK	1:200	SCBT	sc-57012
PDX1	1:50	Invitrogen	PA514824

¹SCBT: Santa Cruz Biotechnology.

2.7.4.1. Scoring

Slides were viewed using a Leica microscope and were scored according to the presence or absence of staining.

2.7.4.2. Clinicopathological features of patient cohort

IHC was performed on a patient cohort of 49 tumour slides, and 3 normal pancreatic slides. Archival PDAC tumours and control normal tissues were obtained courtesy of the Department of Pathology, SVUH Dublin. Ethical approval for IHC analyses of tissues was obtained from the SVUH ethics committee. Clinicopathological details were obtained from patient records. Survival analysis was performed using patient survival data last updated in January 2017. 36.7% of the patient cohort were female, and 63.3% were male, patients had a median age of 64 years, and an age range of 38-81 years at the time of surgery. Patients had an average survival of 25.9 months following surgery, and 14 patients were still alive in January 2017. All patient information is available in Table 2.13. Kaplan-Meier analysis of the patient samples was performed using GraphPad V. 7d (Graphpad Software Inc., San Diego, CA, USA), comparing two groups – patients with no CASP7 staining and patients with CASP7 staining.

Table 2.13: Clinical patient information for PDAC cohort (n=49 patients).

Clinicopathological Feature	Number of Patients (Percent)
Gender	
Male	31 (63.27%)
Female	18 (36.73%)
Age at surgery (median)	64
T stage	
1	1 (2.04%)
2	3 (6.12%)
3	43 (87.76%)
N stage	
0	35 (71.43%)
1	14 (28.57%)
R0 (>1mm)	
Yes	48 (97.96%)
No	1 (2.04%)
Differentiation	
Poor	13 (26.53%)
Moderate	31 (63.27%)
Well	5 (10.20%)
Max pathological axis	2.96 cm
Lymphatic invasion	
Absent	17 (34.69%)
Present	32 (65.31%)
Perineural invasion	
Absent	2 (4.08%)
Present	47 (95.92%)
Portal vein involvement	
No	47 (95.92%)
Yes	2 (4.08%)
Survival Status	
Dead	35 (71.43%)
Alive	12 (24.49%)
Average Survival (Months)	25.9

2.8. Western blotting

Cells were seeded in 6 well plates and grown for 7 days. Cells were washed three times with PBS and 50 μ L/well RIPA buffer (Sigma, R0278) 5 mM Tris-HCl pH 7.4, 1% NP-40, 0.1% SDS, 150 mM NaCl, 1% Triton x-100) containing 1X Protease Inhibitor cocktail (Calbiochem, US1539131-IVL), 2 mM PMSF (Sigma, 10837091001), and 1 mM sodium orthovanadate (Sigma, 5086050004) was added and cells were incubated on ice for 20 minutes. Cells were scraped and lysis buffer collected. Lysates were sheared with a 21-gauge needle and centrifuged at 10,000 RPM for 10 minutes at 4°C. Supernatant was collected and stored at -80°C. Protein quantification was carried out using a bicinchoninic acid (BCA) quantification kit (Pierce, 23225).

Samples were made up to 20 μ g using PBS, loading buffer (4X) (Invitrogen, NP0007) and reducing agent (10X) (Invitrogen, NP0009). Samples were then electrophoretically resolved on Bolt 4-12% Bis-Tris Plus Gels (15-well) (Invitrogen, NW00125BOX). The proteins were then transferred to a nitrocellulose membrane using the Trans-Blot Turbo Transfer System (BioRad, 1704159). Ponceau S (Sigma Aldrich, P7170) was used to determine successful transfer of the protein to the membrane. The membrane was blocked in 1X NET buffer (1.5 M NaCl, 0.05 M EDTA, 0.5 M Tris pH 7.8, 0.5% Triton X-100, 0.25% w/v porcine gelatin) for 1 hour at room temperature with gentle rocking. Primary antibody was made up in NET. The membranes were incubated in primary antibody (Table 2.14) overnight at 4°C with gentle rocking. The primary antibody was removed, and membranes were washed three times for 10 minutes in NET buffer at room temperature. The membranes were then incubated in secondary antibody for 60 minutes. Three further ten-minute washes were performed in NET at room temperature and blots were developed using the Odyssey Imaging system (LI-COR Imaging Systems). Alpha tubulin (Sigma, T6199) was used as a loading control for Western blots performed. ImageJ software was used to perform densitometry analysis.

Table 2.14: Antibody conditions, suppliers and catalogue numbers for antibodies used in Western Blotting.

<i>Antibody</i>	<i>Secondary</i>	<i>Concentration</i>	<i>Source</i>
Anti-mouse secondary	-	1:5000	LiCor
CASP7	Mouse	1:1000	SC-365034

2.9. RNA extraction

Organoids and CLOs were grown as described for 14 days. TRI-reagent® (Sigma, T9424) was used to lyse cells, and RNA isolation was performed using Direct-zol RNA Miniprep Plus Kit (Zymo Research, R2072). RNA samples were quality controlled with an RNA Screen Tape (Agilent) and quantified with Qubit RNA HS Assay kit.

2.9.1. cDNA synthesis

A High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, 4368814) was used to synthesise cDNA from RNA. A master mix was prepared as described in Table 2.15. RNA was diluted to 10 µg in 10 µL in a 0.2 mL PCR tube and 10 µL of the Master Mix was added. A G-Storm Thermocycler (Model GSI, Somerton Biotechnology Centre) was used to perform cDNA synthesis using the conditions listed in Table 2.16. Following synthesis, cDNA was stored at -20°C.

Table 2.15: Reverse Transcription (RT) master mix made using the High-Capacity cDNA RT Kit.

Master Mix	IX (μL)
RT ¹ Buffer	2.0
dNTP ² Mix	0.8
Random Primers	2.0
Reverse Transcriptase	1.0
RNase Inhibitor	1.0
Nuclease Free Water	3.2
Total	10

¹RT: Reverse transcription; ²dNTP: deoxynucleotide triphosphate.

Table 2.16: Conditions used to perform PCR using the G-Storm thermocycler.

Step	Temperature ($^{\circ}\text{C}$)	Time (Minutes)
Annealing	25	10
cDNA Synthesis	37	120
Enzyme Inactivation	85	5
Storage	4	∞

2.9.2. Quantitative reverse transcription PCR (RT-qPCR)

The assay was prepared using 20X TaqMan Gene Expression Assay (Table 2.17), and master mix (Table 2.18) with 20 ng of cDNA, in MicroAmp Fast Optical 96-well reaction plates (Applied Biosystems, 4346907). The plate was sealed using Adhesive PCR Plate Seals (Thermo Scientific, AB0558) and centrifuged briefly. Using an Applied Biosystems 7900 Real-Time PCR System RT-qPCR was performed at 50°C for 2 minutes, 95°C for 20 seconds, and 40 cycles of 95°C for 1 second and 60°C for 20 seconds. Relative quantification was measured by the delta delta Ct method with GAPDH or 18S as an endogenous control. Each biological

replicate was measured in technical triplicate wells. Representative RT-qPCR raw data shown in Appendix B, Supplementary Table 8.1.

Table 2.17: TaqMan Gene expression assays used in this project, and assay ID.

20X TaqMan Gene Expression Assay	Assay ID
NANOG	Hs02387400
POU5F1	Hs04260367
SOX2	Hs01053049
CASP7	Hs00169152
18S	Hs03003631
GAPDH	Hs02786624

Table 2.18: Master mix for RT-qPCR.

PCR reaction mix component	Ix (μL)
20X TaqMan Gene Expression Assay	0.5
2x TaqMan Gene Expression Master Mix	5.0
cDNA Template (1 to 100 ng)	2.0
RNase-free Water	2.5
Total	10.0

2.9.3. RNA-sequencing

RNA library preparation, sequencing and analysis were performed by RealSeq Biosciences Inc. (California, USA).

2.9.3.1. Library preparation

Sequencing libraries were prepared with 200 ng of total RNA for all samples. The Nugen Universal Plus mRNA protocol was followed as recommended by the manufacturer. Libraries were amplified by 15 cycles of PCR. Libraries were sequenced in one NextSeq 550 run with the NextSeq 500/550 High Output Kit

v2.5 (300 cycles), sequencing was done with paired-end 150 nucleotide reads and 8 nucleotide indexes. Libraries were loaded at 1.4 pM, 5% PhiX control was used.

2.9.3.2. RNA-seq data analysis

RNA-seq data analysis was kindly performed by Chenxi Zhang of the Lars Bolund Institute of Regenerative Medicine, BGI-Qingdao, China.

Raw sequencing files were merged for each sample to generate a single FASTQ file per sample. Cutadapt was used to remove adapter sequences. Trimmed reads were aligned against the *Homo sapiens* GRCh37 genome build (hg19) with hisat2. The resulting SAM file was piped into samtools for BAM conversion and sorting. featureCounts was used to generate a gene counts matrix for DESeq2. This counts generation is written to counts.tsv for further analysis. Up/downregulated genes: differentially expressed genes ($p < 0.001$ and $|\log_2 \text{fold-change}| \geq 2$) are detected using DESeq2 after genes with counts less than 10 [278]. The following comparisons were performed: CLOs compared with cell lines; PT127 PDX and CLO compared with cell lines; PT291 organoids and CLO compared with cell lines. The correlation of paired groups - for each sample, reads count is normalised to transcripts per million (TPM). The median was applied if there was more than one sample in a group. Linear regression was performed to identify the relationship of paired groups: PT291 ORG and PT127 PDX; PT127 PDX and PT127 CLO; PT291 ORG and PT291 CLO, ($p < 0.05$).

2.9.3.3. Tiriak et al. PDAC organoid RNA-seq dataset analysis

The dataset from the study Organoid Profiling Identifies Common Responders to Chemotherapy in Pancreatic Cancer by Tiriak et al. [236] was obtained through dbGAP: (phs001611). RNA-sequencing data was available for a total of 49 samples, including 38 PDAC organoids, and 11 normal pancreas organoids. Gene lists containing genes controlled by transcription factors *HNFI1A*, *HNFI1B*, *HNFI4A*, *HNFI4G*, *GATA6*, *NR5A2* and *PDX1* were obtained from <https://maayanlab.cloud/Harmonizome/> (Appendix C). PDAC samples were manually divided into high and low expression of transcription factor of interest using gene counts transcripts per million to TPM, where low expression was ≤ 6.25 ,

and high expression was ≥ 6.25 . Data analysis was performed using iDEP.93 (<http://bioinformatics.sdstate.edu/idep/>), a web application for differential expression and pathway analysis of RNA-seq data [279]. Analysis was performed by looking at differences in downstream gene expression in samples with low and high expression of the transcription factor of interest.

2.10. *In vivo* xenograft study

In vivo work was carried out by Neil Conlon and Justine Meiller.

2.10.1. Tumour induction via subcutaneous implant

PT291 organoids and CLOs were seeded as described in Section 2.3.3.6, and grown for one week prior to implantation. The number of cells in each well was estimated by passaging and performing a cell count on three wells and obtaining the average cell count. Cells were collected from the number of wells required for a cell suspension of 2×10^7 cells. For the cell line, cells were trypsinised, counted and a cell suspension of 2×10^7 was made.

Cell suspensions were centrifuged at 1500 RPM for 10 minutes. Cell pellets were resuspended in 1 mL of ECM. The ECM-cell suspension was plated in 400 μ L ECM domes, and allowed to solidify at 37°C. SCID mice were implanted subcutaneously with the ECM-cell dome. Under anaesthesia (isoflurane, O₂ carrier gas) a small incision was made in the skin of the left flank of the animal. The ECM-cell dome was then drawn into a syringe and was injected in the pocket under the skin and the wound sealed with a single staple. The animals were monitored post-surgery, and staple removal was within 10 days. Animals were monitored weekly for body weight and tumour development.

2.10.2. Tumour measurements

Tumours were measured by height, width, and depth, using electronic callipers. Tumour volume was calculated as outlined in Equation 2.

Equation 2: Calculation of tumour volume using tumour height, depth and width measurements

$$\text{Tumour volume} = \frac{(\text{Height} \times \text{width} \times \text{depth})}{1.9}$$

2.10.3. Tumour retrieval and processing

Mice were culled and tumours were excised when the tumour reached a volume of <1600 mm³, or a tumour axis <15 mm, or due to adverse events. The tumour was then quickly divided into sections and preserved by liquid nitrogen flash freezing and formalin fixation. Flash-frozen samples were stored at -80°C and formalin fixed samples were dehydrated in 50%, 70%, 90%, 100% ethanol, followed by 100% xylene, and then embedded in paraffin. Formalin fixed paraffin embedded (FFPE) blocks were stored at 4°C until sectioning.

2.10.4. Sectioning

Sectioning of *in vivo* tumour samples was carried out using a Reichert-Jung 2030 microtome. Blocks were cut into 5 µm sections, which were floated in a water bath at 40°C and mounted onto a SuperFrost Plus slide (ThermoFisher, 10149870) and allowed to dry at 60°C for 2 hours.

2.11. CUT&RUN

2.11.1. CUT&RUN assay

PT291 CLOs were seeded as described in Section 2.3.3.6, and grown for one week prior to the assay. Cells were passaged, and 5x10⁵ cells were used per reaction. CUT&RUN was performed using CUT&RUN Assay Kit (CST, 86652S) and was carried out as per manufacturer's guidelines, using antibodies in Table 2.19. Following extraction, DNA was purified, using DNA Purification Buffers and Spin Columns

(ChIP, CUT&RUN) (CST, 14209S) as per manufacturer's guidelines. Samples were eluted in 50 µL elution buffer.

Table 2.19: Antibodies used for CUT&RUN.

<i>Antibody</i>	<i>Concentration</i>	<i>Supplier</i>	<i>Catalogue Number</i>
<i>HNFI</i> A Rabbit monoclonal	1:25	CST ¹	89670
<i>HNFI</i> B Mouse monoclonal	1:25	ThermoFisher	MA5-24605

¹CST: Cell Signalling Technology

2.11.2. Sequencing

Sample library preparation and sequencing was performed by Quick Biology Inc. (California, USA) and bioinformatics was performed by Charlotte Andrieu.

2.11.2.1. Library Preparation and sequencing

Library was prepared with KAPA Hyper Prep Kit (KAPA Biosystems, Wilmington, MA) using 25 µL DNA as input. Final library quality and quantity was analysed by Agilent Tapestation and Life Technologies Qubit3.0 Fluorometer. 150 basepair paired-end reads were sequenced on Illumina HighSeq (Illumina Inc., San Diego, CA).

2.11.2.2. Bioinformatics

Briefly, raw paired-end reads for HNFI A, HNFI B or IgG CUT&RUN were trimmed using Sickel, and aligned using Bowtie2, filtered for only properly-paired reads, then molecular complexity was estimated, and libraries were subsampled to equalise based on complexity estimates [280]. PCR duplicates were removed, and MACS2 was used to call narrowPeaks against the IgG negative control as input [281]. Peak centring and annotation was performed using HOMER annotation [282]. Tracks were visualised using Integrative Genomics Viewer (IGV) software v2.8.6, with tracks with ChIP-seq data for HNFI A and HNFI B in HepG2 cells for comparison [283,284].

2.12. SNP selection

2.12.1. Literature search

A systematic and comprehensive search of PubMed was performed on genes from pathways/gene sets: MODY and Pujana ATM PCC. The search was restricted to human studies only with no restrictions placed on language. The search terms (1) gene name OR common variation; (2) PDAC OR pancreatic cancer, and (3) SNP OR polymorphism OR variant, were combined. All genes in the KEGG MODY Pathway and Pujana ATM PCC dataset were used as search terms. Studies that investigated SNPs and their correlation with PDAC prognosis were included for review. Randomised control trials (RCT) and cohort studies were included. Articles were excluded if they were review articles. All results were reviewed and verified by checking the original publication, and the SNPs must have been (1) identified in a population of European descent, (2) a study with $N > 1000$, and (3) statistically significant.

The NHGRI-EBI (National Human Genome Research Institute-European Bioinformatics Institute) catalogue (<https://www.ebi.ac.uk/gwas/>) was used to search for SNPs associated with PDAC in genes in the MODY pathway and Pujana ATM PCC gene set.

2.12.2. Identification of proxy variants

LDlink (<https://ldlink.nci.nih.gov/>) was used to identify putatively functional proxy variants for the SNPs identified. An $R^2 \geq 0.8$ and a RegulomeDB score of ≤ 3 was used as a cut-off for linkage disequilibrium (LD) proxies. Duplicate SNPs were removed, LD pairwise analysis tables were generated using R^2 values from LD-matrix (<https://ldlink.nci.nih.gov/?tab=ldmatrix>) and heatmaps were generated using Prism GraphPad v9.

2.12.3. *In silico* analysis of SNPs

Identified SNPs were further assessed through the following:

- RegulomeDB – identification of the putative regulatory potential and position weight matrixes (<http://regulomedb.org/>).

- NIH Roadmap Epigenomics: to assess the chromatin state in the pancreas (<https://www.ncbi.nlm.nih.gov/geo/roadmap/epigenomics/>).
- UCSC genome browser – for identification of CpG Islands (<https://genome.ucsc.edu/>).
- Haploreg v4.1 – annotation of non-coding genome variants (<https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>).
- GTEx – assessment of tissue specific gene expression and regulation (<https://www.gtexportal.org/home/>).

2.12.4. Prioritisation of SNPs

SNPs were prioritised for functional analysis by assessing scores from *in silico* tools:

- NIH Roadmaps Epigenomic: SNPs occurring in regions with a quiescent/low chromatin state were removed.
- RegulomeDB: SNPs with a low score (≤ 3) were prioritised.
- UCSC Genome browser: SNPs occurring in a CpG Island were prioritised.
- GTEx: SNPs shown to affect gene expression were prioritised.
- Haploreg: SNPs within regions with histone marks (H3K4me1, H3K4me3, H3K27ac, H3K9ac and DNase) in the pancreas were prioritised.

A flowchart for the methods used for SNP identification, *in silico* assessment and prioritisation is shown Figure 2.1.

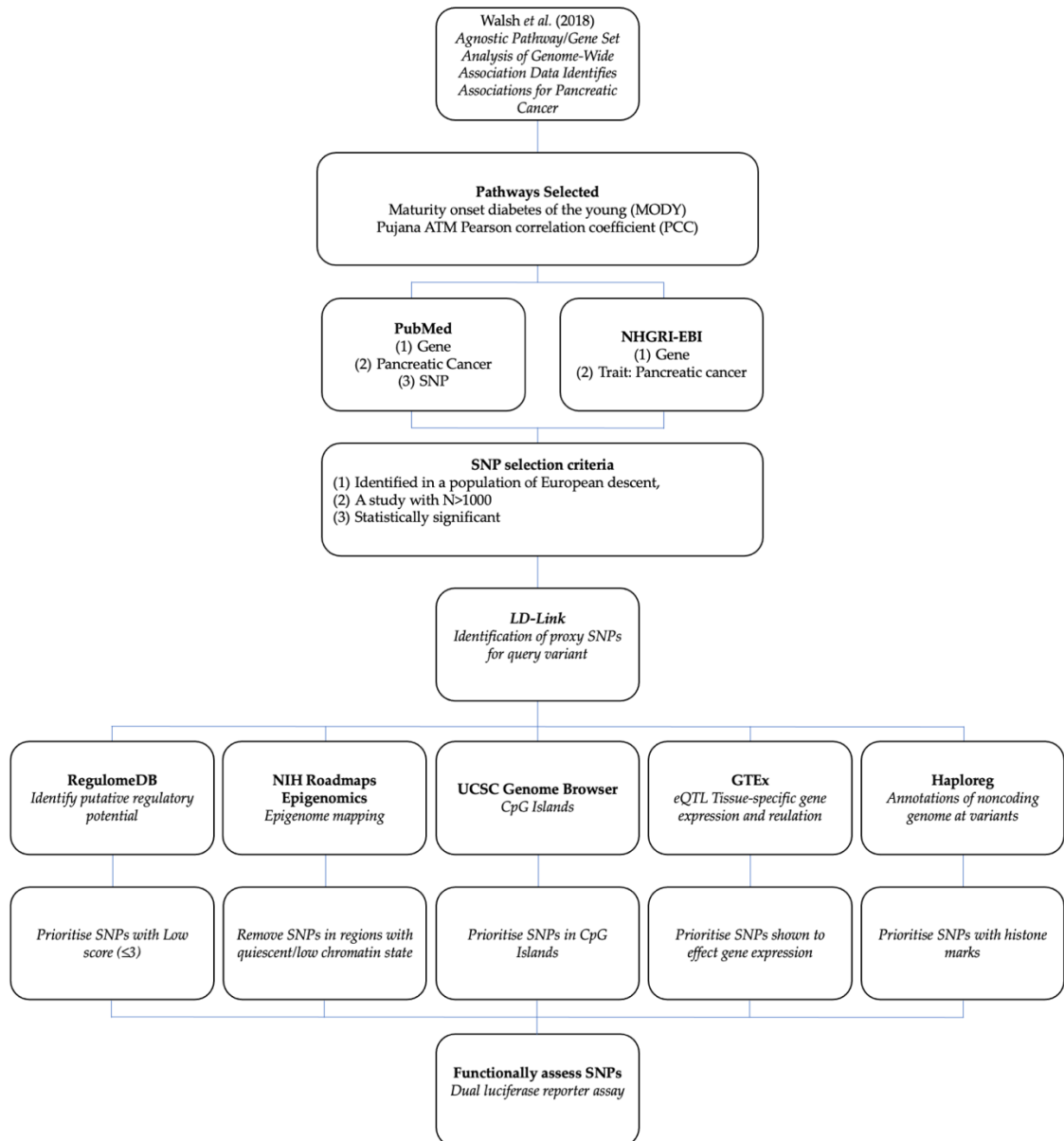


Figure 2.1: The methods used for SNP identification, in silico assessment and prioritisation of SNPs from the MODY pathway and Pujana ATM PCC gene set from the study by Walsh et al. [122].

2.12.5. Position weight matrix (PWM)

HOCOMOCO-v11 was used to generate PWM, and to identify changes in binding motifs due to the presence of the alternative SNP compared to the reference SNP. Fold-change values were according to PERFECTOS-APE [285,286].

2.13. Dual luciferase reporter assays

2.13.1. SNP cloning

Eleven sequences were amplified by PCR using genomic DNA isolated from a healthy donor peripheral blood using Genomic DNA Purification Kit (Thermo Fisher Scientific, Waltham, MA, USA) with the following primers introducing restriction sites (lowercase letters) for HindIII (forward, except BamHI forward for rs4742901 / rs4742902) and BamHI (reverse) (Table 2.20).

Table 2.20: Primer sequences used for genomic DNA cloning of SNP region, including gene, SNP of interest, length of region in base pairs (bp), and forward and reverse primers including restriction sites (lowercase letters) and overhang regions (bold).

Gene	SNP	Length, bp	Forward Primer (5'-3')	Reverse Primer (5'-3')
<i>CASP7</i>	rs3124737	528	CCCaagcttGCCACAAGGTCCTCAATACCA	CGCggatccCACTGCATGGGTAGCCGTTA
<i>CEL</i>	rs488087	1102	CCCaagcttTGTGTGTTTGAGGGATGAGAGTTAC	CGCggatccATACCCCTCCCCTACTGCACCTT
<i>HNFI A</i>	rs2258287	466	CCCaagcttCCCCCAACCCAGTACAGTTTCAT	CGCggatccACTCTATGGTCAAGCCCCCACA
<i>SMC2</i>	rs4742902/ rs4742902	409	CGCggatccCTCGGAACAGCTCTCGGTAG	CGCggatccGCTTAGCCTTGGGAGCAAGA
<i>BCL2L1I</i>	rs2241845	555	CCCaagcttTACTCCGGTAAACACGCCAG	CGCggatccTCTACCTCCCAGGAGACCAAA
<i>BCL2L1I</i>	rs11378324	942	CCCaagcttCTCCGAGTCTACTTAGCCGT	CCCaagcttACACGAAACACCCCTCACACT
<i>PNMT</i>	rs876493	548	CCCaagcttGGGAGTTGTGTGGCAGTTCTA	CGCggatccTTTTTCTACCGCCCCGACCCC
<i>HNFI B</i>	rs718961	592	CCCaagcttTGGTTCTGTCTGGGTAAACAGA	CGCggatccACAGTACTGGGTCCATCAAGC
<i>HNFI B</i>	rs4795218	1836	CCCaagcttGGGCAGCCCATTAGGTCCTCTG	CGCggatccCCACCCACGCCCTCATCCTTAC
<i>NR5A2</i>	rs3790843	1423	CCCaagcttCCCTAGACCCGAAACTGGGGA	CGCggatccGACACACAGAAAGGGTTTGGTT
<i>NR5A2</i>	rs3790844	1956	CCCaagcttTCGCTGCTACGCTTAGGTTT	CGCggatccAAATGCAGCCGGGTCTTACA

PCR was performed using a high-fidelity Platinum™ SuperFi™ DNA Polymerase (ThermoFisher, 12351010) using the master mix outline in Table 2.21. Amplification was carried out in MiniAmp™ Plus Thermal Cycler (ThermoFisher, A37835), using a two-step PCR, with the samples held at 98°C for 30 seconds, followed by 35 cycles of 98°C for 10 seconds, 72°C for 30 seconds per kb and followed by a final extension step of 72°C for 10 minutes. Samples were stored at 4°C. 5 µL of each sample was run on a 1% agarose gel with a low molecular weight gene ladder (ThermoFisher, SM1193) for 30 minutes to confirm the presence of PCR product. A PCR product clean-up was performed using GenElute™ PCR Clean-Up Kit (Sigma, NA1020) and PCR products were eluted in 20 µL of nuclease free water.

Table 2.21: Master mix for PCR with Platinum™ SuperFi™ DNA Polymerase.

<i>Reagent</i>	<i>IX (µL)</i>
Nuclease free water	10.75
5X SuperFi Buffer	5
10mM dNTP	0.5
10 uM forward primer (Table 2.20)	1.25
10 uM reverse primer (Table 2.20)	1.25
Template DNA	1
5x SuperFi GC Enhancer	5
Platinum SuperFi DNA Polymerase (2U/ µL)	0.25

2.13.2. Restriction enzyme digestion

A restriction enzyme digestion was performed on PCR products and the vector, a p2-LucSG vector, a kind gift from Dr Gary Loughran, School of Biochemistry and Cell Biology, University College Cork (Figure 2.2). For the vector, 1 µL of plasmid was used and for the cloned region, 5 µL of the cleaned PCR product was used. These were mixed with 5 µL 10X NEB Buffer 3, 1 µL HindIII-FD and 1 µL BamHI-FD and nuclease free water up to 50 µL. For the digestion of the vector, 5 µL alkaline phosphatase was added to the reaction to prevent self-ligation of the

vector. The reactions were digested for 20 minutes at 37°C, and digestion product clean-up was performed using GenElute™ PCR Clean-Up Kit (Sigma, NA1020) and digestion products were eluted in 20 µL of nuclease free water.

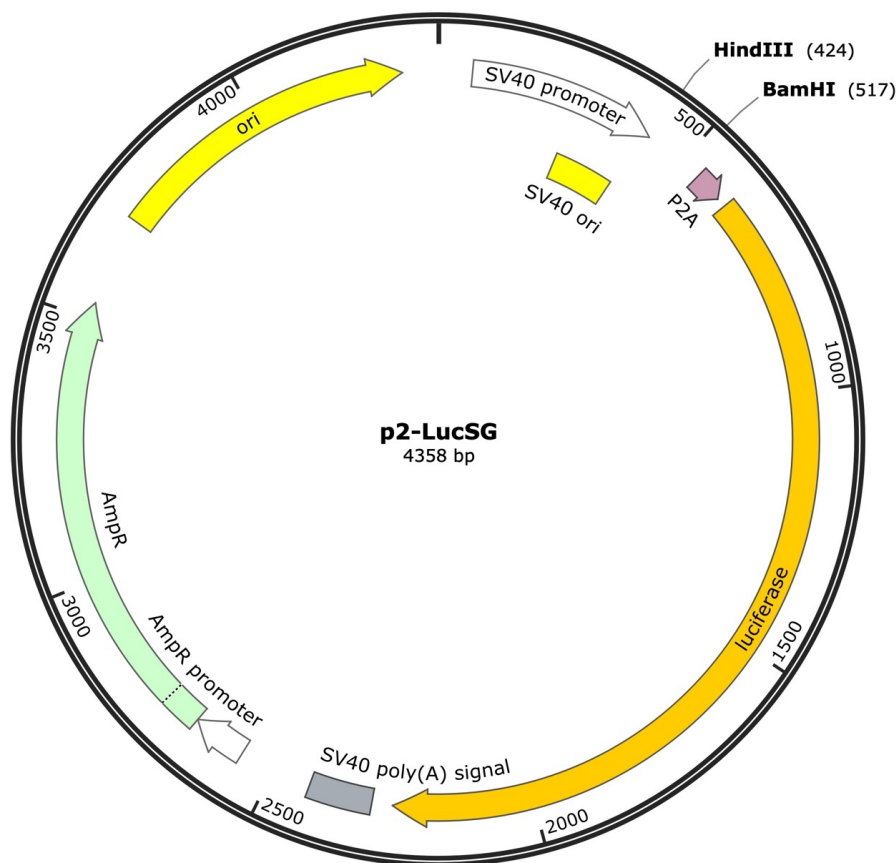


Figure 2.2: Schematic of p2-LucSG plasmid, which contains a luciferase gene, ampicillin promoter (AmpP) and resistance (AmpR), and SV40 promoter. A porcine enterovirus StopGo sequence, P2A, was introduced upstream of the luciferase gene, ensuring the amino acid sequence of the luciferase gene is correct.

2.13.3. Ligation

The cloned region containing the SNP is inserted upstream of the luciferase gene by the HindIII/BamHI restriction sites. The ligation reaction is carried out in a 5 µL reaction with 1 µL of the digested p2-LucSG plasmid, 1 µL of the digested cloned region, 0.5 µL T4 DNA Ligase Reaction Buffer (NEB, B0202S), 0.25 µL T4 DNA Ligase (NEB, M0202S) and 2.25 µL nuclease free water. The reaction was left at room temperature for 1 hour.

2.13.4. Transformation

NEB 5-alpha Competent *E. coli* cells (NEB, C2987H) were transformed with 5 µL of the ligation product by thawing the cells on ice and 5 µL of the ligation product was added. The tube was flicked 5 times to mix cells and DNA and placed on ice for 2 minutes. The mixture was heat shocked at 42°C for 30 seconds and placed on ice for two minutes. 150 µL of the SOC outgrowth media was added to the cells, and 100 µL of the mixture was spread onto an ampicillin-agar plate and incubated overnight at 37°C.

2.13.5. PCR screening

Three single colonies were picked using a tip, dipped in PCR tubes containing 50 µL LB medium with ampicillin (1:1000) and then tips were rinsed in a PCR tube containing 20 µL nuclease free water. The PCR tubes containing LB were placed at 37°C for two hours, then stored at 4°C. The tubes with water were placed 98°C for 10 minutes to lyse bacteria and screened using the PCR conditions outlined in Section 2.13.1 to check plasmids were the correct size. The colonies of the correct size were grown in 5 mL LB-ampicillin at 37°C, at 250 RPM overnight. Plasmid DNA extraction was carried out using a PureLink™ HiPure Plasmid Miniprep Kit (ThermoFisher, K210002) and plasmids were quantified using the NanoDrop, and stored at -20°C.

2.13.6. Site directed mutagenesis

Alternative SNP alleles were introduced using Q5® Site Directed Mutagenesis Kit (NEB, E0554S) according to manufacturer's instructions. Primers were designed using <http://nebasechanger.neb.com/> and PCR was carried out at the temperatures defined in Table 2.22. Following PCR, a KLD (Kinase, Ligase and DpnI) enzyme reaction was performed. 5 µL of the KLD enzyme treated product was transformed, screened and extracted using the method outlined in Section 2.13.4. All constructs were verified by Sanger sequencing.

Table 2.22: Forward and reverse primers, and PCR temperature for site directed mutagenesis (SNP) to introduce alternative SNP, which is in lowercase.

SNP	Forward Primer (5'-3')	Reverse Primer (5'-3')	Temperature (°C)
rs2258287*	GCTTGATAACtGCGGTGTTCTC	CTCCTTCATGACGTCACG	62
rs3124737	CCTTTGTGTGGTGTTGTATG	AAATGTCCTCTTATTGATTTTAG	57
rs4742901	CCGAGGCCAGcACGGGGCGGA	ACCTCGGGGAGCGTTTTTATCTGAAGTTTCC	72
rs4742902	AGTAAGCCACgGCCAGCACCG	TTTTCTCTCCTCAAAGCGACAGTTAGG	69
rs488087	AGGCCACTCCtGTGCCCCCCA	CGGAGTCCCCCTGTGGGGG	72
rs718961	AGAAATGAGGgAAGTGGGgATTc	TCTTACTAAAAGAACAAATTCTGG	59
rs3790844	CTAAAACTGGgAAGTCTGTcG	TGACCTGTGTTTTCCACAC	59
rs2241845	CCGGGAGCTGgGGACCTGCTC	GCCGGCGCCGAAAGGCAC	72
rs11378324	ggggggGCGCCAGCCACGCGTTGG	CAACCCCGGCCCTCCTGTC	72

*T allele used for SDM, as sequence was inserted into vector in the 3' to 5' direction as the nearest coding region, C12orf43, is transcribed from the reverse strand

2.13.7. Dual luciferase reporter assay

Dual luciferase reporter assays were performed using p2Luc-SG vectors containing regions of interest with reference and alternative SNP of interest to assess if the SNP results in a change in luciferase gene expression.

Seeding: ASPC1, HEK293T, and PANC1 cells were used for this assay. Cells were grown to 70% confluency, and a cell suspension of 2.5×10^4 cells/mL was made. Cells were seeded at 100 μ L per well in a black-sided 96-well plate (Corning, CLS3603). Plate was placed overnight in a 37°C, 5% CO₂ incubator.

Transfection: Assays were carried out using 25 ng p2Luc-SG vectors and 5 ng control Renilla vector and 0.5 μ L X-tremeGENE transfection reagent, were mixed in 25 μ L Opti-MEM (Gibco, 11058021) and incubated at room temperature for 15 minutes. The transfection mixture was added to the cells, and a GFP plasmid was used as a transfection control. The plate was incubated at 37°C, 5% CO₂ overnight.

Dual luciferase reporter assay: The Stop and Glow (Promega, E2920) dual luciferase reporter assay System was used. Reagents were prepared according to protocols provided. The 96 well plate was removed from the incubator and equilibrated to room temperature. Dual Glo Substrate (100 μ L) was added to each and the plate was incubated at room temperature for 10 minutes. The plate was read using a plate reader (Biotek with Gen4 software) with a gain of 100, and integration time of 1 second to determine Firefly luciferase (p2Luc-SG) luminescence. Stop and Glo reagent (100 μ L) was added, and the plate was incubated at room temperature for 10 minutes. The plate was read using the same settings to determine Renilla luciferase luminescence. Relative luminescence was calculating the relative Firefly to Renilla for each well, then normalising to the p2Luc-SG vector with no insertion.

2.14. CRISPR

2.14.1. Golden Gate Assembly

2.14.1.1. CRISPR sgRNA construction

All CRISPR gRNAs were designed with the online software tool CRISPOR (<http://crispor.tefor.net/>). sgRNA was constructed using oligos (Table 2.23) and 10X NEBuffer 2 (NEB, B7002S). The reaction was made as described in Table 2.24 in a PCR tube and placed in a MiniAmp™ Plus Thermal Cycler (ThermoFisher, A37835). The reaction was heated to 95°C for 5 minutes, then set to decrease by 0.5°C every cycle for 150 cycles (to 20°C), then held at 20°C for 5 minutes. Products were stored at -20 °C.

Table 2.23: Oligo sequences for the generation of sgRNA.

<i>Gene</i>	<i>Same-sense</i>	<i>Anti-sense</i>
<i>HNFI A</i>	CACCAGGACGAGACGGACGACGAT	AAACATCGTCGTCCGTCTCGTCCT
<i>HNFI B</i>	CACCGGCCGCTTGTCCGGCGACGA	AAACTCGTCGCCGACAAGCGGCC
<i>HNFI 4G</i>	CACCGTGTTGTTGACAAGGACAAA	AAACTTTGTCCTTGTCAACAACAC
<i>NR5A2</i>	CACCGTAAGGGCCGACCGAATGCG	AAACCGCATTCCGGTCGGCCCTTAC
<i>PDXI</i>	CACCTCGTACGGGGAGATGTCCGG	AAACCCGGACATCTCCCCGTACGA

Table 2.24: Master mix for CRISPR sgRNA construction.

<i>Reagent</i>	<i>Ix (μL)</i>
SS oligo (100 μM)	1
AS oligo (100 μM)	1
10X NEBuffer 2	2
ddH ₂ O	16
Total	20.0

2.14.1.2. Ligation

CRISPR sgRNA is inserted into plasmid lentiCRISPR v2 (Addgene, Plasmid #52961) using the mixture described in Table 2.25. The reaction was added to a PCR tube, and the MiniAmp™ Plus Thermal Cycler was set to the times outlined in Table 2.26. A transformation was performed as outlined in Section 2.13.4 using 5 μ L of the ligation product.

Table 2.25: Components and volumes required for ligation of sgRNA to lentiCRISPRv2 plasmid.

Reagent	Ix (μL)
100ng lentiCRISPRv2 plasmid	1
sgRNA	1
T4 ligase buffer	2
T4 ligase	0.5
BsmBI	0.5
ddH ₂ O	15
Total	20.0

Table 2.26: Temperature conditions for ligation of plasmid.

Temperature	Time	Number of Cycles
37°C	5 min	10
22°C	10 min	
37°C	30 min	1
75°C	15 min	1
4°C	∞	—

2.14.1.3. PCR screening

Single colonies were picked, and lysates were prepared using the conditions outlined in Section 2.13.5. A PCR was set up using the mixture in Table 2.27 and the MiniAmp™ Plus Thermal Cycler was set to the times outlined in Table 2.28.

Table 2.27: PCR reaction for screening of plasmids.

Reagent	1x (μL)
Forward Primer (U6-F) 5' - GAGGGCCTATTTCCCATGATTC - 3'	0.6
Reverse - AS oligo (Table 2.23)	0.6
dNTP	0.3
10X Dream Buffer	1.5
Lysate	1
Dream Taq	0.08
ddH ₂ O	10.92
Total	15

Table 2.28: Thermocycler conditions for screening PCR.

Temperature	Time	Number of Cycles
95°C	3 min	1
95°C	20 sec	35
58°C	20 sec	
72°C	30 sec	
72°C	7 min	1
4°C	∞	—

PCR products were run at 100 V for 20 minutes on a 1.5% agarose gel made with SafeView and imaged using UV. For clones with a band present, a PCR product clean-up was performed using GenElute™ PCR Clean-Up Kit (Sigma, NA1020).

Products were sent for Sanger sequencing with Eurofins. 30 μ L LB cultures verified by PCR and sequencing put into 5 mL LB with ampicillin overnight. Plasmid DNA extraction was carried out using a PureLink™ HiPure Plasmid Miniprep Kit (ThermoFisher, K210002) and plasmids were quantified using the NanoDrop, and stored at -20°C.

2.14.2. Lentivirus protocol

Day 1: Seeding cells – HEK293T cells were seeded at a concentration of 5×10^5 per well in a 6 well plate.

Day 2: Co-transfect HEK293T cells with packaging plasmids – media was changed one hour prior to transfection. The plasmid DNA (Table 2.29) was mixed with nuclease free water to a volume of 76.9 μ L, and 8.55 μ L of CaCl_2 was added. To a new 2 mL tube, 85.5 μ L 2X HBS solution (280 mM NaCl, 50 mM HEPES, 1.42 mM $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$ (pH 7.05), Adjust pH to 7.05 with 10 M NaOH) was added. The DNA- CaCl_2 solution was added to the HSB and quickly vortexed to mix. The mixture was incubated at room temperature for 15 minutes, then added dropwise to cells and mixed by swirling.

Table 2.29: Plasmids used for lentivirus transfection of HEK293T cells.

<i>Plasmid</i>	<i>DNA concentration</i>	<i>Source</i>
pRSV-Rev	513 ng	Addgene, 12253
pMD2.G	614 ng	Addgene, 12259
pMDGP-Lg/p-RRE	2.222 μ g	Addgene, 12251
Lentiviral vector plasmid	2.222 μ g	NICB
d2EGPF	68.4 ng	NICB

Day 3: Media changed on transfected cells.

Day 4: Harvest virus into 1.5 mL tubes by filtration of media through 0.45 μ m filter and freeze at -80°C.

2.14.3. Transduction

Day 1: A 24-well plate was seeded with the cell line to be edited at a concentration of 1×10^5 cells/mL. Organoids were seeded at a concentration of 1×10^4 cells per well for three days prior to transduction.

Day 2: The lentivirus containing media was thawed quickly. 8 μ g/mL polybrene was added to the growth media required for cells and lentivirus. An equal volume of growth media with polybrene, and lentivirus containing media with polybrene was added to cells, the plate was rocked gently to mix.

Day 4: Selection – media containing puromycin was added to the cells for selection. Cells were maintained in puromycin media until all cells in negative control were dead. CRISPR edited cells grown to confluency and passaged into a 6 well plate. CRISPR-knockout confirmed by TIDE assay.

[This page is intentionally left blank]

Chapter 3. Establishment and characterisation of novel *in vitro* PDAC models – PDX derived organoids and cell line organoids

3.1. Introduction

One of the limitations in understanding of the disease progression and the development of effective treatments for PDAC is due the lack of representative *in vitro* patient tumour models. For over 50 years established 2D cell lines have been the most widely used model for the development and testing of chemotherapeutics [187]. Their importance in cancer research is indisputable, however, their use as a robust clinical model is questionable [287]. Three-dimensional *in vitro* models such as organoids can address the limitations of growing cancer in 2D [218]. These 3D systems can model the physiology, shape, dynamics and cell make-up of the cancer and produce a relevant and highly adaptable model system [219]. Unlike established and primary cell lines, organoids display cell heterogeneity after several passages and organoids can be expanded indefinitely and cryopreserved [219]. However, the use of organoids in cancer research presents a number of challenges, including difficulty obtaining patient samples; specialist training is required for the successful culture of cells; expensive media and reagents, cell maintenance and performing experiments is extremely time consuming in comparison to 2D assays.

Through the development of methodologies for the establishment of PDAC organoids from PDX tumour samples, this study has addressed a number of these issues particularly overcoming issues relating to obtaining patient samples. The development of a matched isogenic 2D primary cell line and recapitulated primary cell line into organoids – a novel model termed cell line organoids (CLOs) – overcomes the high costs associated with organoid culture. The characterisation

and validation of the new models was performed by assessing the stem cell, cancer stem cell and PDAC marker expression, proliferation and response to gemcitabine. RNA-sequencing was performed to compare the transcriptomic signatures of organoids, CLOs, cell lines and the primary PDX tumour. Finally, an *in vivo* study was carried out to assess if the organoid lines maintain tumorigenicity, and to investigate if implanted organoid and CLO are representative the original PDX and patient tumour.

3.1.1. Aims

Aim 1: To develop a biologically relevant model of PDAC and establish and validate a method for developing organoids from PDAC PDX samples.

Aim 2: To ensure that these novel models were representative of PDAC, phenotypic and molecular characterisation of organoids, cell lines and CLOs through assessing drug response, morphology and mRNA and protein expression of markers of interest.

Aim 3: Transcriptomic analysis of PDX derived organoids, CLOs and cell lines was performed to identify if novel models developed from different patients had similar expression signatures.

Aim 4: An *in vivo* study was performed to assess tumorigenicity of novel models and assess if there is differential expression of markers of interest.

3.2. Results

3.2.1. Conditioned media

Organoids require growth factor enriched media to grow in a stem like state. Two of the most important growth factors identified for the self-renewal of stem cells in PDAC organoid culture are Wnt3a and Rspondin1 [226]. While these growth factors are commercially available in recombinant form, it is costly to maintain the high volume of media for large-scale cultures required for standard assays using organoids. To determine the concentration of Rspondin3 and Wnt3a in L-WRN conditioned media a dual luciferase reporter assay was performed, where a beta-catenin promoter is upstream of the luciferase gene, and transcription is activated by Wnt3a or Rspondin3.

As per the methodology described in Section 2.2, a cell line that was generated by the Stappenbeck lab which secretes the factors Wnt3A, R-spondin3 and noggin (L-WRN) into the medium was grown, and conditioned media obtained was tested for use in organoid culture media using a TOPFlash assay [288,289]. Figure 3.1 shows that the concentration of Rspondin3 and Wnt3a in the conditioned media, made from L-WRN cells from Passage 3 to Passage 6, and in a spiked recombinant Wnt3a/Rspondin3 control, which contains the minimum required levels of proteins for organoid growth. The levels of proteins in the conditioned media is higher than the minimum required levels for the organoids to grow, as indicated by the spiked recombinant Wnt3a/Rspondin3 control. To determine the stability of the conditioned media during storage, the media was maintained at 4°C for one week and the assay was repeated. Results showed that there is a significant reduction in levels of Wnt3a and Rspondin3 in the media of all samples, indicating that the proteins are not stable long term at 4°C. Therefore, the conditioned media was frozen in aliquots at -20°C.

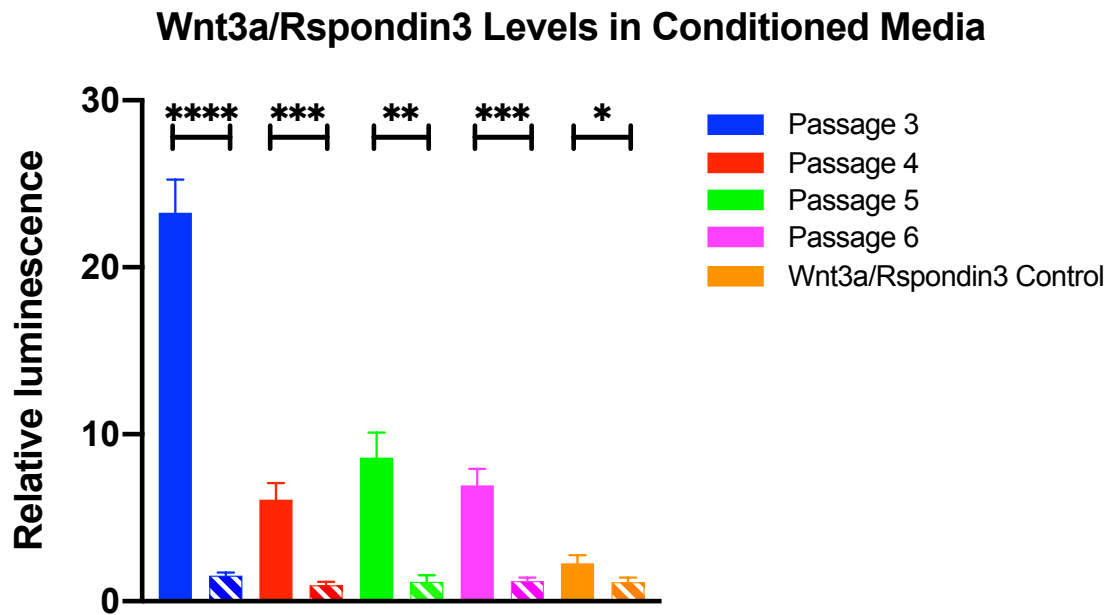


Figure 3.1: TOPFlash Assay – a dual luciferase reporter assay to determine the concentration of WNT3a/Rspondin3 in conditioned media required for the growth of organoids. Media generated from four different passages (P3-P6), and a spiked recombinant Wnt3a/Rspondin3 control was tested on day 0 (freshly thawed) indicated by the solid colours and day 7 (one week at 4°C) indicated by the hatched boxes. The spiked recombinant Wnt3a/Rspondin3 control contains the minimum level of proteins required for organoid growth. A Student's *t*-test was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

3.2.2. Development of PDX organoids and isogenic matched primary cell lines

These protocols were developed as per the methodology described in Section 2.3. In this study, the protocol from Boj *et al.* [226] was modified to develop a novel protocol for the establishment organoids from PDX tumours (Figure 3.2). A mouse cell depletion kit was used to remove any infiltrating mouse cells from the PDX tumour sample. Organoids were plated on polyHEMA-coated plates to prevent cells attaching to the cell culture plate.

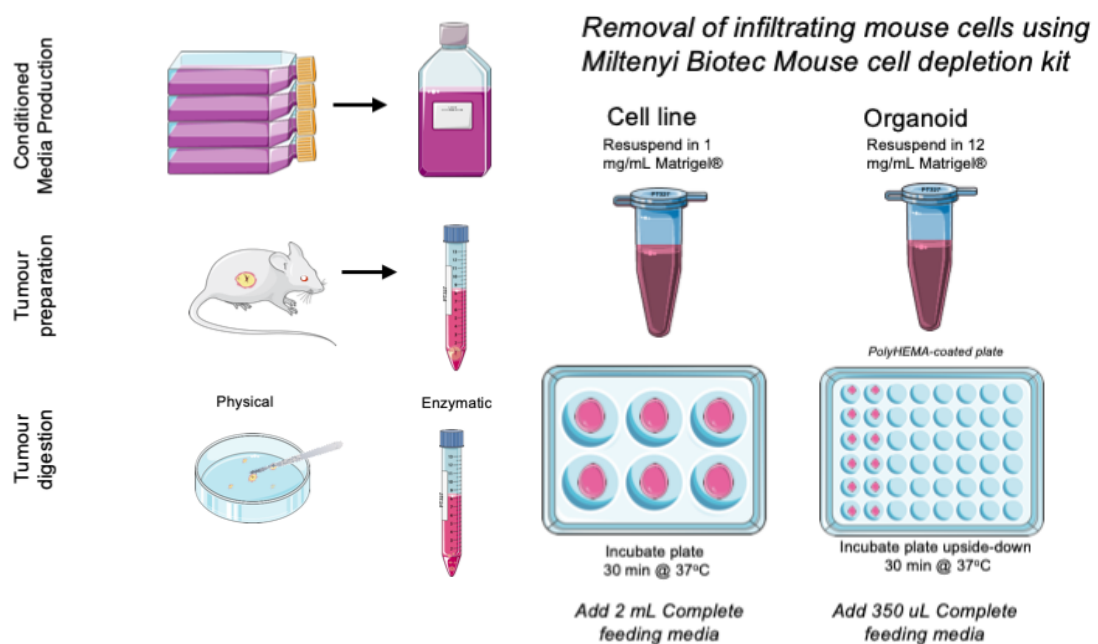


Figure 3.2: Schematic diagram protocol for generating PDAC organoids and isogenic matched 2D primary cell lines from a single PDX tumour sample.

PDAC primary cell lines were simultaneously developed by plating the tumour cells in lower concentration ECM gel on adherent wells, which allowed the cells to attach and proliferate, as shown in Figure 3.3.

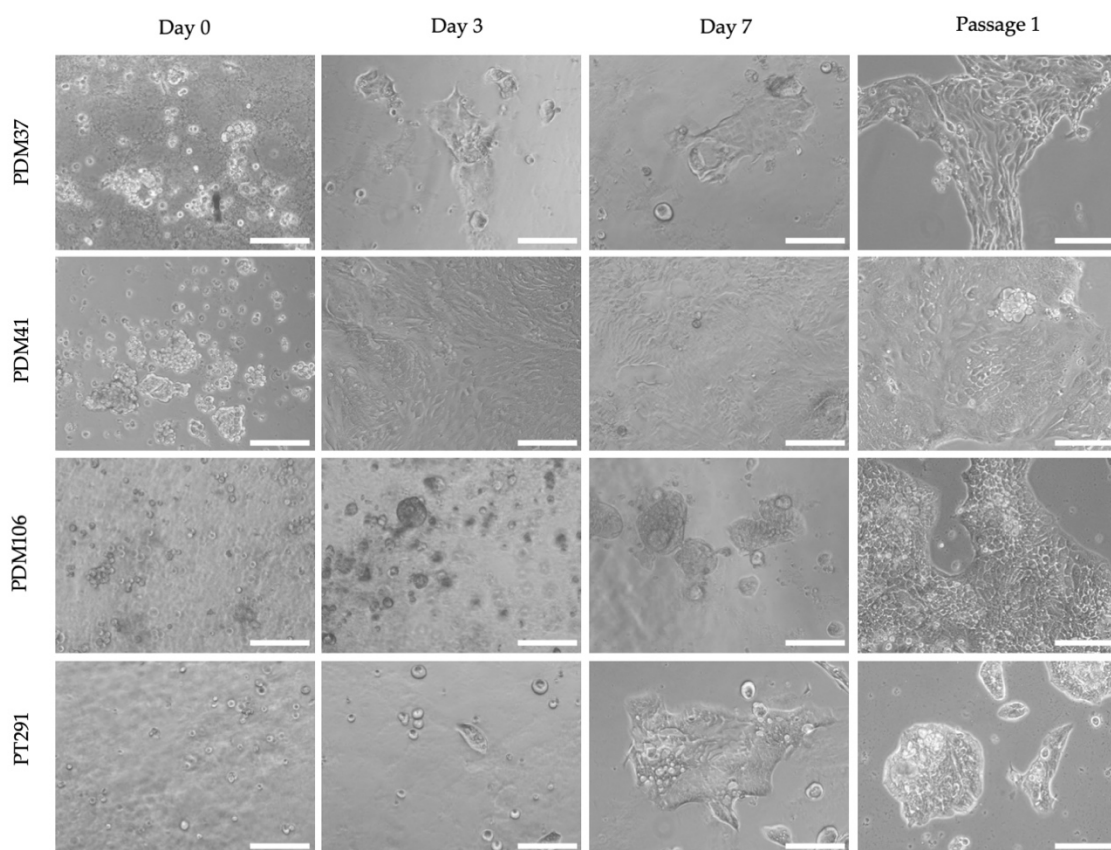


Figure 3.3: Brightfield images taken throughout the generation of primary cell lines from organoids/tumour cells taken at day 0, day 3, day 7 and passage 1 of newly generated cell lines PDM37-CL, PDM41-CL, PDM106-CL, and PT291-CL. Images taken at 100X magnification, scale bars 400 μ m.

3.2.3. Confirmation of primary cell lines and organoid culture derived from human PDAC PDX tumour cells

A two-step, nested PCR method was used to confirm the absence of mouse cells in the primary cell lines and organoids in PT291 and PT127 [290]. Universal primers amplified a region common in both the human and mouse mitochondrial genomes. Mouse specific primers were used to identify if mouse-contaminating cells were present in the samples. PCR results, as shown in Figure 3.4 confirmed that no contaminating mouse mitochondrial DNA is present in the PDX derived organoids, 2D primary cell lines or CLOs, therefore confirming human PDAC origin.

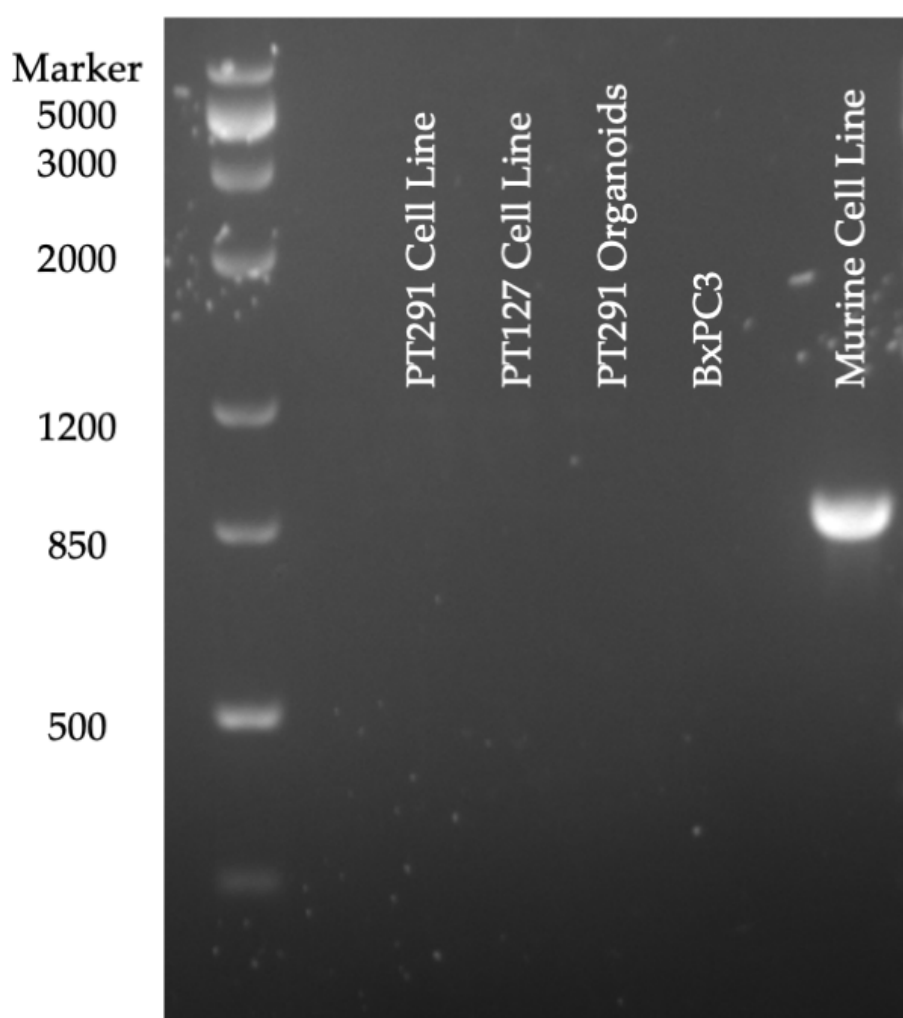


Figure 3.4: Electrophoresis gels of samples PT291 primary cell line, PT127 primary cell line, PT291 organoids, BxPC3 (negative control) and Murine cell line C2C12 (positive control) for products of nested PCR using mouse specific primers.

3.2.4. Generation of CLOs as novel models to study PDAC

The 2D primary cells were cultured for at least two passages then recapitulated as 3D organoid cultures – CLOs. These CLOs were cultured using the same methods as used for the growth of organoids as shown in Figure 3.5. CLOs were cultured for at least two passages prior to use in assays to ensure the removal of cells which did not form CLOs.

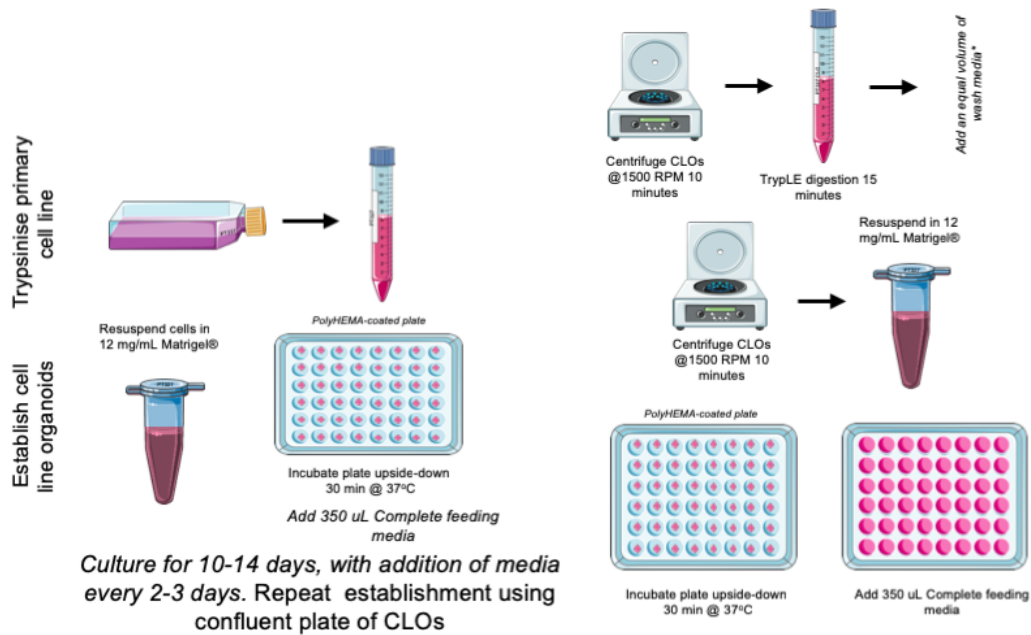


Figure 3.5: Schematic diagram protocol for generating CLOs from 2D primary cell lines.

3.2.5. Characterisation of novel models

3.2.5.1. Morphology of organoids

The growth and development of all organoid lines was monitored over a 10-day period, with representative images taken on day 0, day 3, day 7 and day 10 (Figure 3.6).

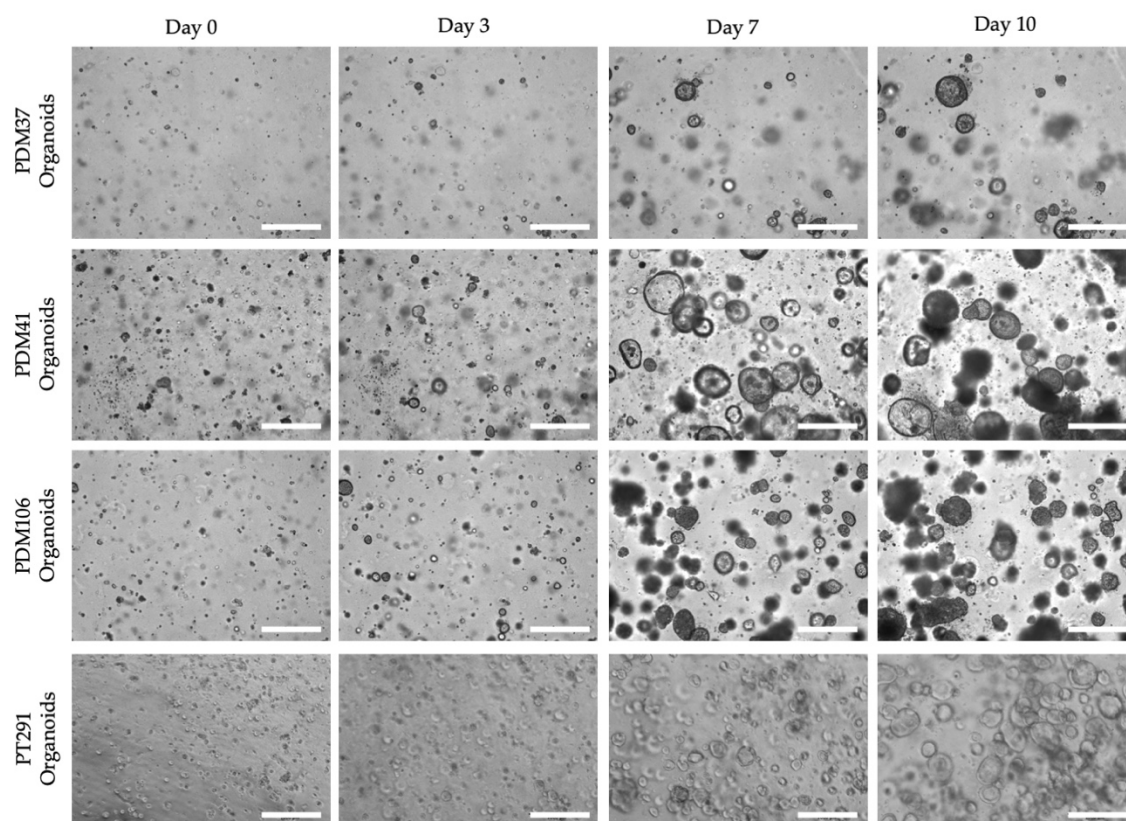


Figure 3.6: Brightfield images at Day 0, Day 3, Day 7, Day 10 of PDM37, PDM41, PDM106 and PT291 organoids. Images taken at 100X magnification. Scale bar 400 μm .

3.2.5.2. Morphology of primary cell lines

Isogenic cell lines generated from organoids were imaged after passaging (Figure 3.7).

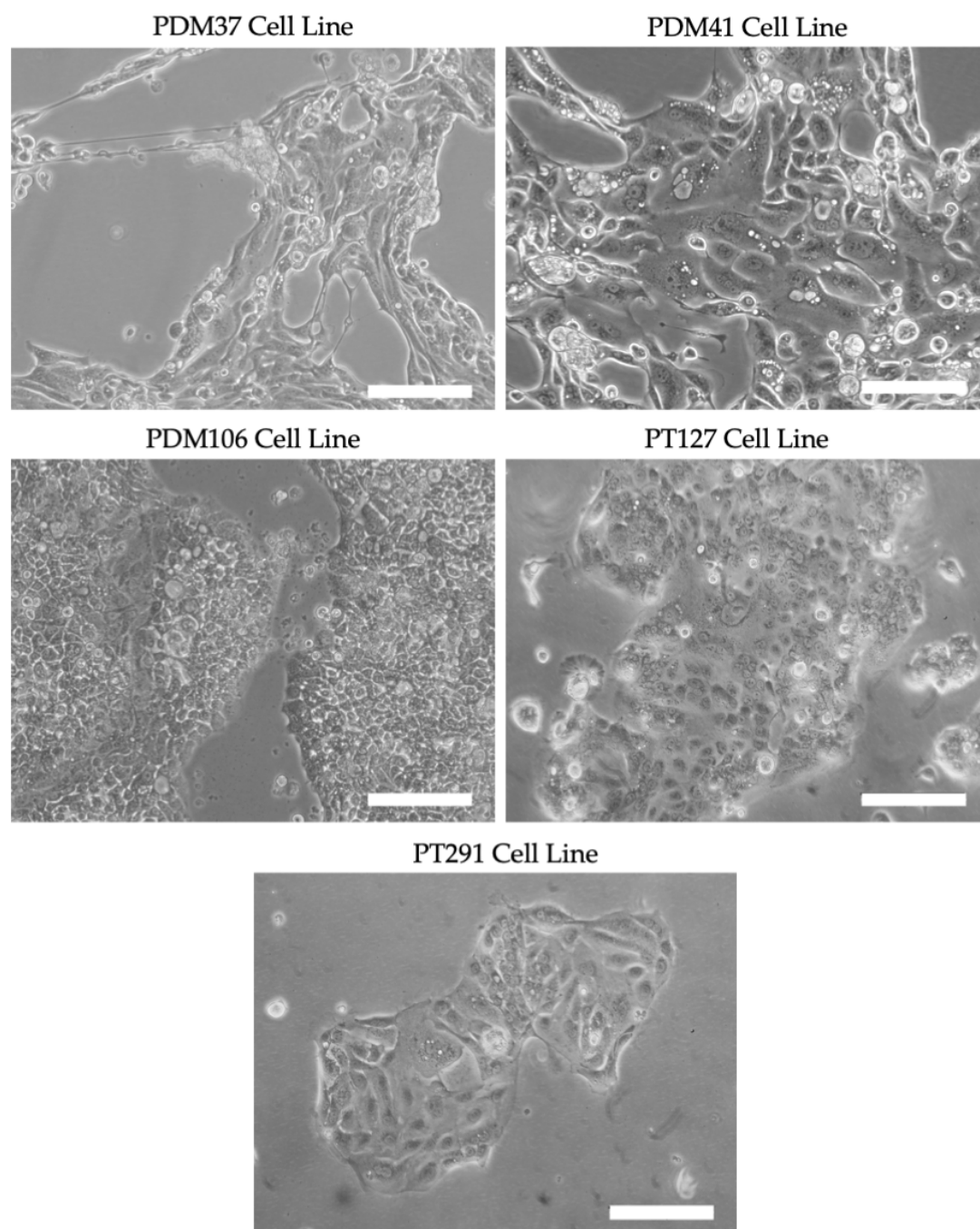


Figure 3.7: Brightfield imaging of newly generated cell lines PDM37-CL, PDM41-CL, PDM106-CL, and PT291-CL, and primary cell line PT127-CL, which was previously generated in the NICB. Images taken at 100X magnification. Scale bars 400 μ m.

3.2.5.3. Morphology of CLOs

The newly generated CLOs develop rapidly and resemble the matched organoid lines, maintaining a similar size and the morphology of the original cell lines. from day 0 to day 10 (Figure 3.8).

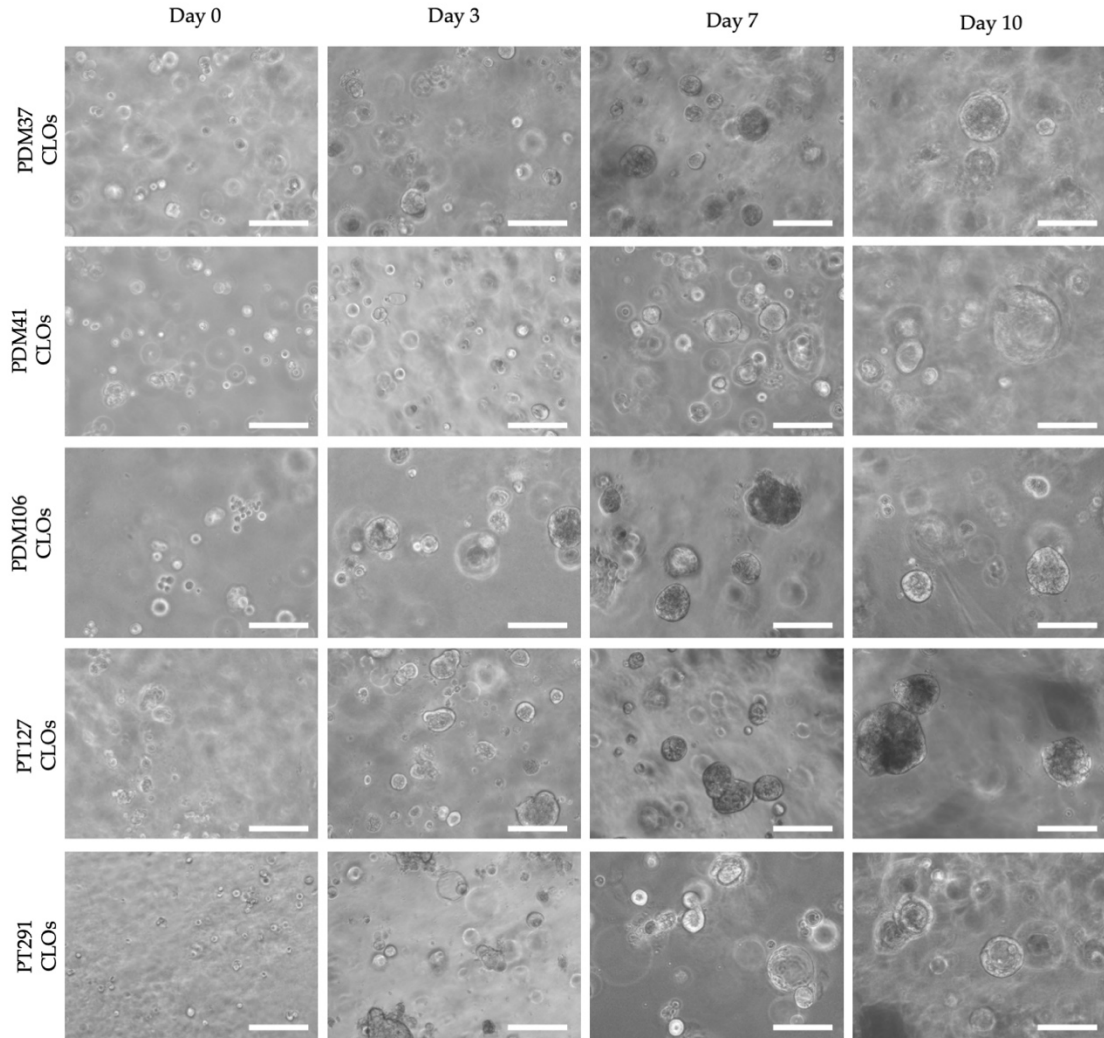


Figure 3.8: Brightfield images of the generation of CLOs at day 0, day 3, day 7 and day 10 of PDM37-CLO, PDM41-CLO, PDM106-CLO, PT127-CLO and PT291-CLO.

Images taken at 100X magnification. Scale bars 400 μ m.

3.2.5.4. Organoid mutation status

Mutational analysis of the common PDAC mutations in genes *KRAS*, *CDKN2A*, *SMAD4* and *TP53* was performed by Sanger sequencing, as per the methodology described in Section 2.4. All models were found to have missense mutations in *KRAS* – PDM37 and PT291 display the mutation G12D, the most common mutation in PDAC. PDM41 and PDM106 have a G12V *KRAS* mutation, and PT127 has a G13D *KRAS* mutation (Table 3.1). Only PDM37, had a mutation in *SMAD4* Q442*, which results in a stop, mutations were not identified in codon 1 or 2 of the other samples (Table 3.2). Mutations were also observed in *CDKN2A*, PDM37 had a deletion in codons R87-D92, all other samples had a *29G>C mutation in the 3'UTR of the *CDKN2A* (Table 3.3). *TP53* mutations were observed in all samples, PDM41, PT127 and PT291 had multiple mutations in the gene (Table 3.4).

Table 3.1: Mutations occurring in the *KRAS* gene in patient samples PDM37, PDM41, PDM106, PT127 and PT291.

<i>Patient</i>	<i>Mutation</i>	<i>Consequence</i>
PDM37	G12D ¹	Missense variant
PDM41	G12V	Missense variant
PDM106	G12V	Missense variant
PT127	G13D	Missense variant
PT291	G12D	Missense variant

¹Data obtained from HCMI Catalogue (Model: HCM-CSHL-0090-C25)

Table 3.2: Mutations occurring in the *SMAD4* gene in patient PDM37.

<i>Patient</i>	<i>Mutation</i>	<i>Consequence</i>
PDM37	Q442* ¹	Stop

¹Data obtained from HCMI Catalogue (Model: HCM-CSHL-0090-C25)

Table 3.3: Mutations occurring in the CDKN2A gene in patient samples PDM37, PDM41, PDM106, PT127 and PT291.

Patient	Mutation	Consequence
PDM37	R87_D92 del ¹	Inframe deletion
PDM41	*29G>C	SNV
PDM106	*29G>C	SNV
PT127	*29G>C	SNV
PT291	*29G>C	SNV

¹Data obtained from HCM Catalogue (Model: HCM-CSHL-0090-C25)

Table 3.4: Mutations occurring in the TP53 gene in patient samples PDM37, PDM41, PDM106, PT127 and PT291.

Patient	Mutation	Consequence
PDM37	K139R fs*311 ¹	Frameshift Variant
PDM41	P47R	Missense variant
	599delA	Frameshift deletion
	M237I	Missense variant
PDM106	P47R	Missense variant
PT127	P72R	Missense variant
	599delA	Frameshift deletion
PT291	599delA	Frameshift deletion
	710_711insA	Insertion

¹Data obtained from HCM Catalogue (Model: HCM-CSHL-0090-C25)

3.2.6. Proliferation assays

3.2.6.1. Proliferation cell lines

Proliferation assays were carried out using the Incucyte Live Cell Imaging System to assess the growth of the newly generated cell lines (Figure 3.9) as per the methodology described in Section 2.5.1. The optimal seeding density for each cell line is indicated in Table 3.5. Cell lines PDM41 and PT127 reached 100% confluency after 4 days and 9 days respectively. After 10 days, cell lines PDM37, PDM106 and PT291 only achieved a maximum confluence of $47 \pm 5.5\%$ and $54 \pm 6\%$ respectively.

Table 3.5: Optimal cell seeding density for cell lines for assays carried out in 96 well plate.

Cell line	Cell seeding density
PDM37-CL	5000/well
PDM41-CL	5000/well
PDM106-CL	5000/well
PT127-CL	1000/well
PT291-CL	5000/well

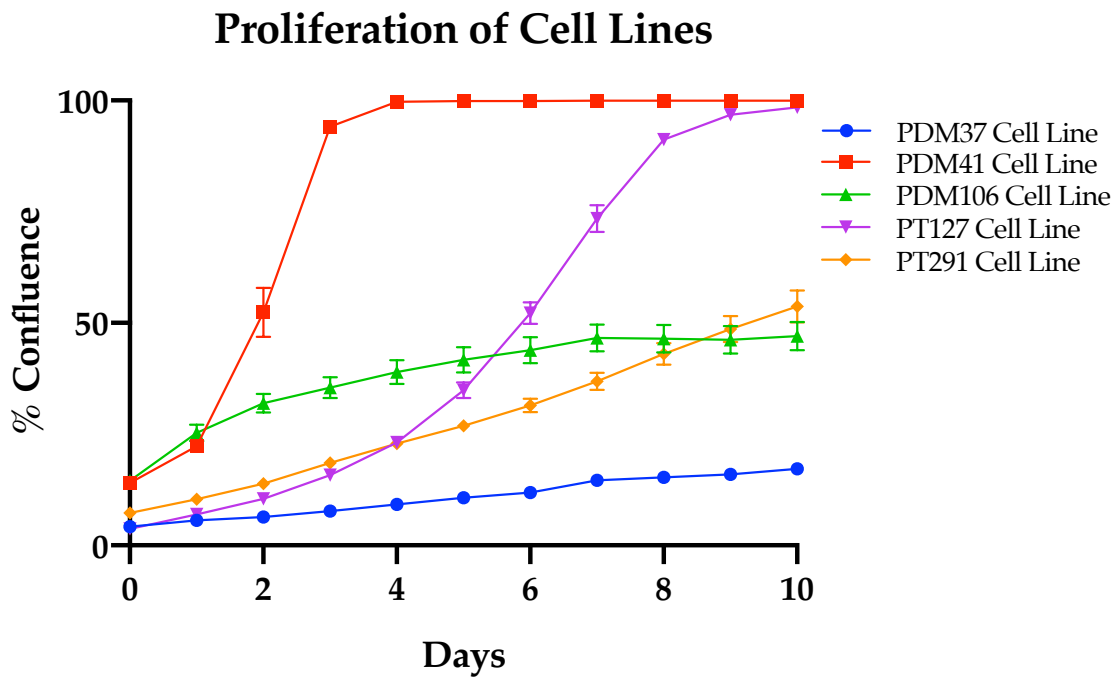


Figure 3.9: Proliferation of cell lines PDM37-CL, PDM41-CL, PDM106-CL, PT127-CL and PT291-CL over 10 days. Cell growth is determined by percentage confluence. Error bars represent standard deviation of biological triplicate experiments.

3.2.6.2. Proliferation of organoids and CLOs

Proliferation assays were carried out using the Incucyte to assess the growth of organoid and CLO models over a period of 9 days (Figure 3.10 - Figure 3.14) as per the methodology described in Section 2.5.2. Proliferation was quantified by measuring the total area of organoids at each timepoint and was calculated relative to the first timepoint. Table 3.6 shows the final fold-change total area. Organoid and matched CLO models had similar proliferation rates in PDM37, PDM106 and PT291. PDM41 organoids had a fold-change of 2.3 ± 0.8 , whereas the matched CLOs which had a fold-change of 4 ± 1 relative to the first timepoint.

Table 3.6: Fold-change of total area cover by organoid and CLO models compared to first timepoint after 9 days.

<i>Patient</i>	<i>Organoid</i>	<i>CLO</i>
PDM37	4.3 ± 1.6	4.5 ± 1.5
PDM41	2.3 ± 0.8	4 ± 1
PDM106	6 ± 3	6.7 ± 1.5
PT127	—	2.6 ± 0.8
PT291	3.2 ± 1	2.1 ± 0.2

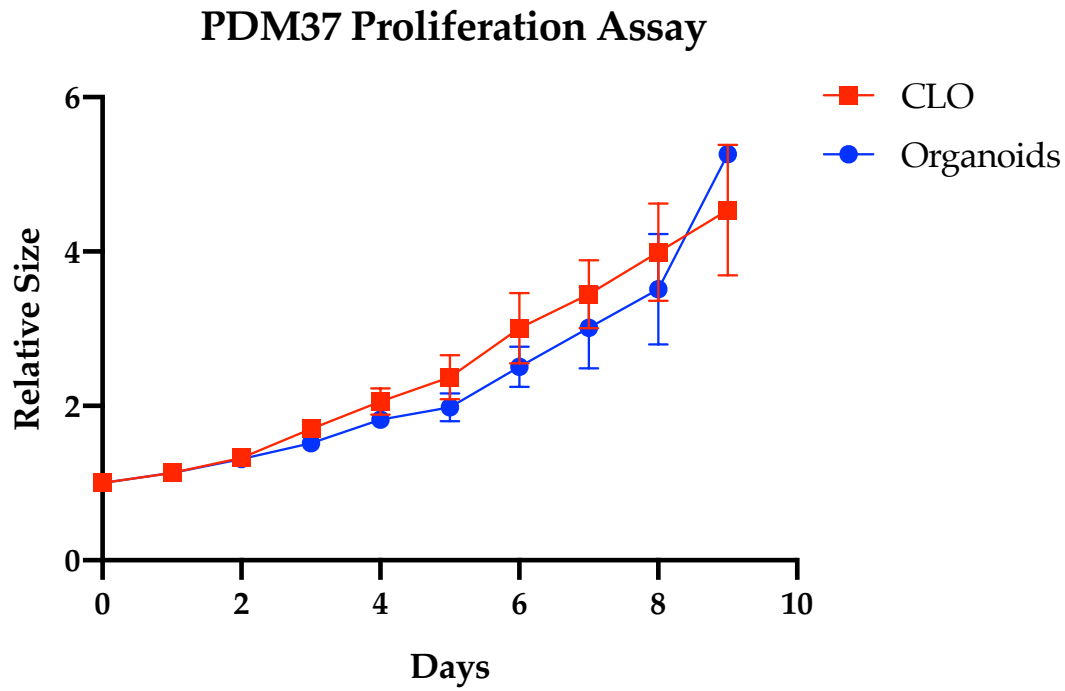


Figure 3.10: Proliferation of PDM37 CLO and organoids, over a 9-day period. Cell growth is determined total area of organoids, relative to day 0. Error bars represent standard error the of mean of biological triplicate experiments.

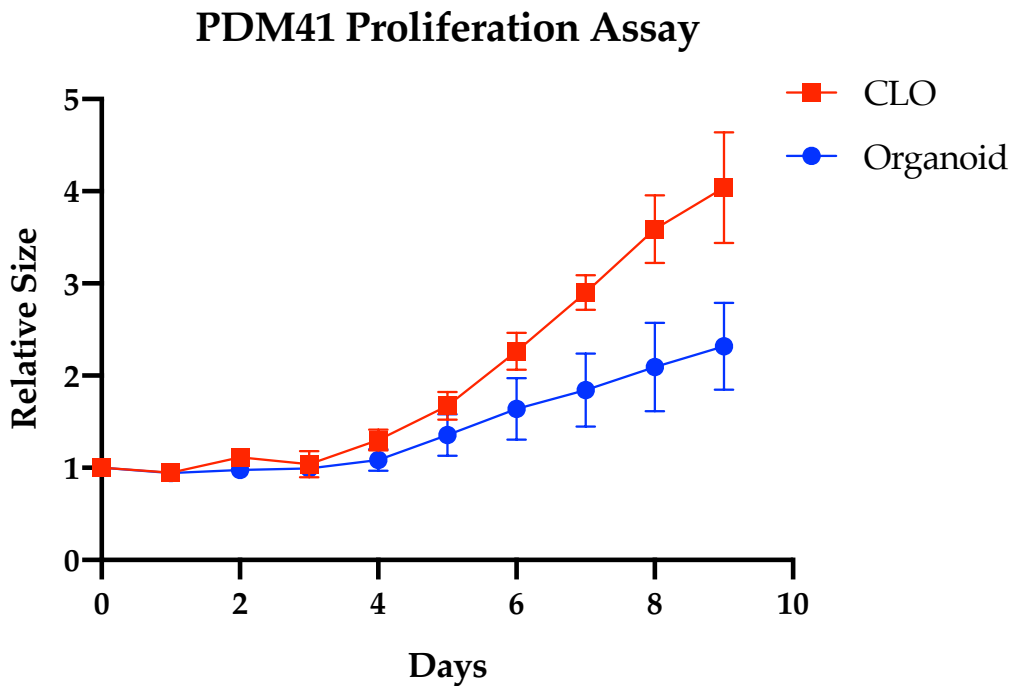


Figure 3.11: Proliferation of PDM41 CLO and organoids, over a 9-day period. Cell growth is determined total area of organoids, relative to day 0. Error bars represent standard error the of mean of biological triplicate experiments.

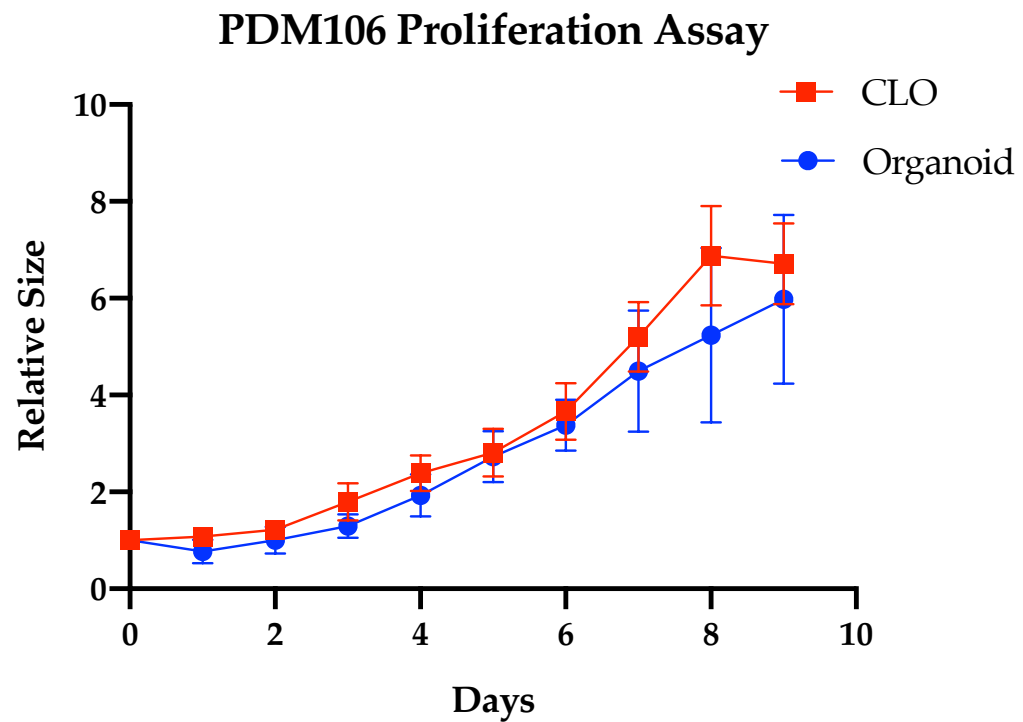


Figure 3.12: Proliferation of PDM106 CLO and organoids, over a 9-day period. Cell growth is determined total area of organoids, relative to day 0. Error bars represent standard error the of mean of biological triplicate experiments.

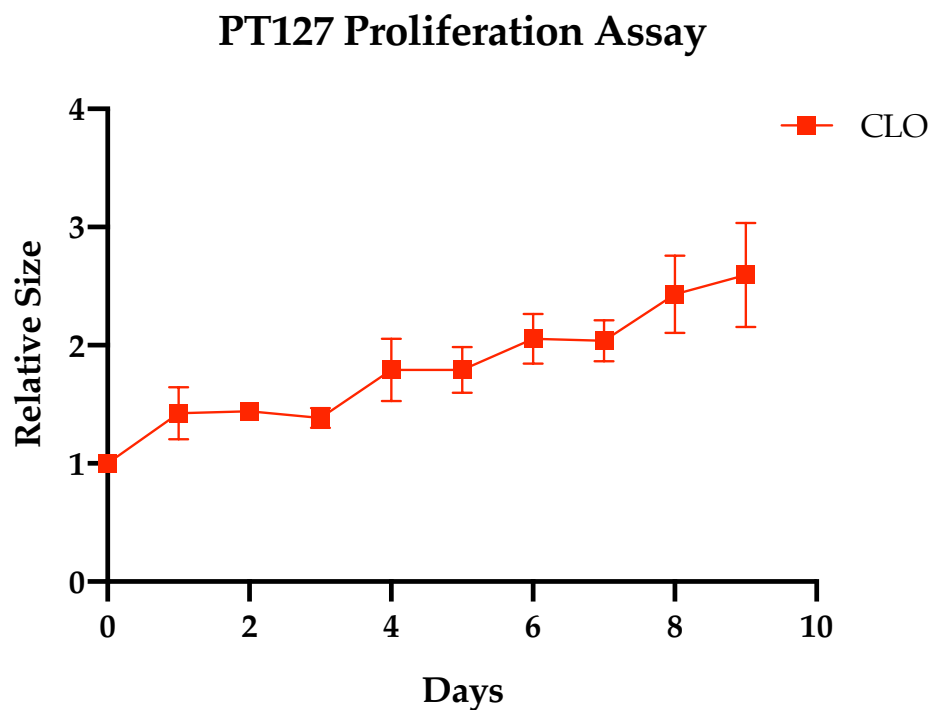


Figure 3.13: Proliferation of PT127 CLOs, over a 9-day period. Cell growth is determined total area of organoids, relative to day 0. Error bars standard error the of mean biological triplicate experiments.

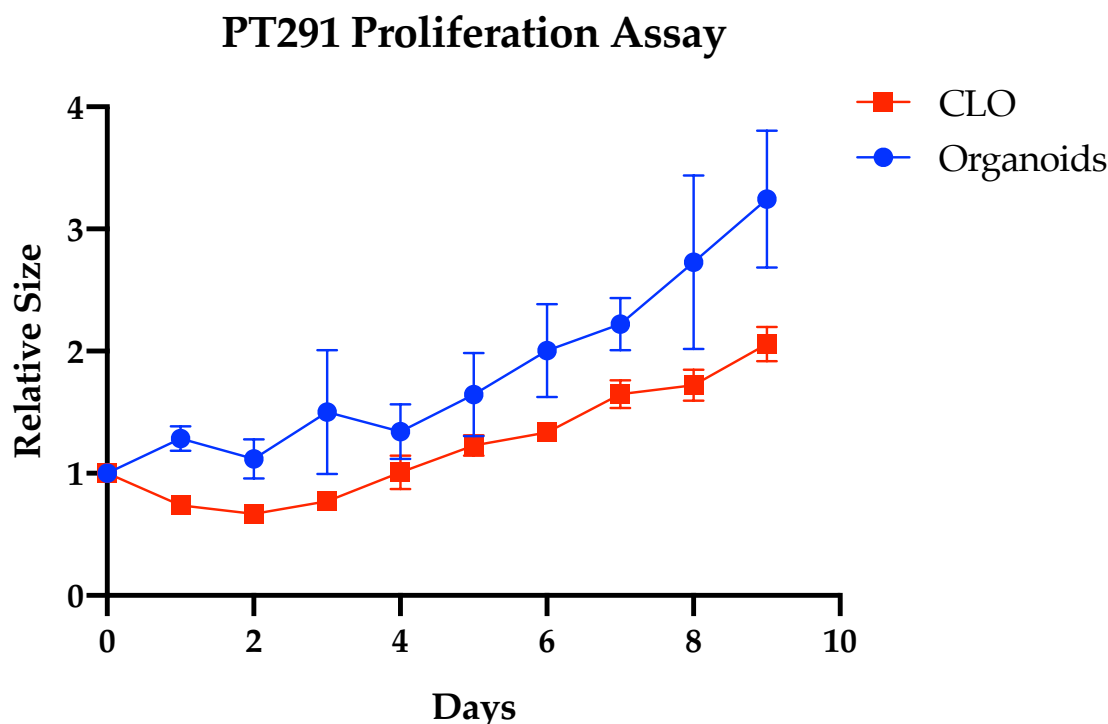


Figure 3.14: Proliferation of PT291 CLOs and organoids, over a 9-day period. Cell growth is determined total area of organoids, relative to day 0. Error bars represent standard error the of mean of biological triplicate experiments.

3.2.7. Gemcitabine response assays

Assays were performed to assess the response to gemcitabine treatment in all organoids, CLOs and cell lines, as per the methodology described in Section 2.6. The organoids and CLOs for PDM37, PDM41 and PT291 had IC_{50} values <25 ng/mL, PDM106 organoids and CLOs had an IC_{50} of <50 ng/mL (Figure 3.15 - Figure 3.18). PT127 CLOs did not achieve an IC_{50} within the concentrations used in this experiment (Figure 3.19). In comparison, isogenic matched cell lines PDM37-CL, PDM106-CL, PT127-CL and PT291-CL had an IC_{50} >800 ng/mL, with only PDM41-CL achieving an IC_{50} within the range set out in this experiment (Figure 3.15 - Figure 3.19).

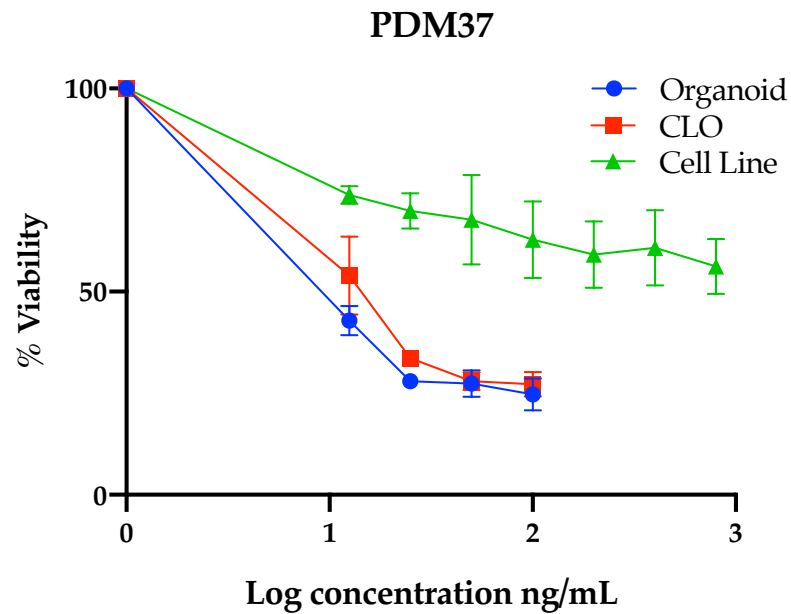


Figure 3.15: Growth inhibition of PDM37 organoids (blue), CLOs (red) and cell lines (green) after 5 days of treatment with gemcitabine. Viability of organoids and CLOs was quantified by Cell Titre Glo, cell lines were quantified by acid phosphatase. Percentage growth was calculated relative to untreated control cells. Error bars represent standard error the of mean of biological triplicate experiments.

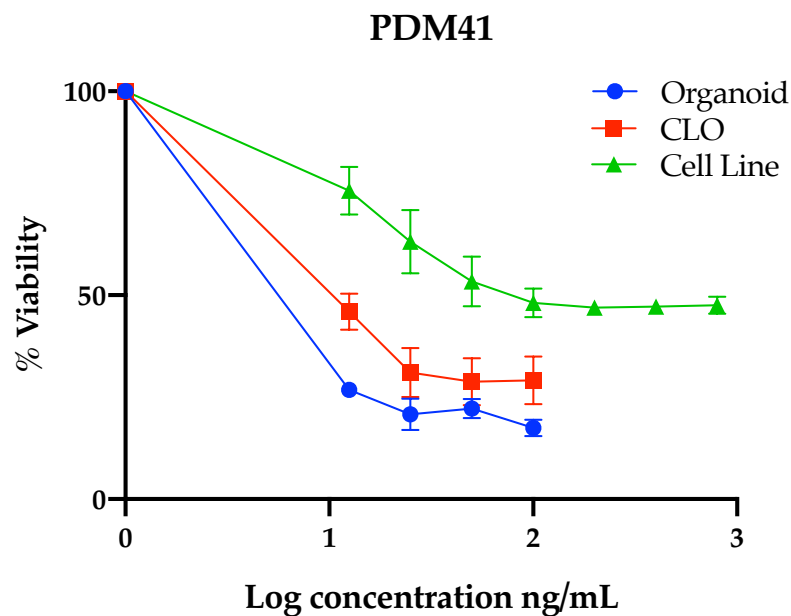


Figure 3.16: Growth inhibition of PDM41 organoids (blue), CLOs (red) and cell lines (green) after 5 days of treatment with gemcitabine. Viability of organoids and CLOs was quantified by Cell Titre Glo, cell lines were quantified by acid phosphatase. Percentage growth was calculated relative to untreated control cells. Error bars represent standard error the of mean of biological triplicate experiments.

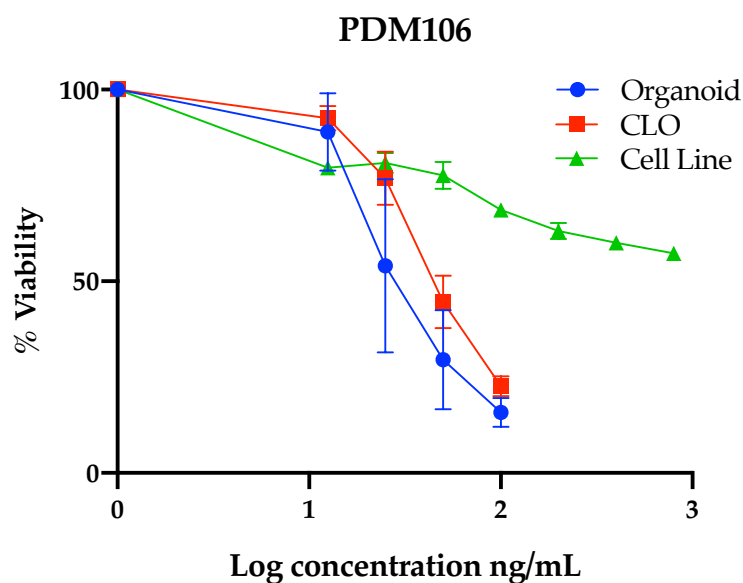


Figure 3.17: Growth inhibition of PDM106 organoids (blue), CLOs (red) and cell lines (green) after 5 days of treatment with gemcitabine. Viability of organoids and CLOs was quantified by Cell Titre Glo, cell lines were quantified by acid phosphatase. Percentage growth was calculated relative to untreated control cells. Error bars represent standard error the of mean of biological triplicate experiments.

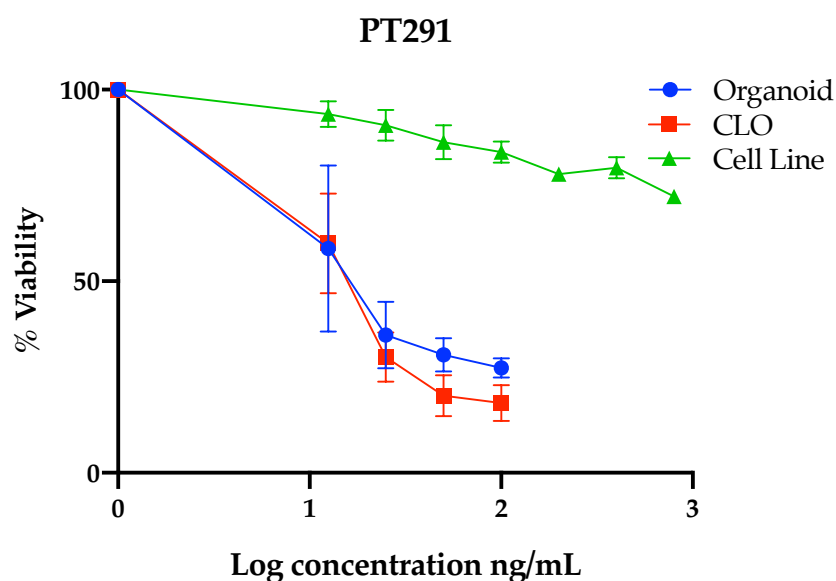


Figure 3.18: Growth inhibition of PT291 organoids (blue), CLOs (red) and cell lines (green) after 5 days of treatment with gemcitabine. Viability of organoids and CLOs was quantified by Cell Titre Glo, cell lines were quantified by acid phosphatase. Percentage growth was calculated relative to untreated control cells. Error bars represent standard error the of mean of biological triplicate experiments.

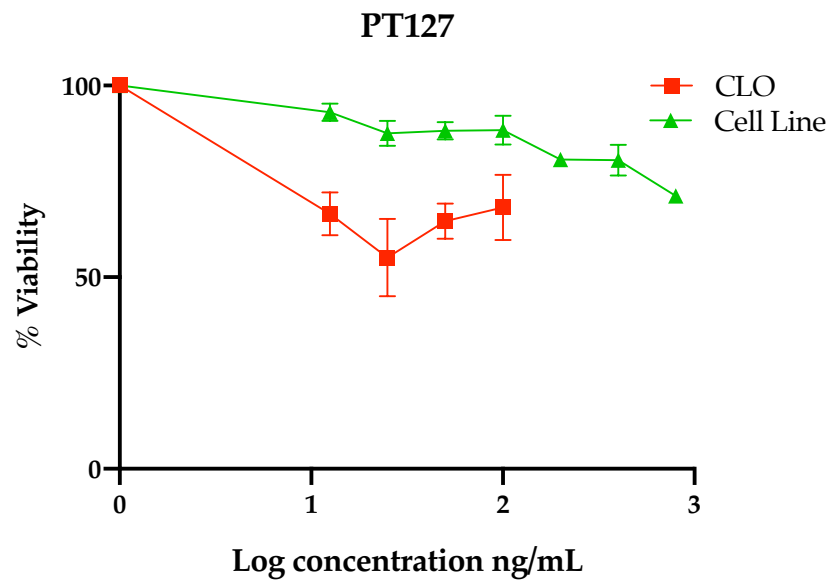


Figure 3.19: Growth inhibition of PT127 CLOs (red) and cell lines (green) after 5 days of treatment with gemcitabine. Viability of organoids and CLOs was quantified by Cell Titre Glo, cell lines were quantified by acid phosphatase. Percentage growth was calculated relative to untreated control cells. Error bars represent standard error the of mean of biological triplicate experiments.

3.2.8. Stem cell marker expression

To test for the presence of stem cell markers responsible for self-renewal and propagation in the cell lines, organoids and CLOs, RT-qPCR was performed as per methodology described in Section 2.9.2. The markers tested were *NANOG*, *OCT4* and *SOX2* which have roles in the self-renewal and maintenance of pluripotency in undifferentiated stem cells [291–293]. In PDM37, expression of *NANOG* and *OCT4* was increased in CLOs and organoids compared to the matched cell line, however, expression of *SOX2* was decreased in the organoids and matched CLOs (Figure 3.20). PDM41 models followed a similar trend, with increased expression of *NANOG* and *OCT4* in the CLO and organoid models compared to the cell line, and reduced expression of *SOX2* in CLO models, however, increased expression of *SOX2* was observed in PDM41 organoid models compared to the cell line. Expression of markers *OCT4*, *NANOG* and *SOX2* were increased in PDM106 and PT291 organoids and CLOs compared to the 2D primary cell lines (Figure 3.23 and Figure 3.24). PT127 PDX tumour had a similar expression profile to the PT127 CLOs compared to its 2D primary cell line (Figure 3.23).

PDM37 Stem Cell Marker Expression

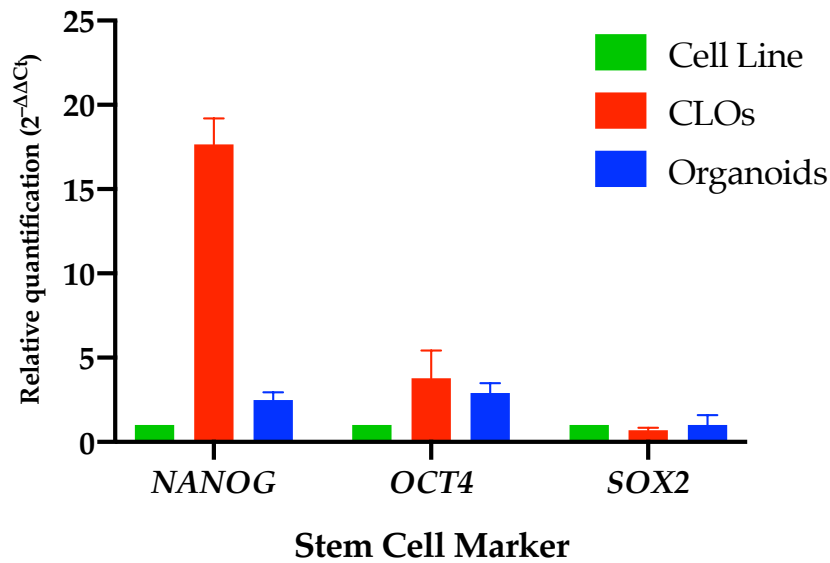


Figure 3.20: Relative quantification of stem cell markers NANOG, OCT4 and SOX2 assessed using RT-qPCR in PDM37 cell line, organoids and CLOs. Expression is shown relative to the expression in matched cell line. 18S rRNA used as endogenous control. Error bars represent standard deviation of biological triplicate experiments.

PDM41 Stem Cell Marker Expression

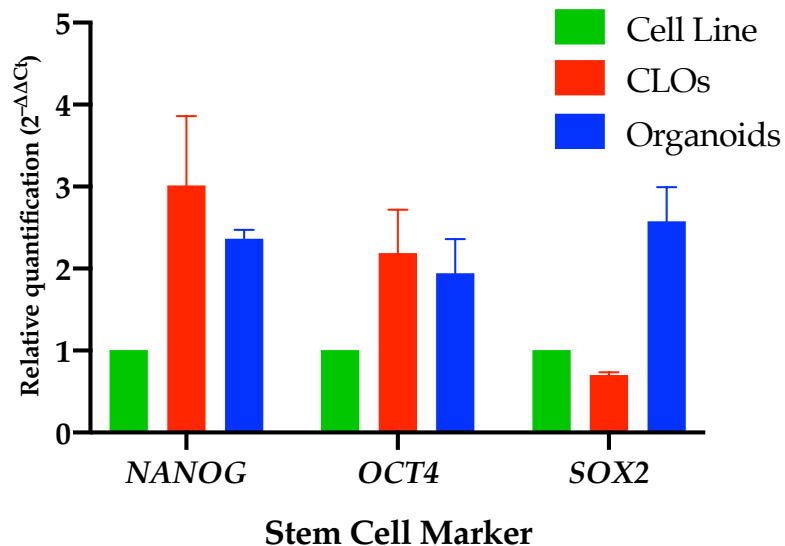


Figure 3.21: Relative quantification of stem cell markers NANOG, OCT4 and SOX2 assessed using RT-qPCR in PDM41 cell line, organoids and CLOs. Expression is shown relative to the expression in matched cell line. 18S rRNA used as endogenous control. Error bars represent standard deviation of biological triplicate experiments.

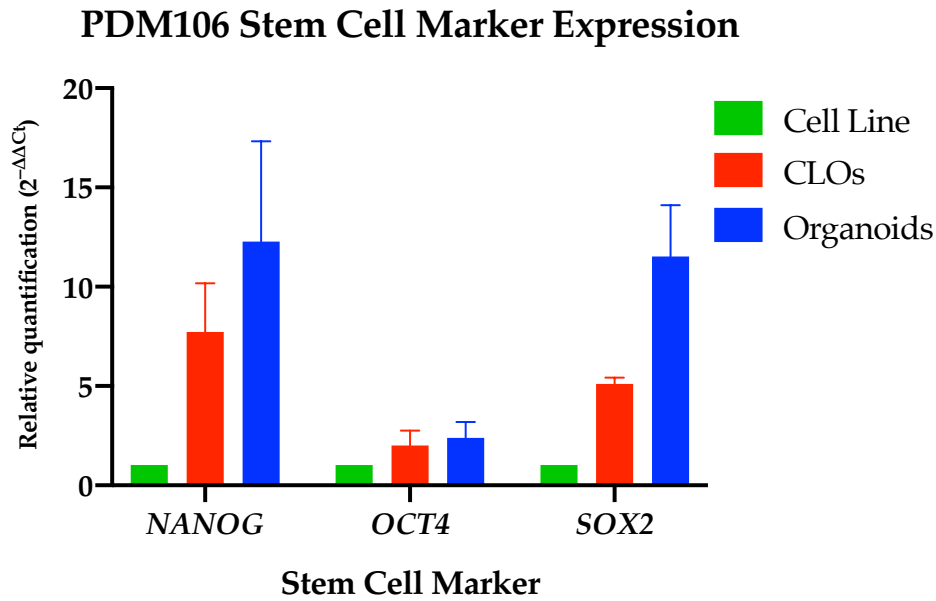


Figure 3.22: Relative quantification of stem cell markers NANOG, OCT4 and SOX2 assessed using RT-qPCR in PDM106 cell line, organoids and CLOs. Expression is shown relative to the expression in matched cell line. 18S rRNA used as endogenous control. Error bars represent standard deviation of biological triplicate experiments.

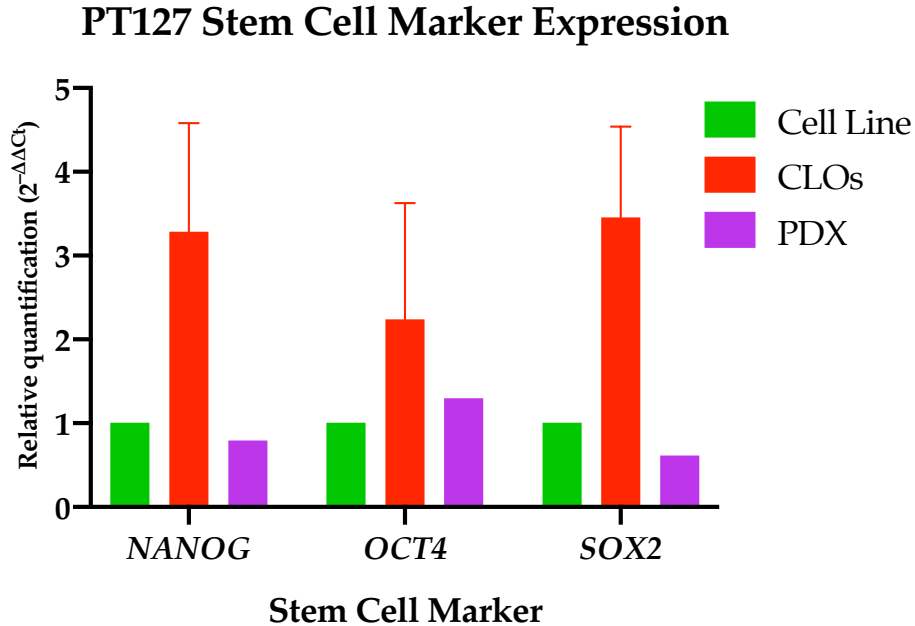


Figure 3.23: Relative quantification of stem cell markers NANOG, OCT4 and SOX2 assessed using RT-qPCR in PT127 cell line, CLO and PDX models. Expression is shown relative to the expression in matched cell line. 18S rRNA used as endogenous control. Error bars represent standard deviation of biological triplicate experiments.

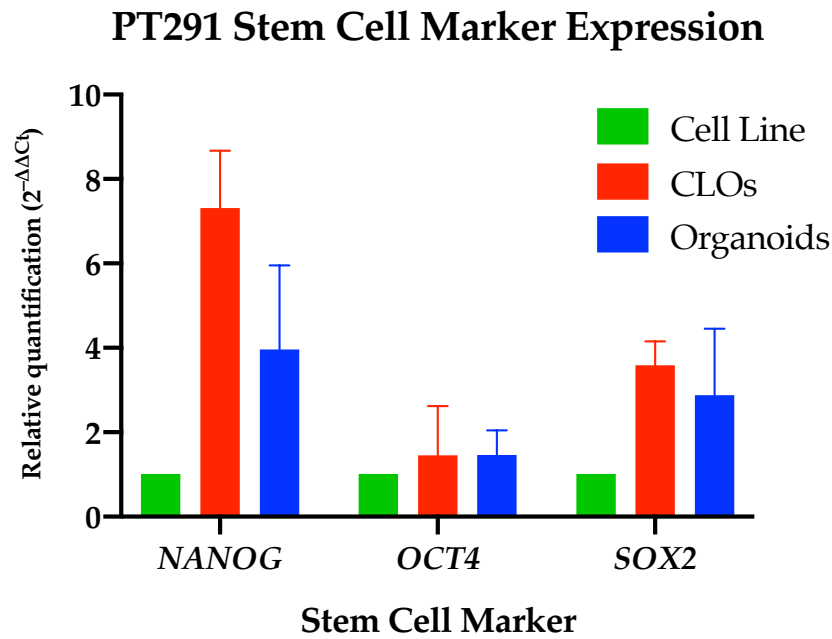


Figure 3.24: Relative quantification of stem cell markers NANOG, OCT4 and SOX2 assessed using RT-qPCR in PT291 cell line, organoids and CLOs. Expression is shown relative to the expression in matched cell line. 18S rRNA used as endogenous control. Error bars represent standard deviation of biological triplicate experiments.

3.2.9. Immunofluorescent analysis of PDAC, stem cell and cancer stem cell markers

Immunofluorescence was performed to determine whether the expression profile of key pancreatic cancer markers (MASPIN, PDX1, PANCK,), stem cell markers (CXCR4) and cancer stem cell markers (ALDH1A1, HCAM, EpCAM) were maintained in cell lines, organoids and CLOs was performed as per methodology described in Section 2.7. PDM37 cell models were excluded as there were not sufficient cells to make cell blocks. ALDH1A1 expression is similar in cell line and CLO models, and had increased expression in organoid models (Figure 3.25). CXCR4 expression is increased in organoid and CLO models compared to cell lines (Figure 3.26). In PDM106 and PT291, EpCAM expression is similar in all models, and in PDM41 and PT127, increased expression was observed in CLOs compared to organoid and cell line models (Figure 3.27). The trend of HCAM expression was the same in all groups of models, with increased expression in the CLO and organoid models in comparison to the cell lines (Figure 3.28). MASPIN, a serine protease inhibitor, has increased expression in organoid and CLO models

compared to the cell lines, with the exception of PDM106, where low expression was observed in the CLO (Figure 3.29). PANCK expression is similar in all cell models in PDM106 and PT291, and in PDM41 and PT127, increased expression in CLOs and organoids compared to the matched cell line was observed (Figure 3.30). Increased expression of PDX1 is observed in the cell line and CLO compared to the matched organoids in PDM106, PT127 and PT291 (Figure 3.31).

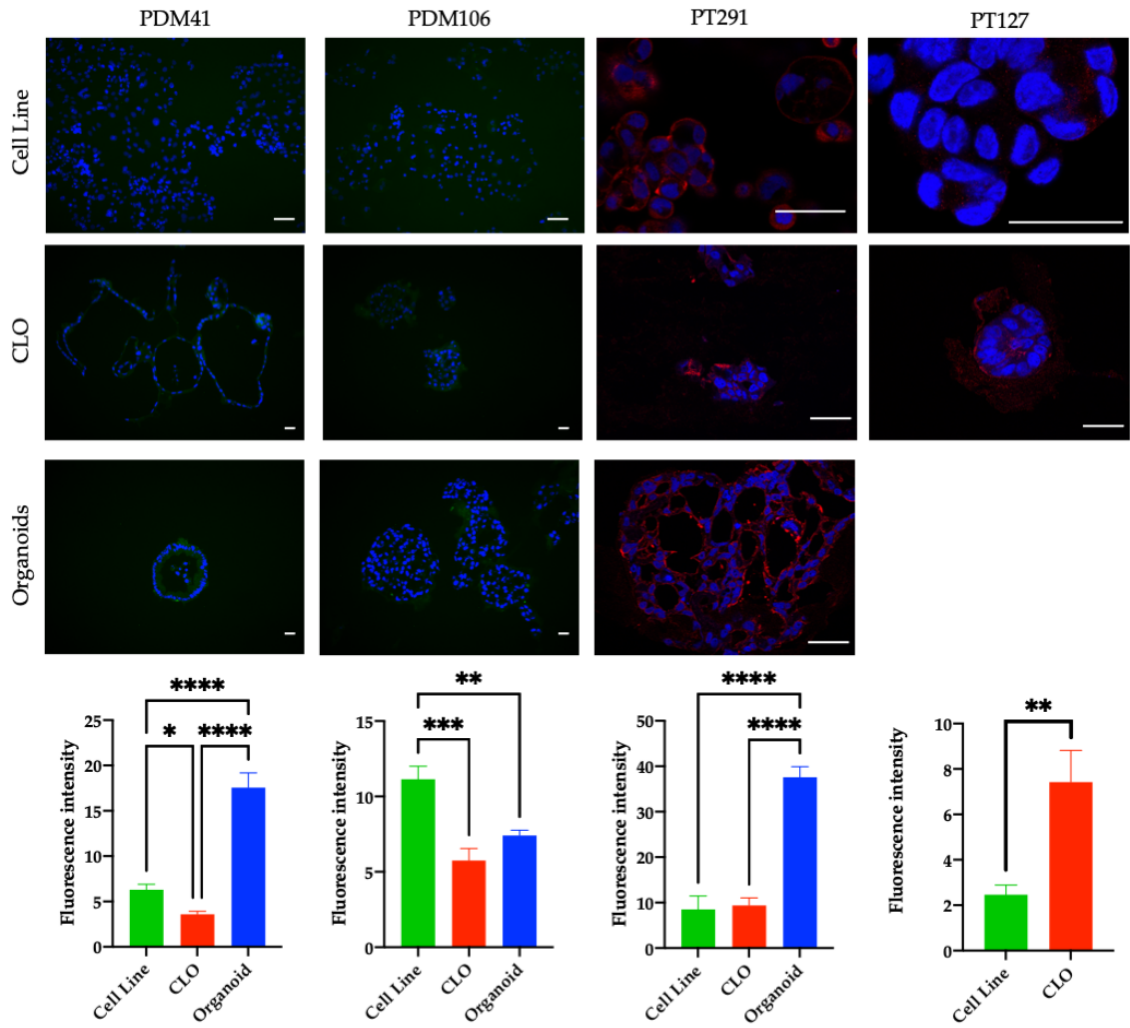


Figure 3.25: Images of immunofluorescence staining of ALDH1A1 (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$, **** denotes $p < 0.0001$.

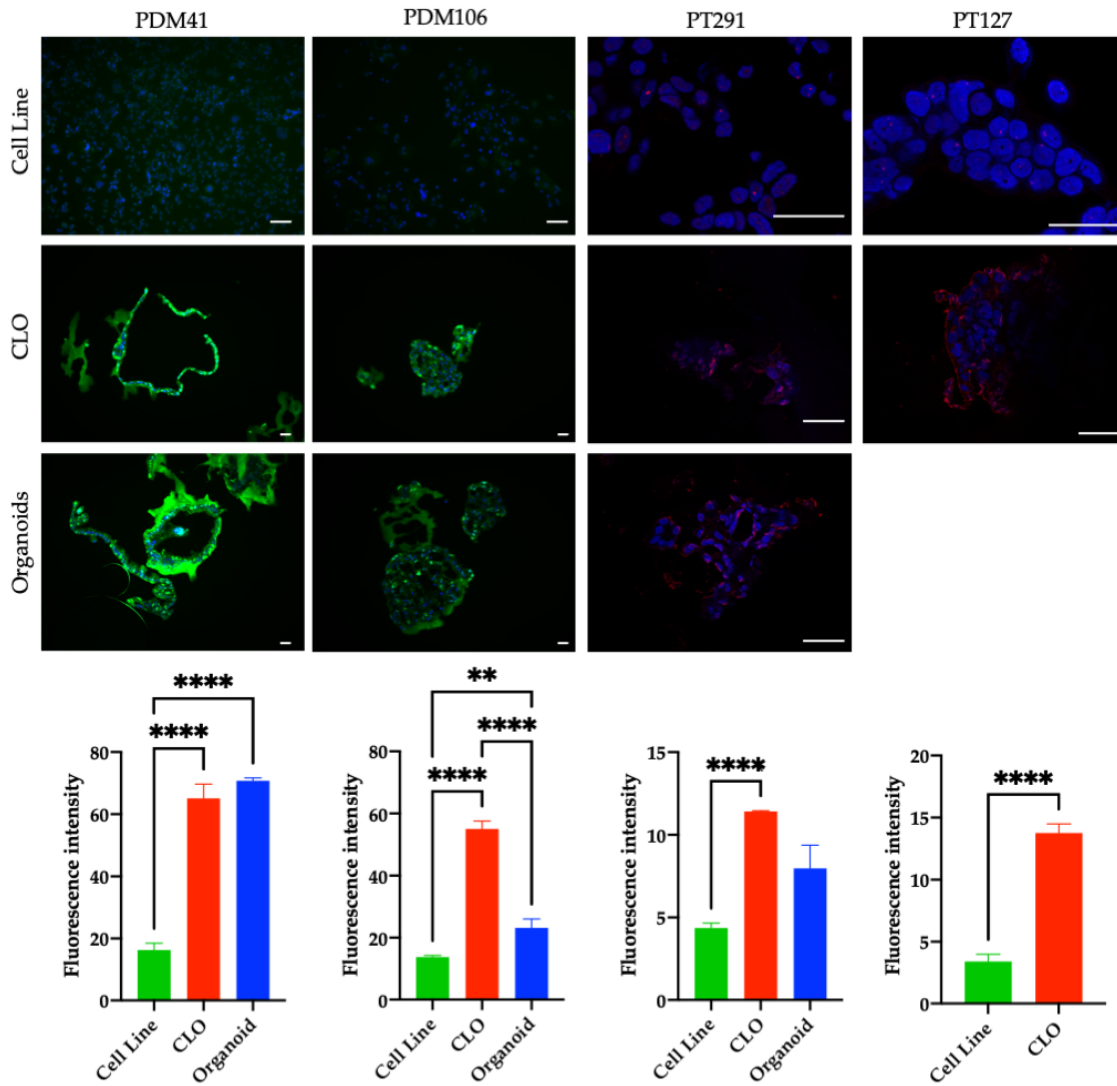


Figure 3.26: Images of immunofluorescence staining of CXCR4 (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

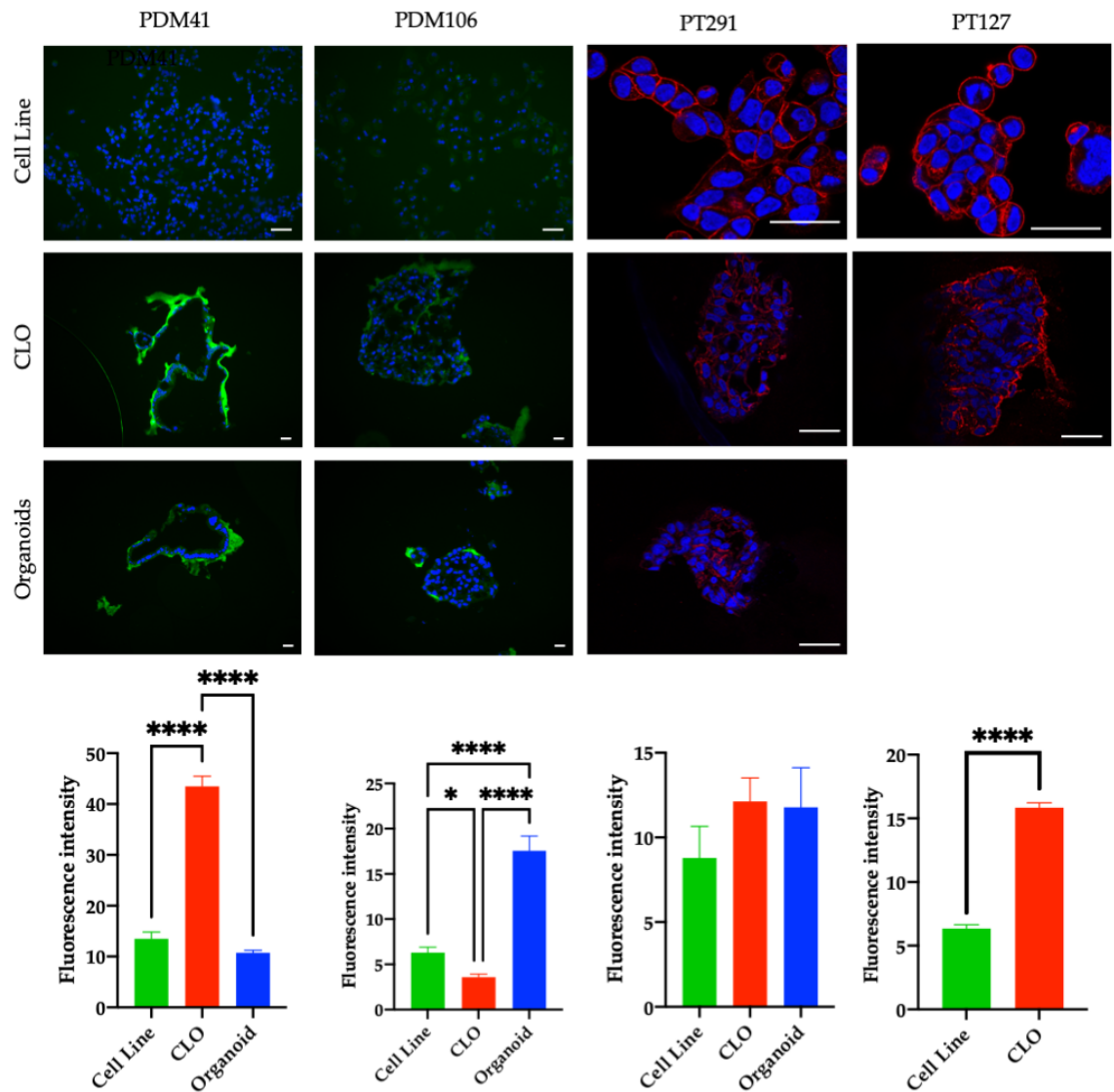


Figure 3.27: Images of immunofluorescence staining of EpCAM (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bares 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

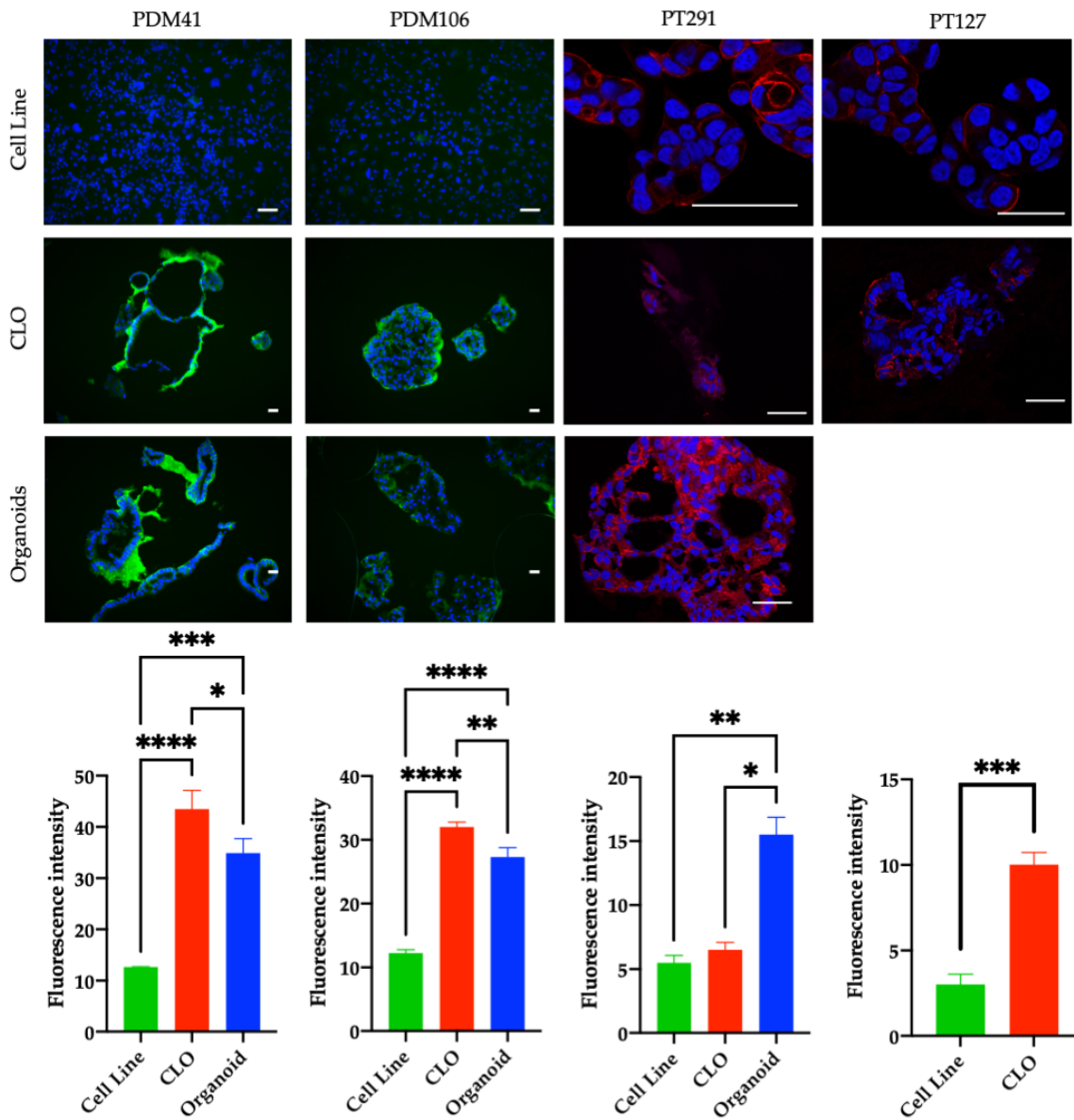


Figure 3.28: Images of immunofluorescence staining of HCAM (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

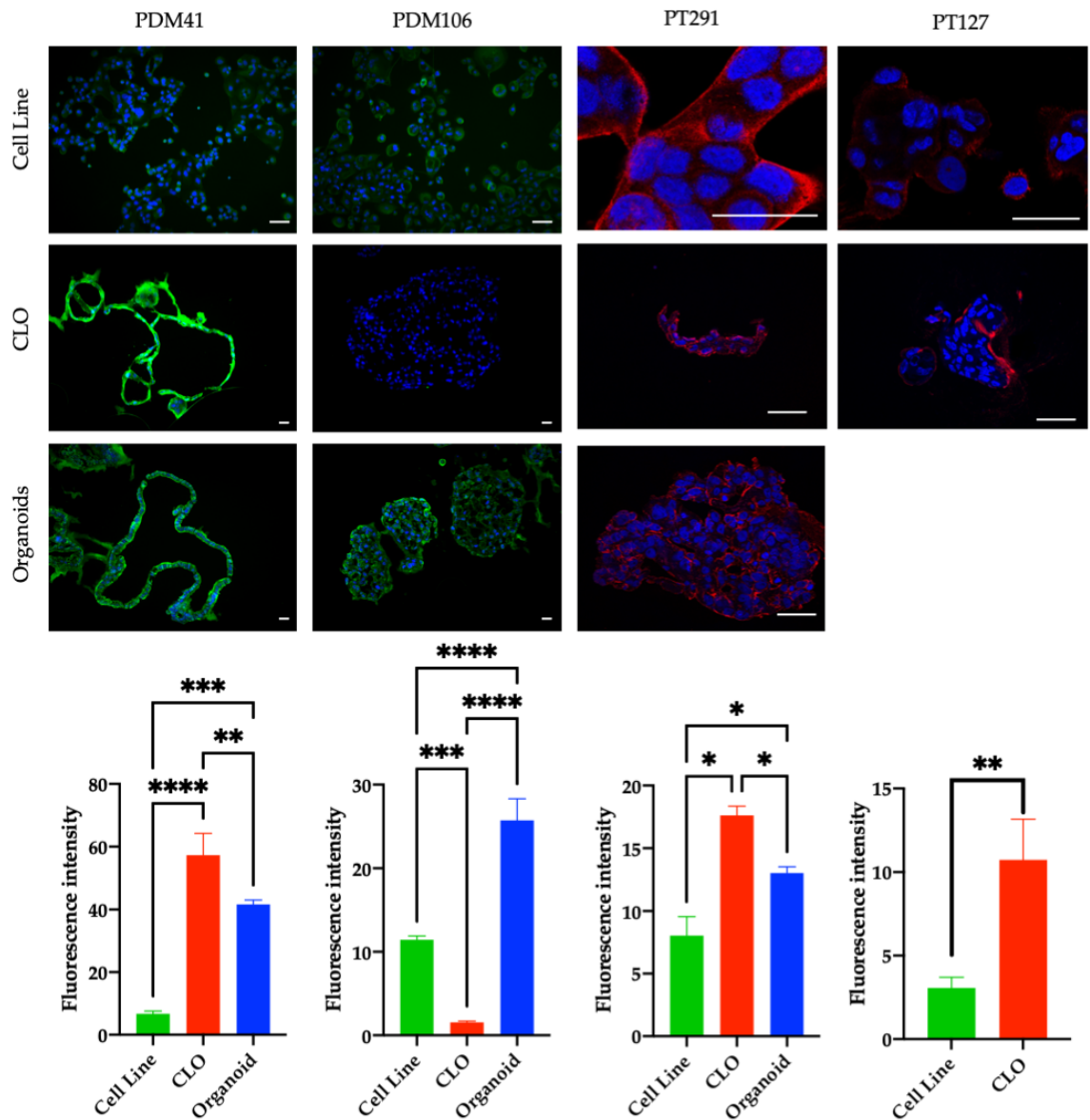


Figure 3.29: Images of immunofluorescence staining of MASPIN (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

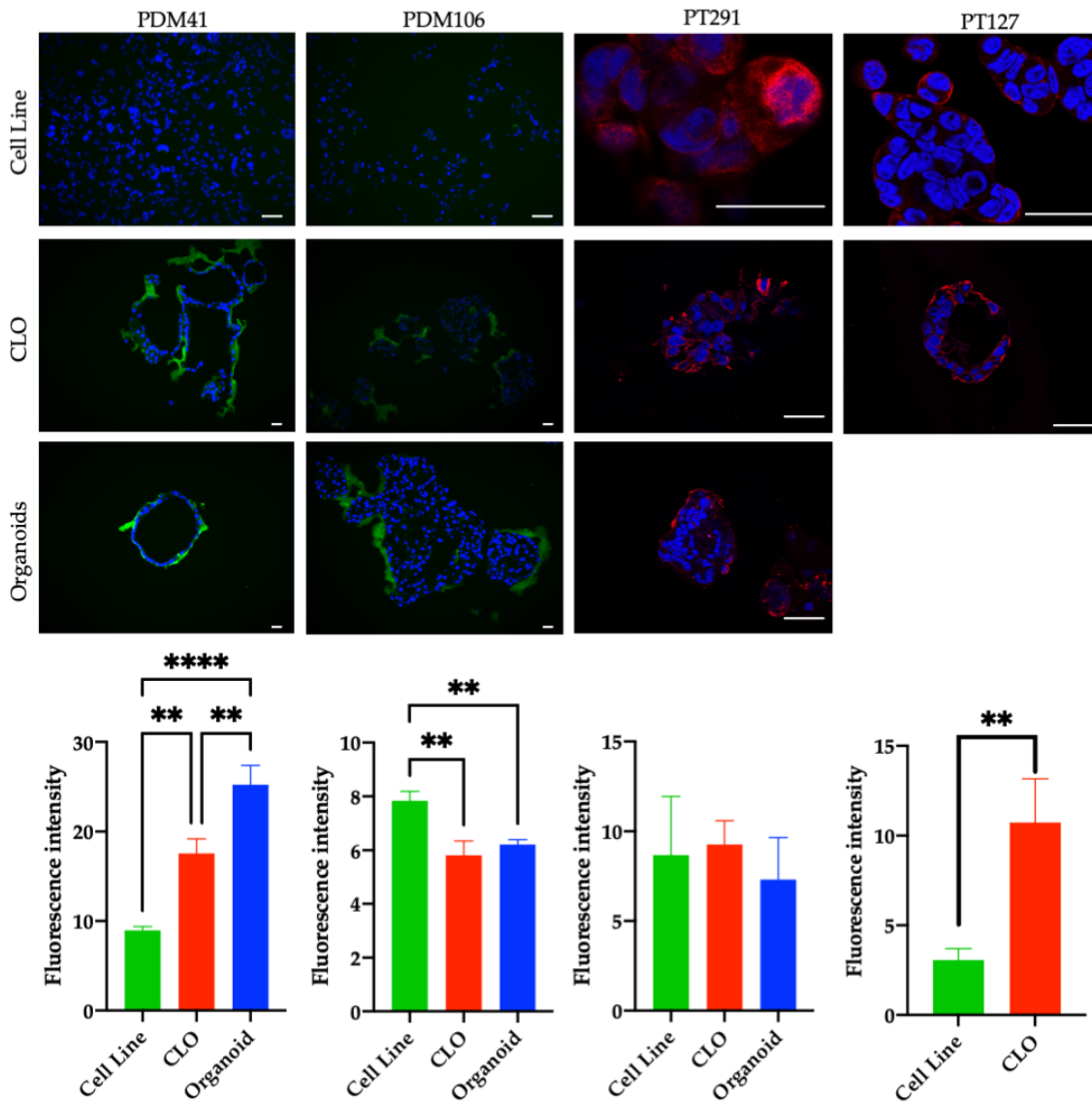


Figure 3.30: Images of immunofluorescence staining of PANCK (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

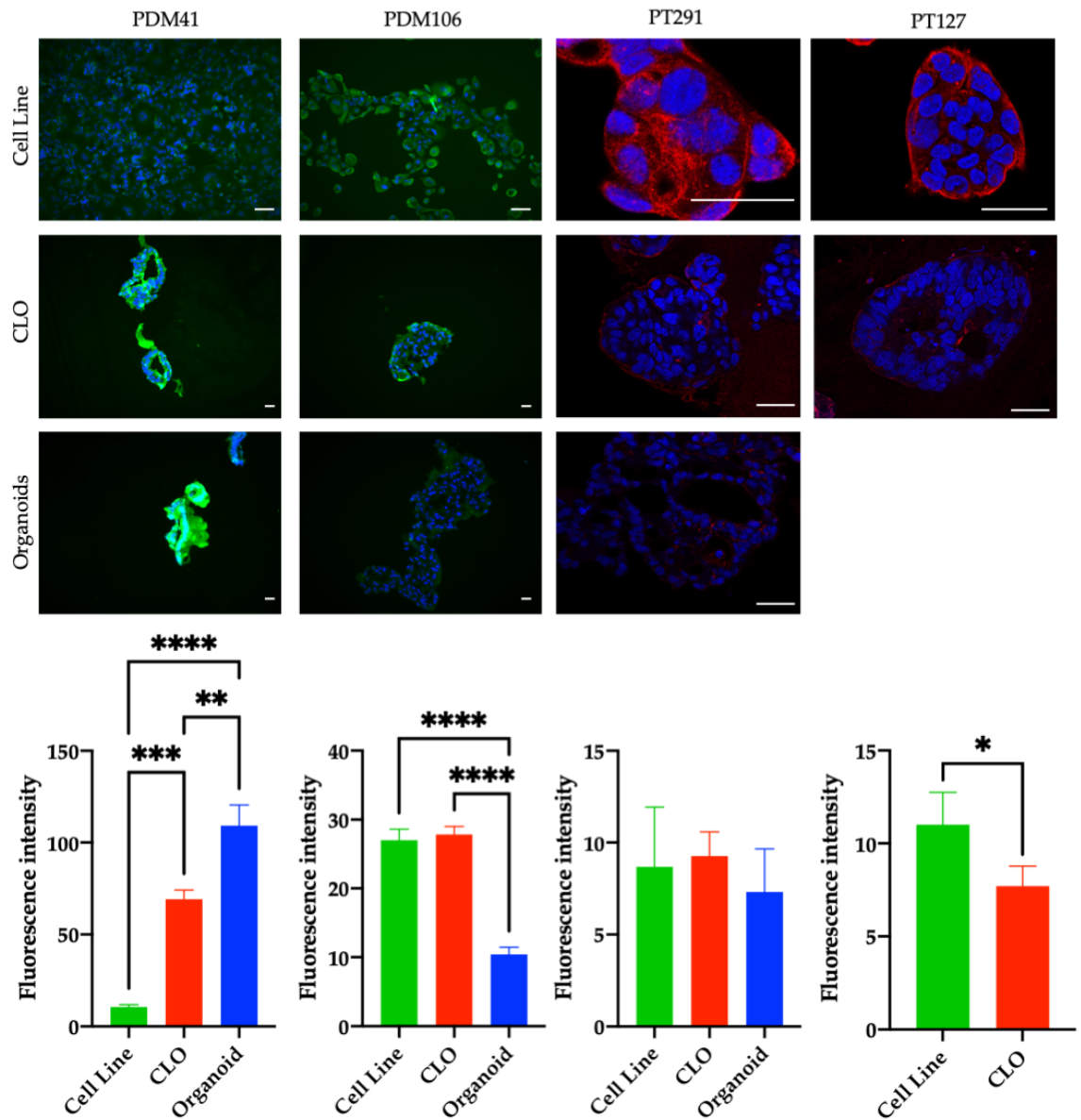


Figure 3.31: Images of immunofluorescence staining of PDXI (green/red) and counterstained with DAPI (blue) in matched cell lines, CLOs and organoids generated from patients PDM41, PDM106, PT291, and matched cell lines and CLOs generated from patient PT127. Scale bars 100 μm (PDM41 and PDM106) and 40 μm (PT291/PT127). A repeated measures ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

3.2.10. Transcriptomics

RNA-sequencing was performed to identify transcriptome signatures of genes dysregulated among PDX derived organoids/PDX tumour and corresponding CLOs compared to 2D primary cell lines in patient samples PT127 and PT291 was performed as per methodology described in Section 2.9.

3.2.10.1. Principal component analysis of sequenced samples

Principal component analysis (PCA) reveals PT291 organoids and CLOs clustered together; while PT127 PDX tumour and CLOs clustered together, separate from their corresponding isogenic matched 2D primary cell line (Figure 3.32).

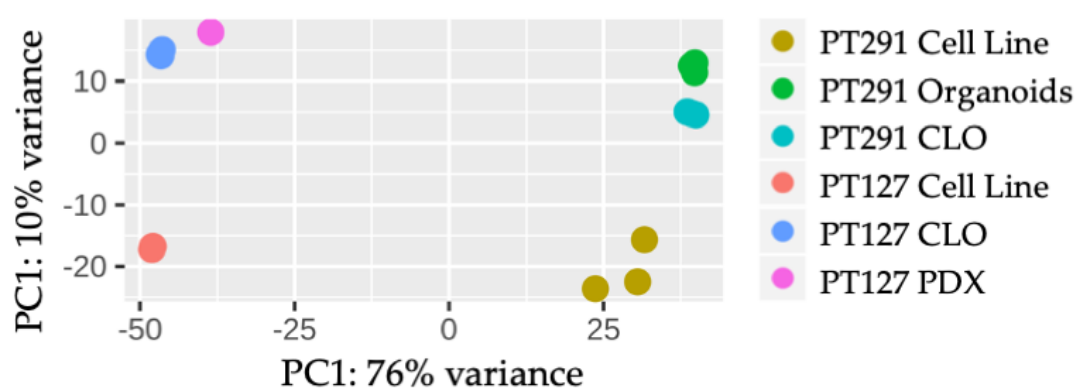


Figure 3.32: Principal component (PC) analysis was performed to compare the expression profiles of PT127 cell line, PT127 CLO, PT127 PDX, PT291 cell line, PT291 CLO and PT291 organoids.

3.2.10.2. Volcano Plots of PT127 PDX vs. CLO and PT291 organoids vs. CLO

A volcano plot was used to visualise the overall distribution of differentially expressed genes in the PT127 PDX vs. CLO (Figure 3.33) and in the PT291 organoids vs. CLO (Figure 3.34). In PT127, 1506 genes were significantly upregulated ($p < 0.001$; \log_2 fold-change ≥ 2), and 1594 genes were significantly downregulated ($p < 0.001$; \log_2 fold-change ≥ 2) in the PDX in comparison to the CLO. In PT291, 519 genes were significantly upregulated ($p < 0.001$; \log_2 fold-change ≥ 2), and 302 genes were significantly downregulated ($p < 0.001$; \log_2 fold-change ≥ 2) in the organoids compared to the matched CLOs.

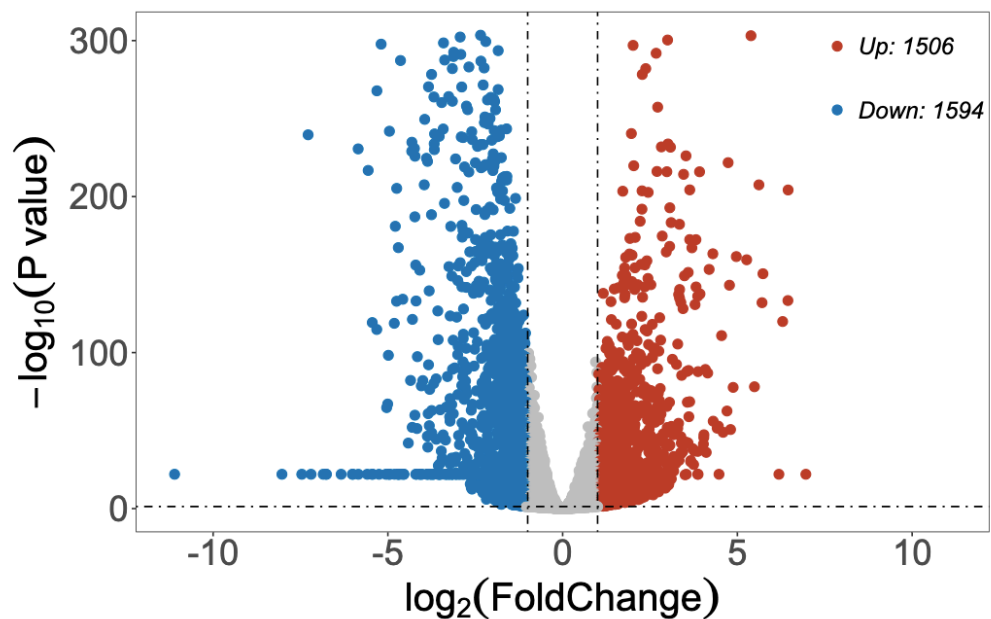


Figure 3.33: Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) ($p < 0.001$; \log_2 fold-change ≥ 2) in PT127 PDX vs. CLO. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ transformed P-values from logistic regression analysis.

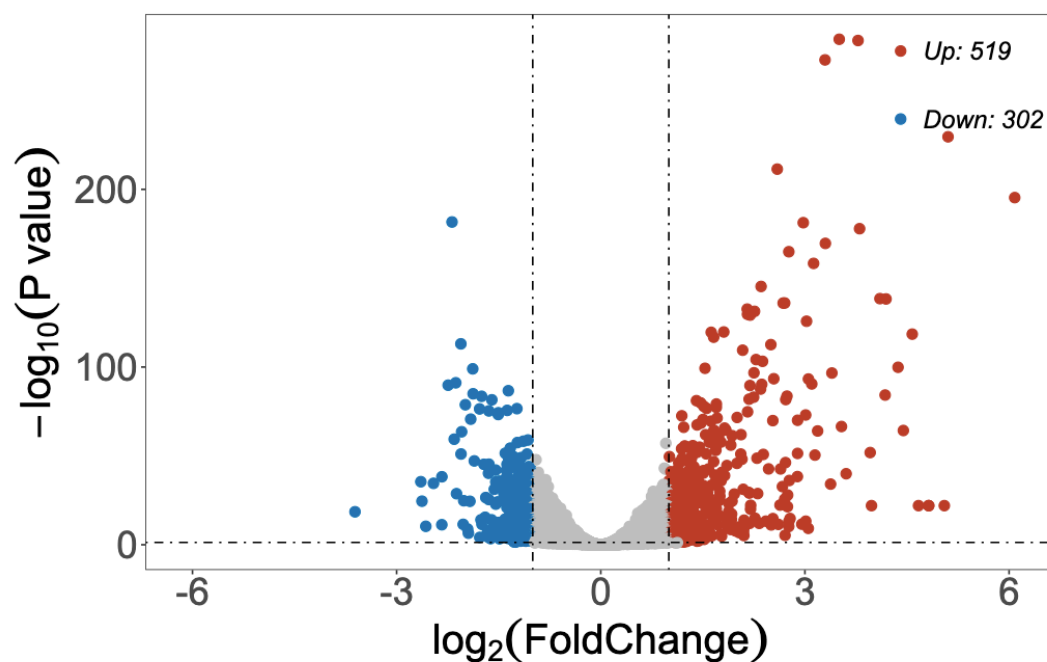


Figure 3.34: Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) ($p < 0.001$; \log_2 fold-change ≥ 2) in PT291 organoid vs. CLO. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ transformed P-values from logistic regression analysis.

3.2.10.3. Hierarchical clustering of PT127 and PT291 CLOs and cell lines

To identify if cell models resulted in similar overall mRNA expression patterns, hierarchical clustering was performed on all differentially expressed genes in the grouped CLOs and 2D primary cell lines (PT291 and PT127) (Figure 3.35). Organoid and PDX models were excluded from hierarchical clustering as RNA-sequencing data was only available for one patient sample for each model, and thereby a direct comparison could not be made.

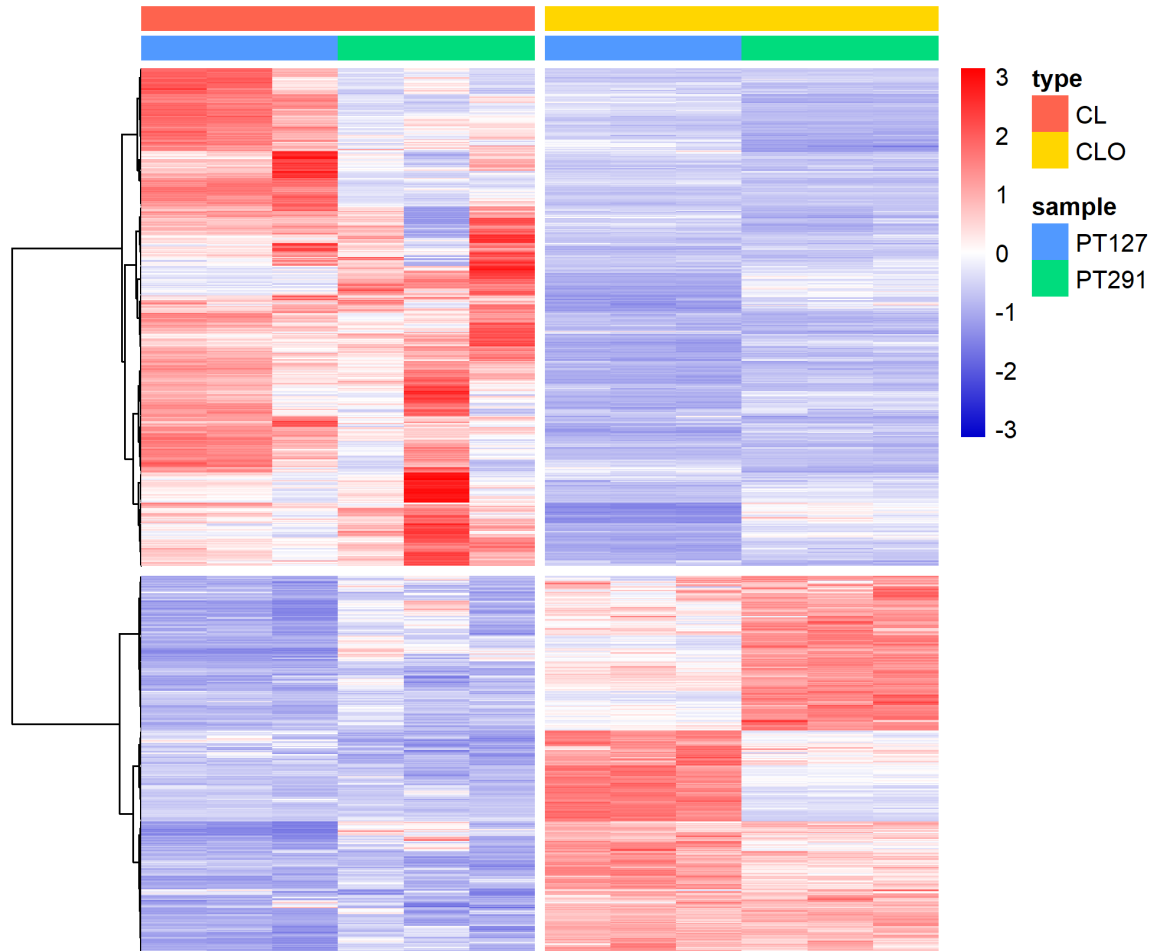


Figure 3.35: Clustering of all differentially expressed genes (fold-change ≥ 2 , $p\text{-adj} < 0.001$) in CLOs and cell lines (PT127 and PT291).

3.2.10.4. Scatterplot of differentially expressed genes in CLOs vs. cell lines

Figure 3.38 scatterplot reveals the most differentially expressed genes between grouped PT291 and PT127 CLOs compared to grouped PT291 and PT127 cell line. *FOSB* (log2 fold-change -6.50, p-adj 6.65E-50), *CAVI* (log2 fold-change -5.77 p-adj 4.89E-75) and *AHNAK2* (log2 fold-change -4.90, p-adj 9.25E-20) were identified as the three top downregulated genes in the grouped CLOs compared to the grouped 2D primary cell lines. *TFF3* (log2 fold-change 4.87, p-adj 5.86E-32), *PLEKHB1* (log2 fold-change 3.80, p-adj 2.59E-13) and *LTB* (log2 fold-change 3.63, p-adj 2.32E-11) were the top upregulated genes.

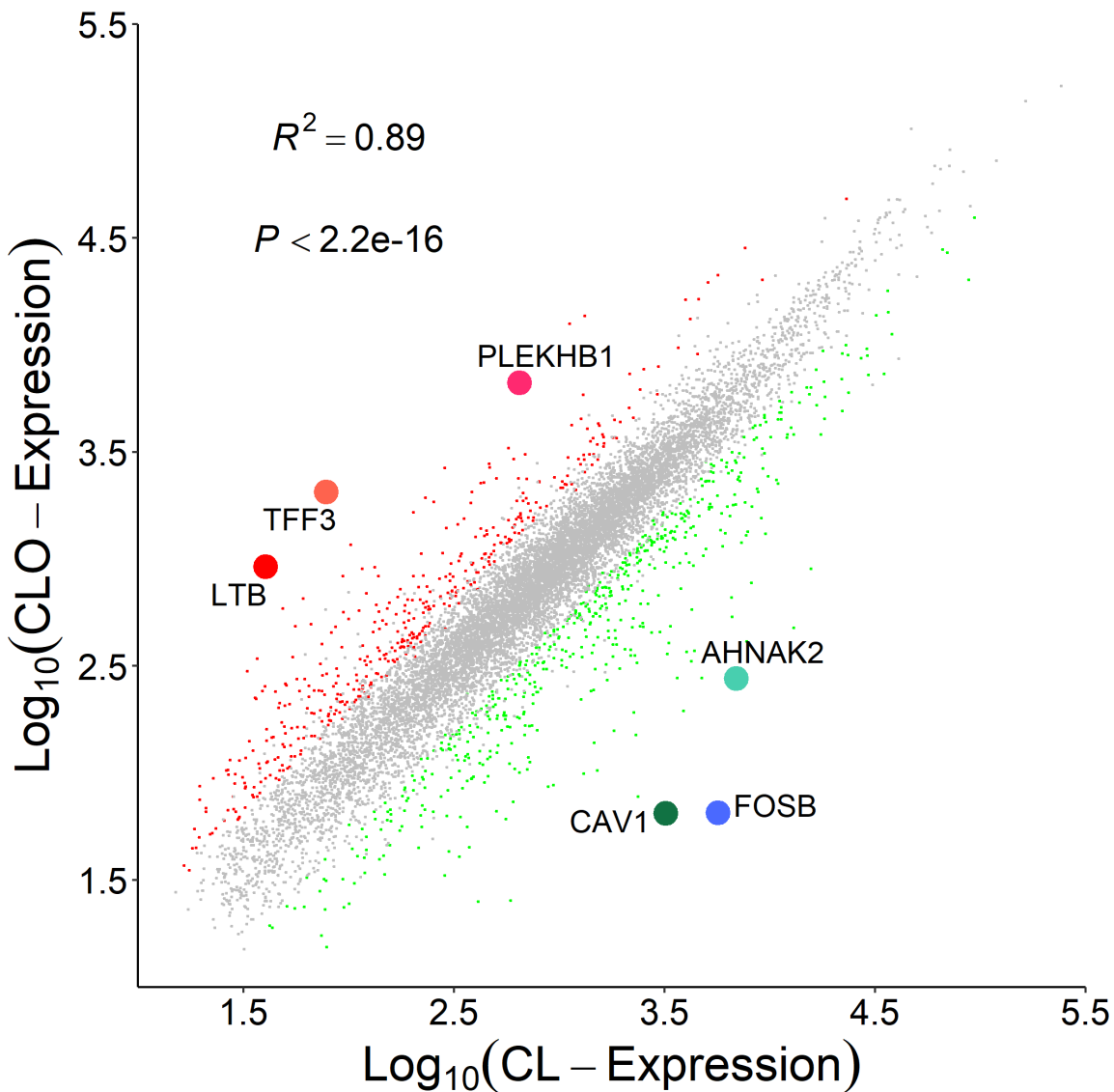


Figure 3.36: Scatterplot – grouped PT291 and PT127 CLO vs CL, genes with the largest log2 fold-change are highlighted.

3.2.10.5. Differential gene expression between organoid/PDX and CLO vs. cell line models

Analysis was performed to identify genes which were upregulated in PT127 PDX and CLO models, and in PT291 organoid and CLO models. PT127 PDX and CLOs had a total of 160 genes upregulated, and PT291 organoids and CLOs had a total of 66 genes upregulated, with 23 overlapping genes (Figure 3.37 A & C). PT127 PDX and CLOs had a total of 137 genes downregulated, and PT291 organoids and CLOs had a total of 119 genes downregulated, with 37 overlapping genes (Figure 3.37 B & C).

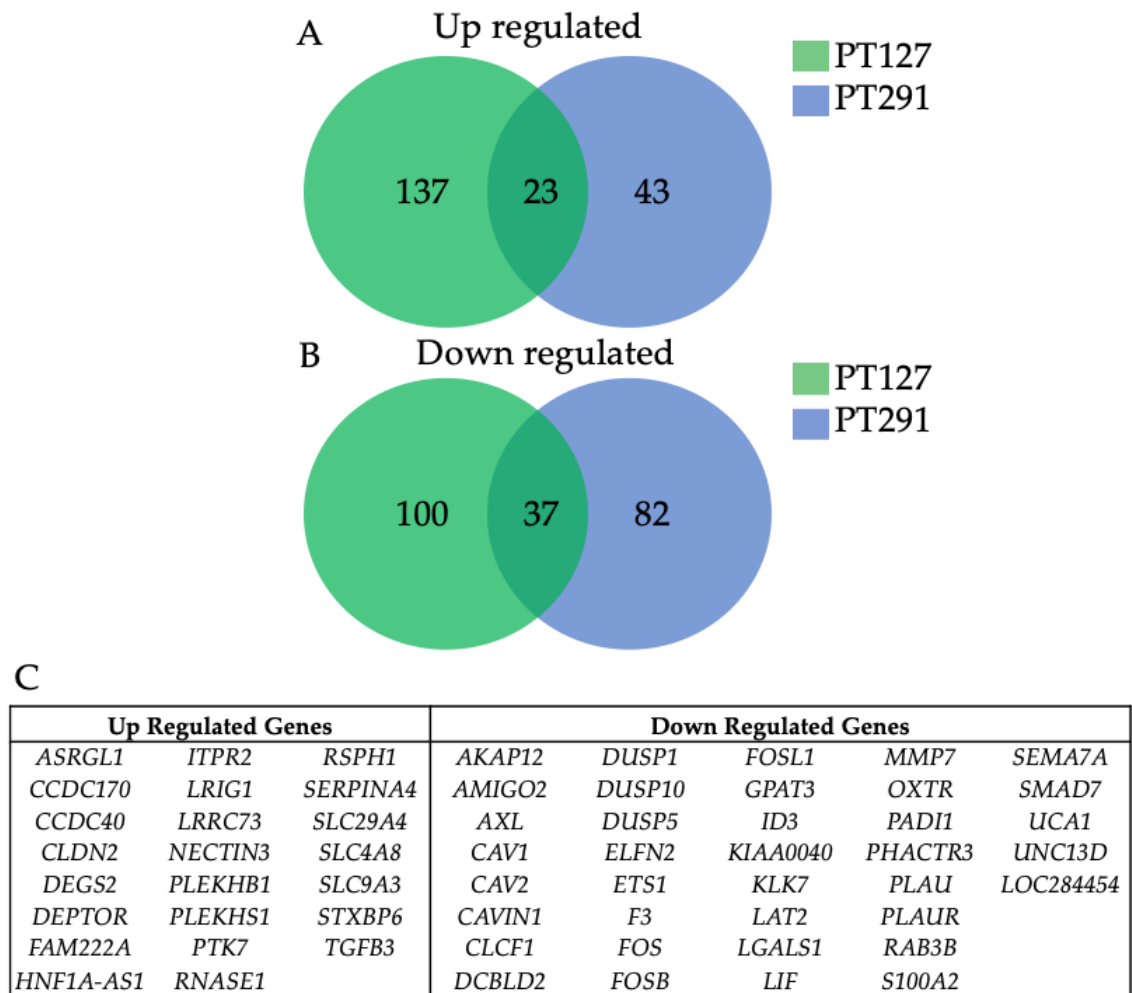


Figure 3.37: (A) Venn diagram – genes upregulated in both PT127 PDX and CLO and PT291 ORG and CLO, (B) Venn diagram – genes downregulated in both PT127 PDX and CLO and PT291 ORG and CLO, and (C) List of upregulated and downregulated genes in both PT127 PDX/CLO and PT291 organoid/CLO.

3.2.10.6. Transcriptomic signature in PDX tumour, organoid and CLO models

To further identify a specific transcriptomic signature common among the PDX tumour/ organoid and CLOs, the transcriptomic signature of the PT127 PDX derived tumour and CLO, and the PT291 organoid and CLO were compared to their respective 2D primary cell lines. Overall, 144 genes were upregulated, and 130 genes were downregulated (\log_2 fold-change ≥ 2 , p-adj 0.001) in both PT127 PDX tumour/CLO compared to the isogenic matched cell line in 2D (Appendix D - Supplementary Table 8.3 and Supplementary Table 8.4). In PT291, 62 genes were upregulated, and 105 genes downregulated in the organoids and CLOs compared to the isogenic matched 2D cell line (Appendix D- Supplementary Table 8.5 and Supplementary Table 8.6). The 10-most significantly ($p\text{-adj} \leq 7.15\text{E-}42$) dysregulated genes in PT127 PDX and CLO compared to the 2D cell line are shown in Figure 3.38 A; *KIF12* had the most upregulated \log_2 fold-change (\log_2 fold-change 5.42, $p\text{-adj}=1.62\text{E-}62$) and *FOSB* (\log_2 fold-change -7.24, $p\text{-adj} 1.82\text{E-}55$) had the most downregulated fold-change. The 10 most significant ($p\text{-adj} \leq 5.79\text{E-}14$) up and downregulated genes in PT291 organoids and CLOs compared to the isogenic matched 2D cell line are outlined in Figure 3.38 B. *SERPINA5* (\log_2 fold-change 4.21, $p\text{-adj}=2.17\text{E-}14$) had the most upregulated \log_2 fold-change, and *AHNAK2* (\log_2 fold-change -5.32, $p\text{-adj}=1.28\text{E-}110$) displayed the most downregulated \log_2 fold-change.

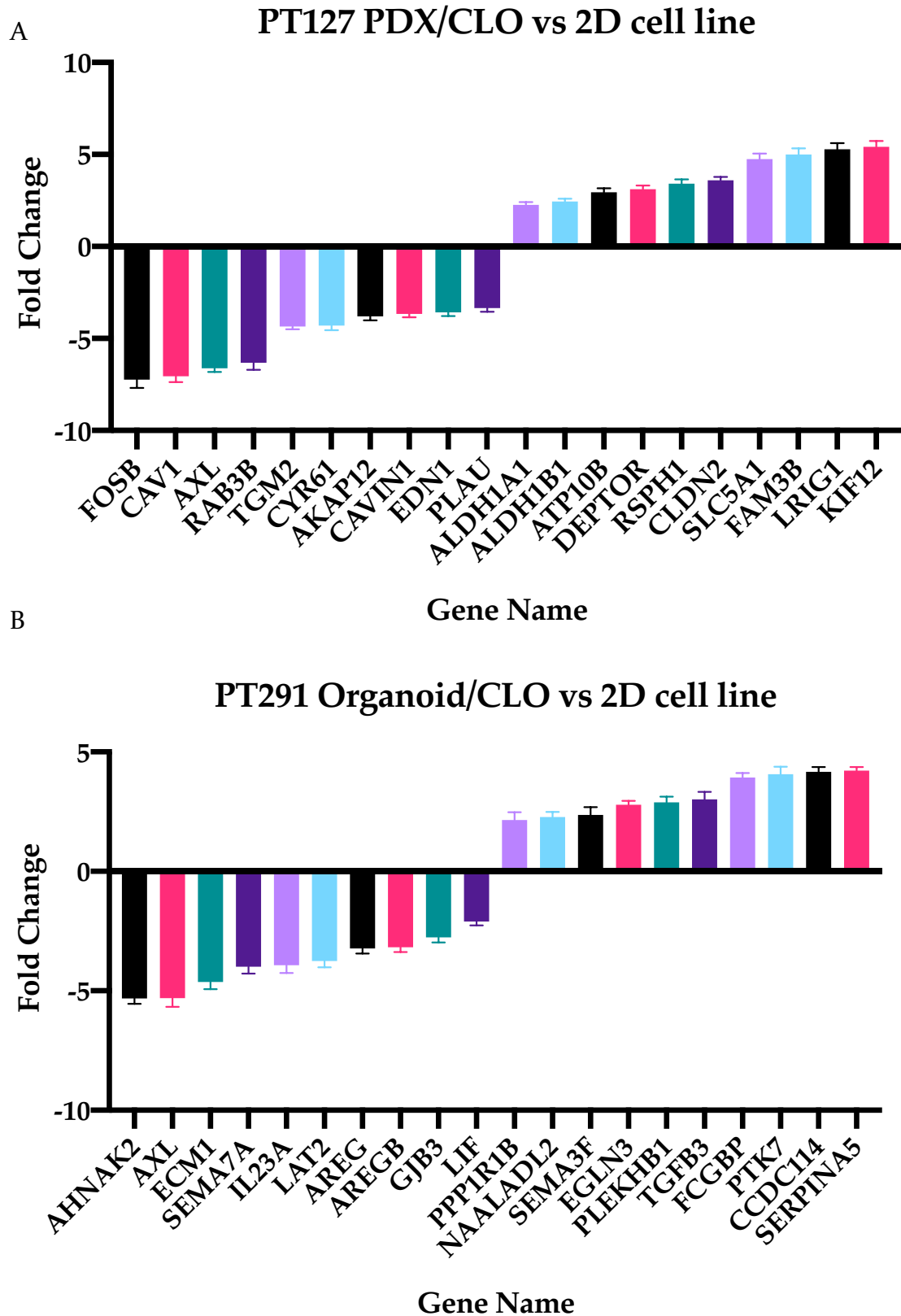


Figure 3.38: (A) The 10 most significant ($p\text{-adj} \leq 7.15E-42$) up and downregulated genes in PT127 PDX/CLO versus the matched 2D cell line (B) The 10 most significant up and downregulated PT291 organoid/CLO versus 2D cell line.

3.2.11. Subtyping of PT291 cancer models

Using a classification method from Moffitt *et al.* [65] the PT291 models cell line, organoid and CLO were subtyped into the PDAC cancer groups of classical and basal. This classification method uses mRNA counts of gene expression and compares the expression levels of 14 gene pairs (Table 3.7). If expression of Gene A is higher than Gene B, this increases the odds of the sample being in the basal subtype (Table 3.7). The odds ratio of model being basal for PT291 cell line is 15.47; 11.26 for CLO, and 6.89 for PT291 organoids (Figure 3.39).

Table 3.7: A table from Moffitt *et al.* [65] which determines the increase in odds ratio of sample being basal subtype. Gene comparison in bold was excluded as gene count data was not available for CTSL2.

<i>Gene A</i>	<i>Gene B</i>	<i>Coefficient</i>	<i>Estimated Increase in odds of basal class membership when A > B</i>
<i>CDI09</i>	<i>GPR160</i>	0.87	2.38
<i>SLC2A1</i>	<i>AGR2</i>	1.22	3.39
<i>KRT16</i>	<i>SLC44A4</i>	0.52	1.68
<i>CTSL2</i>	<i>TMEM45B</i>	1.43	4.17
<i>KRT6A</i>	<i>BCAS1</i>	0.70	2.01
<i>B3GNT5</i>	<i>VSIG2</i>	0.41	1.51
<i>MET</i>	<i>TFF3</i>	0.72	2.06
<i>CHST6</i>	<i>PLA2G10</i>	0.80	2.24
<i>SERPINB5</i>	<i>HPGD</i>	0.76	2.13
<i>DCBLD2</i>	<i>PLSI</i>	1.40	4.07
<i>IL20RB</i>	<i>FAM3D</i>	1.33	3.79
<i>PPP1R14C</i>	<i>SYTL2</i>	1.58	4.85
<i>NABI</i>	<i>PLEKHA6</i>	0.41	1.50
<i>MSLN</i>	<i>CAPN9</i>	1.58	4.83
(Intercept)		-7.16	

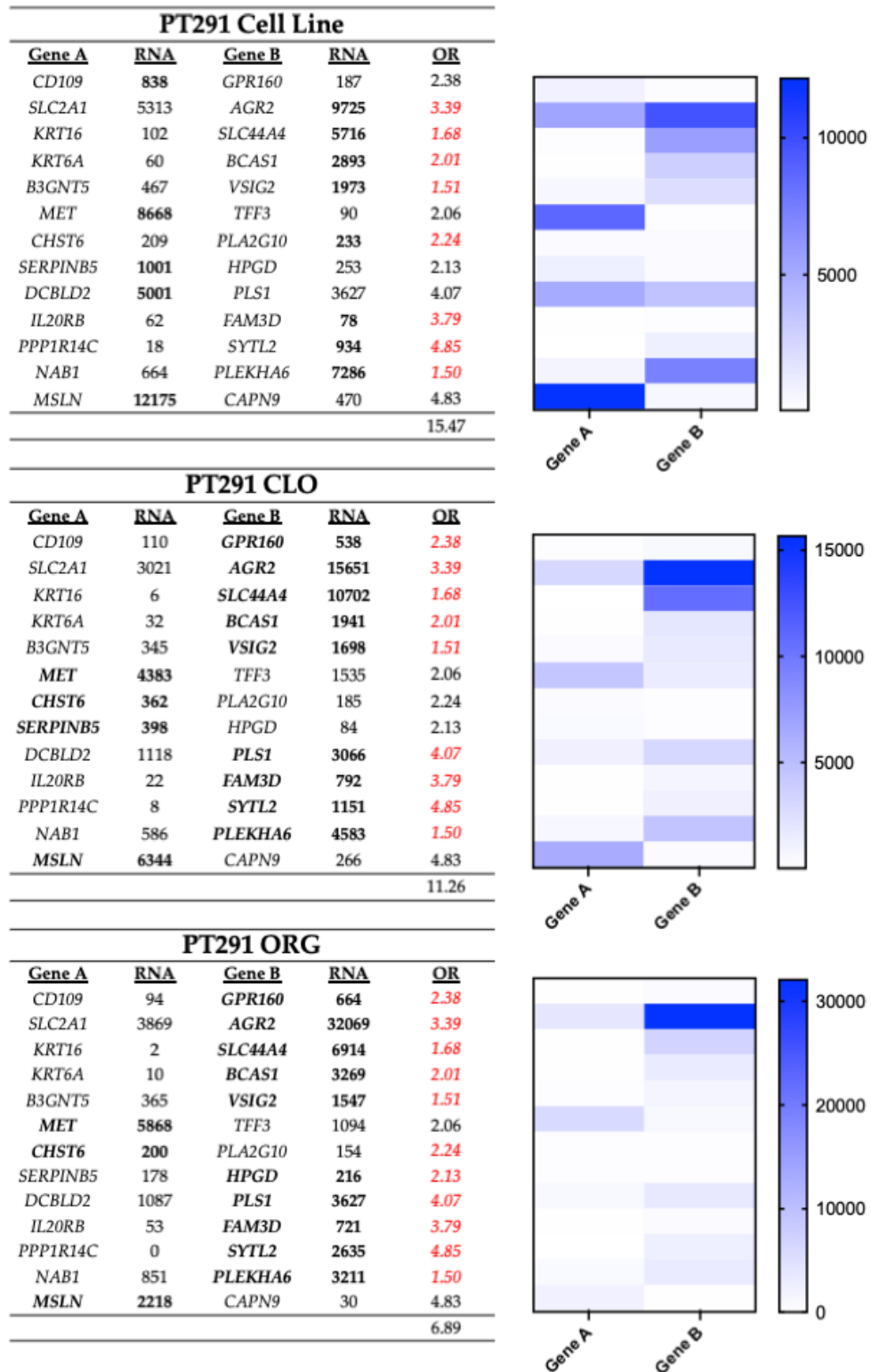


Figure 3.39. Tables of genes, gene counts and increase in OR of PT291 cell line, CLO and organoids to determine the increase in odds ratio of sample being basal subtype. Numbers in red were excluded from the calculation of total OR. Heatmaps show levels of gene expression of each model.

3.2.12. Xenograft *in vivo* study

A xenograft *in vivo* study was performed to characterise the growth and tumour development of PT291 organoids, CLOs and cell lines when subcutaneously implanted into SCID mice was performed as per methodology described in Section 2.10. CLOs and organoids were grown in the lab for one week prior to implantation, and mice were implanted with 5×10^6 cells in a 1:1 ratio with ECM. Mice had an average weight of 22.1 g (range: 21 g - 24.8 g) at the time of tumour inoculation.

3.2.12.1. Tumour development

Tumours were allowed to develop, and when of sufficient size, were measured with callipers. All cell models resulted in tumour growth (Figure 3.40). After 23 days, two of the three mice implanted with organoids, and one of the three mice implanted with CLOs had a measurable tumour. After an additional 9 days, one of the mice implanted with the cell line had a measurable tumour. In this study, all mice in the organoid and CLO groups developed tumours, however, only one of the mice implanted with the PT291 cell line developed a tumour which was large enough to measure (Figure 3.41).

For comparison, the growth of the original PDX tumour has been included for comparison in Figure 3.42 [294]. PT291 was implanted into three mice, with only one animal generating a tumour. Tumour was first observed 71 days post implantation and measurable from 160 days post implantation. The days post implantation in Figure 3.42 have been adjusted to when the tumour was measurable, so tumour growth is comparable to tumours in the current study.

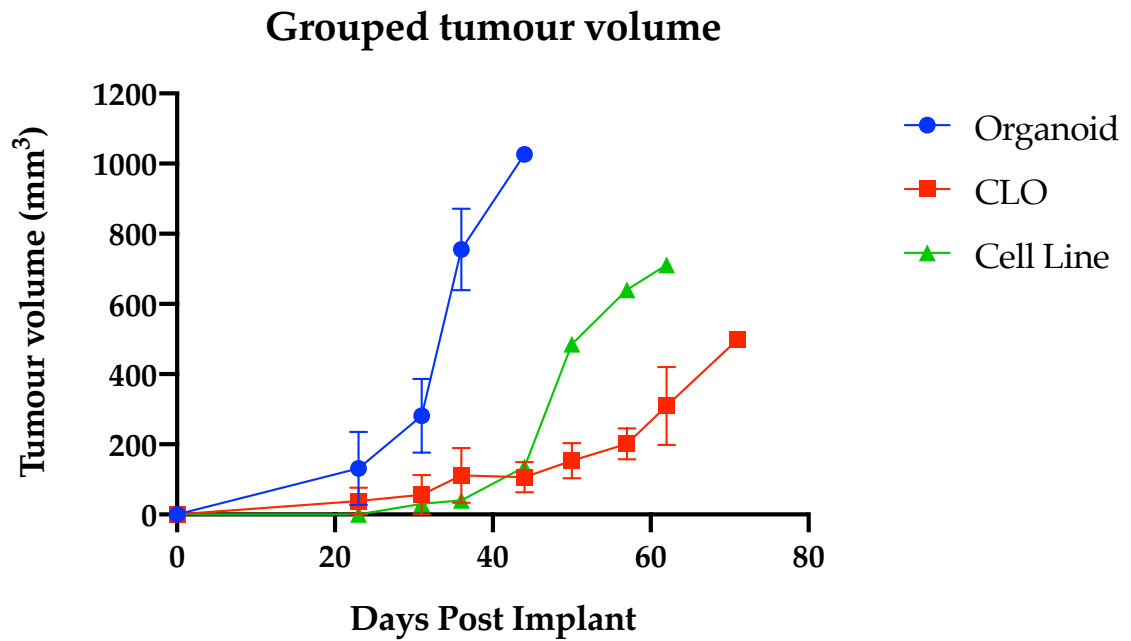


Figure 3.40: Average tumour growth of PT291 organoid ($n=3$) (blue), CLO ($n=3$) (red), and cell line ($n=1$) (green) in mice. Tumour dimensions were measured by callipers and tumour volume was calculated as $(\text{Height} \times \text{Width} \times \text{Depth})/1.9$. Error bars represent standard error of the mean.

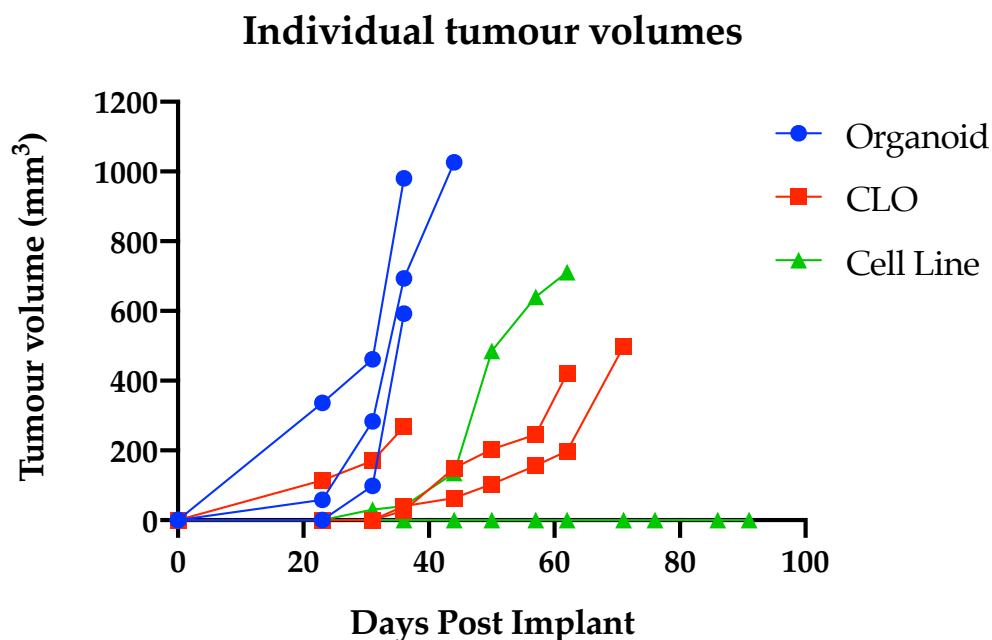


Figure 3.41: Individual tumour growth of PT291 organoid (blue), CLO (red) and cell line (green) in mice. Tumour dimensions were measured by callipers and tumour volume was calculated as $(\text{Height} \times \text{Width} \times \text{Depth})/1.9$. Error bars represent standard error of the mean.

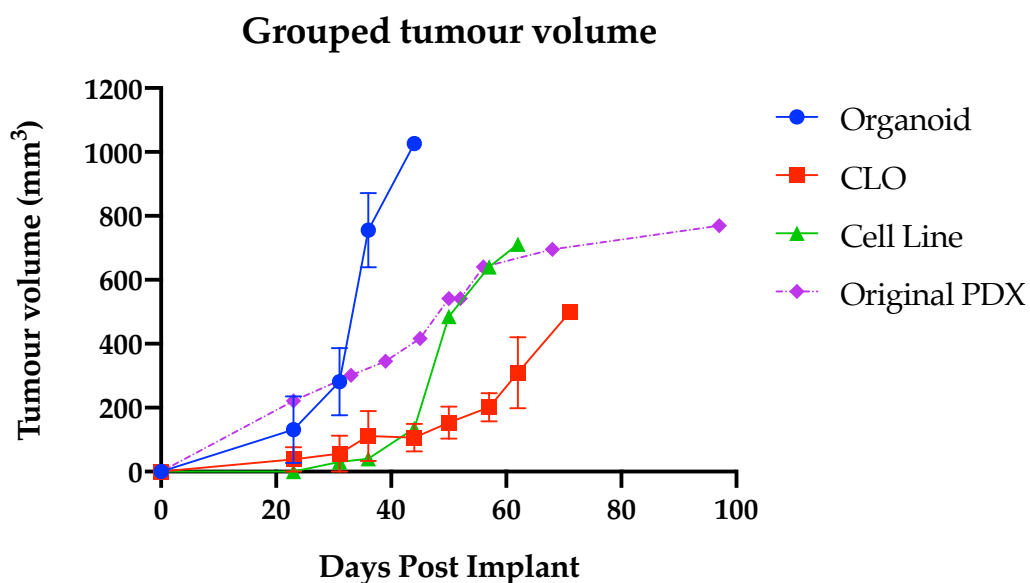


Figure 3.42: Average tumour growth of PT291 organoid (n=3) (blue), CLO (n=3) (red), cell line (n=1) (green) and original PDX tumour (n=1) (purple) in mice. The days post implant for the original PDX tumour have been adjusted to when the tumour was measurable, so the growth curve is comparable to tumours in current study. Tumour dimensions were measured by calipers and tumour volume was calculated as $(\text{Height} \times \text{Width} \times \text{Depth})/1.9$. Error bars represent standard error of the mean.

3.2.12.2. The effect of tumour burden on general health

The health of all mice was monitored daily and mouse body weight was used as an indicator of overall health of the animal, with a 10% decrease in weight an indication that the mouse is stressed, and humane euthanasia is required [295]. Mice were weighed weekly and no significant effect of implantation of cells and growth of tumour on mouse weight was observed (Figure 3.43 and Figure 3.44).

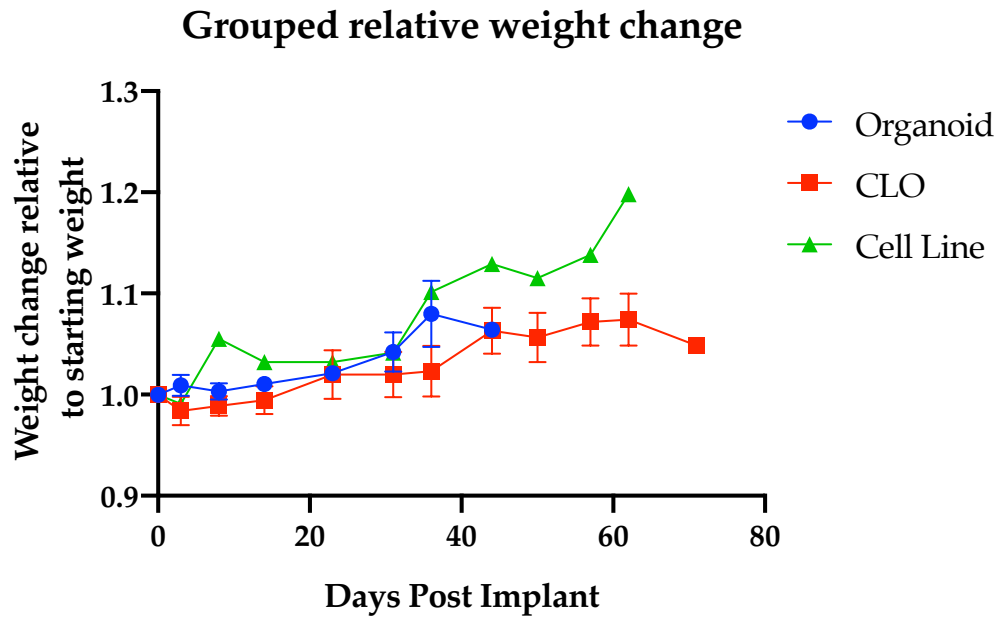


Figure 3.43: The weight of each mouse in groups PT291 organoid (n=3) (blue), CLO (n=3) (red) and cell line (n=1) (green) was normalised to its weight at tumour implantation and measured throughout the study. Error bars represent standard deviation.

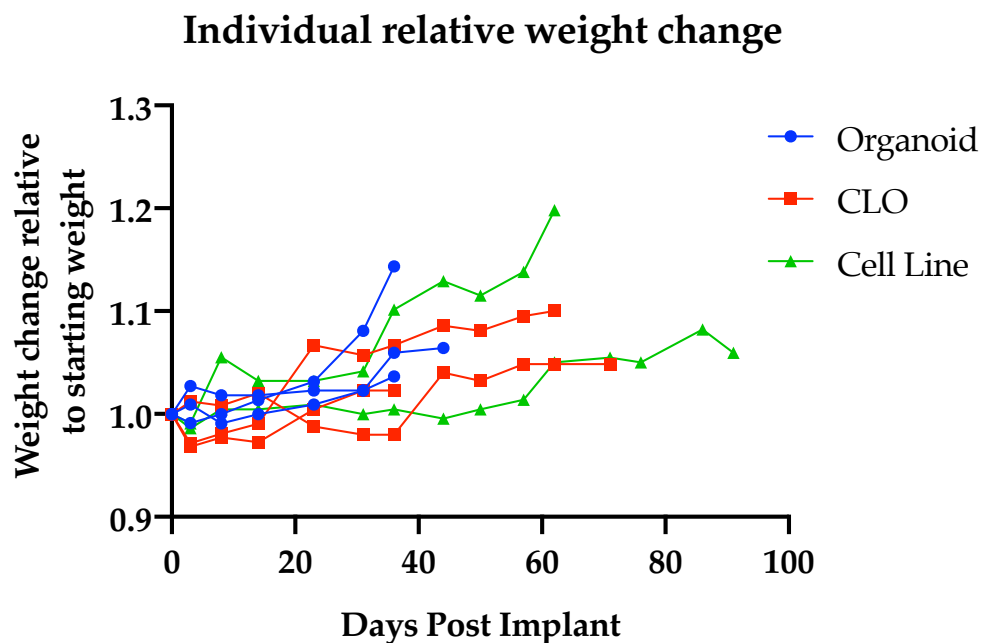


Figure 3.44. The weight of each mouse in groups PT291 organoid (blue), CLO (red) and cell line (green) was normalised to its weight at tumour implantation and measured throughout the study. Error bars represent standard deviation.

3.2.12.3. H&E Staining of tumours

There were clear morphological differences in the tumours formed from the organoids, CLOs and cell lines. The tumour from the implanted cell line (Figure 3.45 A) which was fast-growing, and tumours derived from the organoids (Figure 3.47 A-C) formed as large cell clusters, and no tumour-like structures. Whereas the tumours derived from the CLOs (Figure 3.46 A-C) and the slow-growing cell line (Figure 3.45 B) xenograft formed multiple independent PanIN-like lesions when implanted subcutaneously. Assessment by pathologist, Dr Maura Cotter, confirmed the xenografts derived from the CLOs and slow-growing cell line were adenocarcinoma, however, no definite tumour could be confirmed in organoid and the fast-growing cell line xenografts, and cells in mass were potentially classed as lymphocytes.

Cell Line

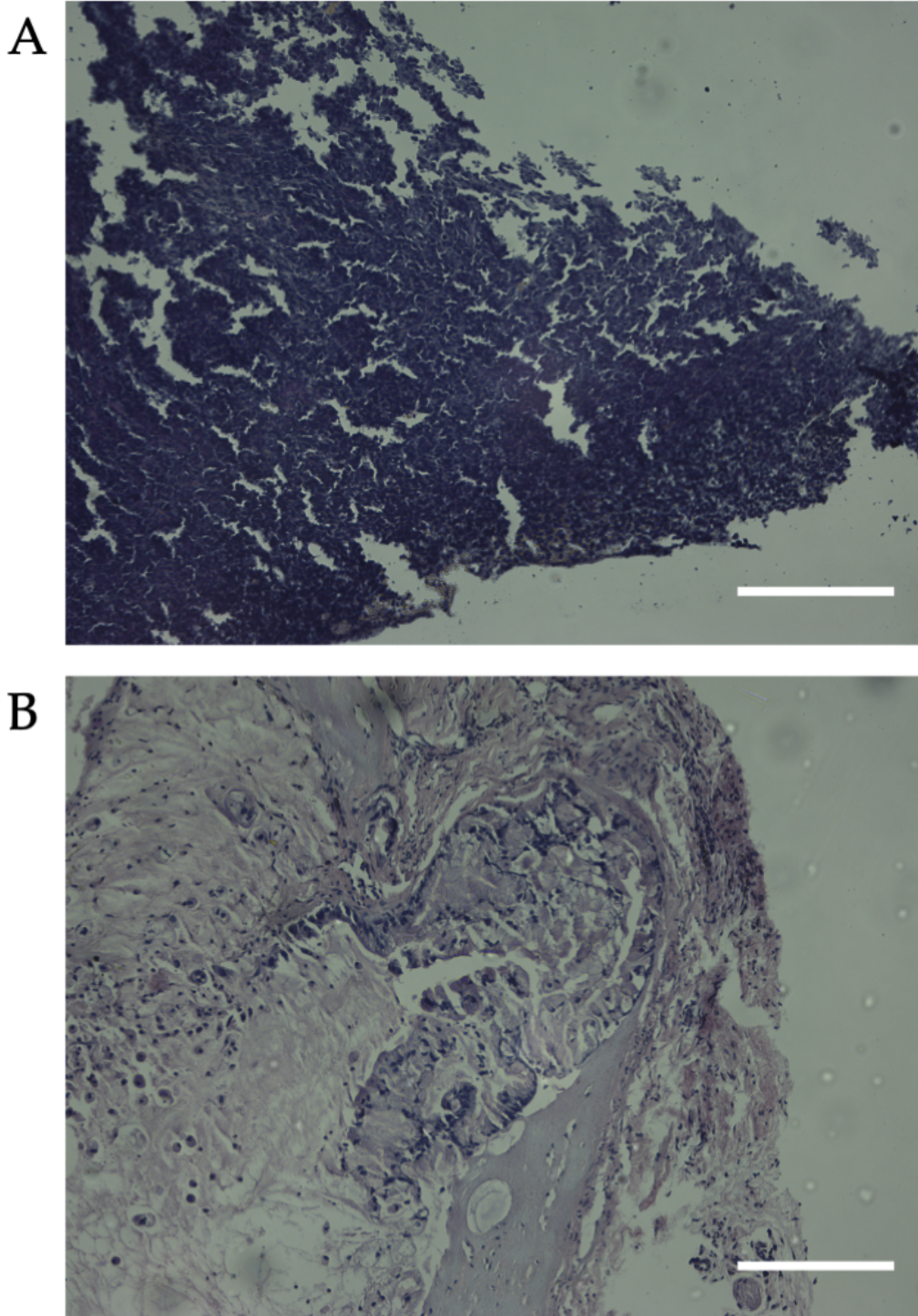


Figure 3.45: H&E staining of (A) fast-growing PT291 cell line xenograft, and (B) slow-growing PT291 cell line xenograft. Images taken at 100X magnification. Scale bars 200 μ m.

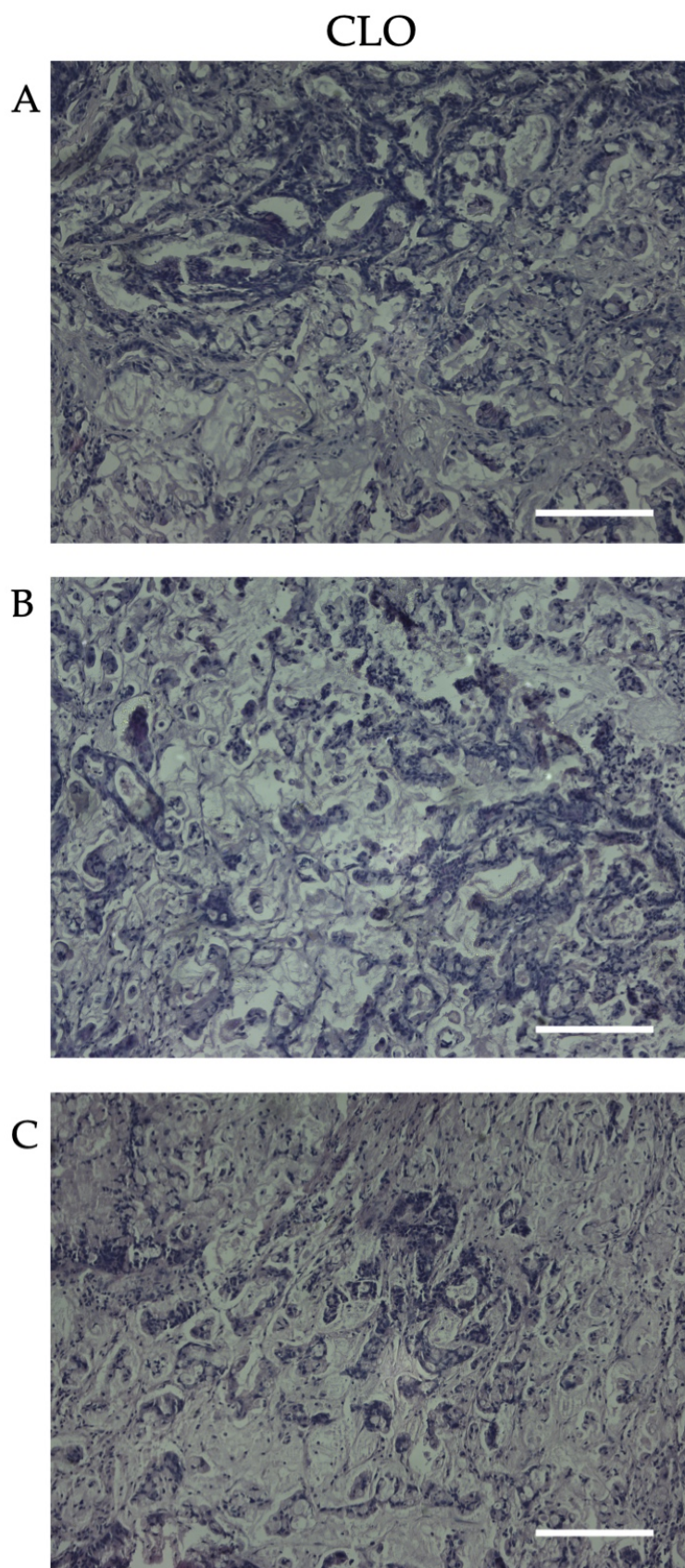


Figure 3.46: H&E staining of PT291 CLO tumour after implantation in SCID mice. Images taken at 100X magnification. Scale bars 200 μ m.

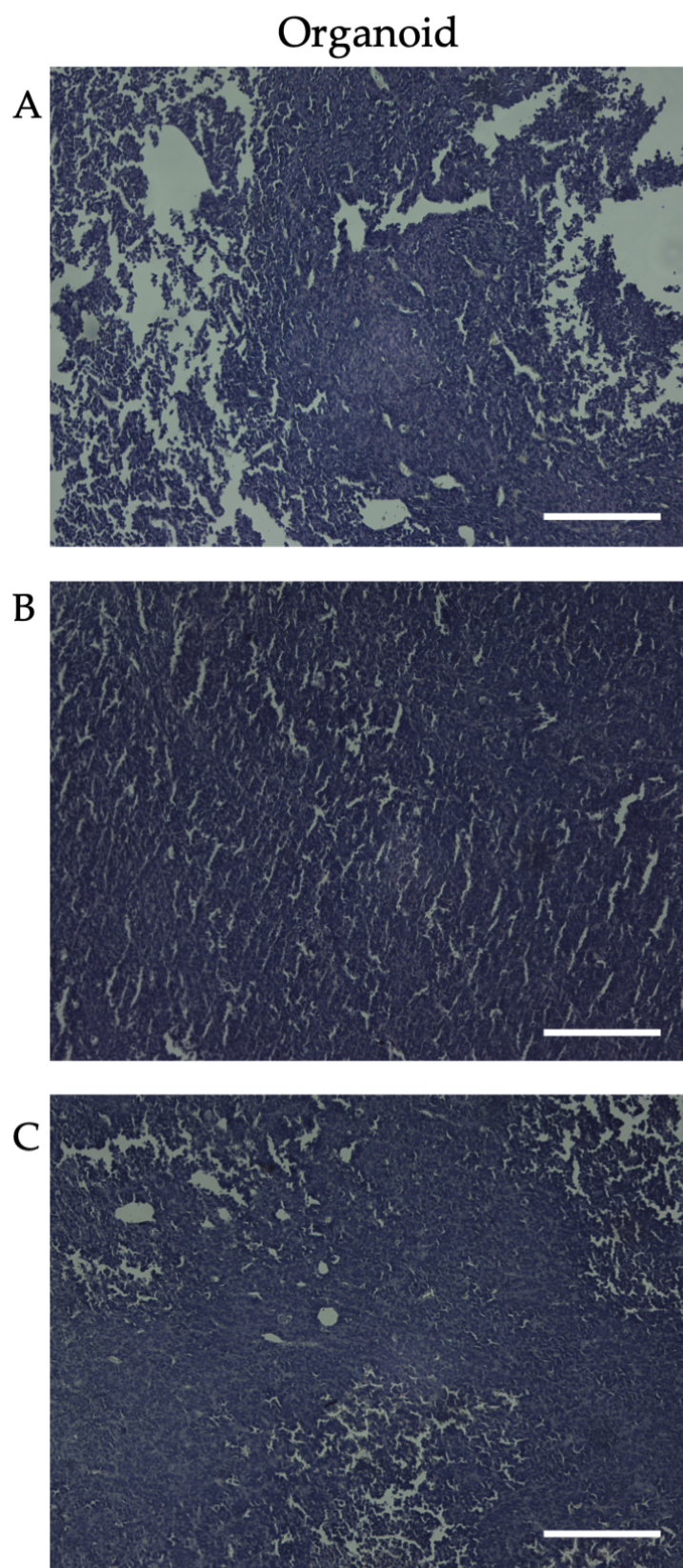


Figure 3.47: H&E staining of PT291 organoid tumour after implantation in SCID mice. Images taken at 100X magnification. Scale bars 200 μ m.

3.2.12.4. IHC CLO xenograft

IHC was performed on a CLO tumour which had been implanted subcutaneously into a SCID mouse. The tumour had very strong staining for cancer stem cell marker ALDH1A1 but low-level staining for CXCR4 (Figure 3.48). Low levels of membrane staining were detected with EpCAM whereas no staining was observed for the cell surface adhesion receptor HCAM (CD44) (Figure 3.49). Pancreatic markers MASPIN and PDX1 were strongly expressed (Figure 3.50). Strong nuclear staining was observed with the proliferation marker Ki67 in <50% of the tumour (Figure 3.51).

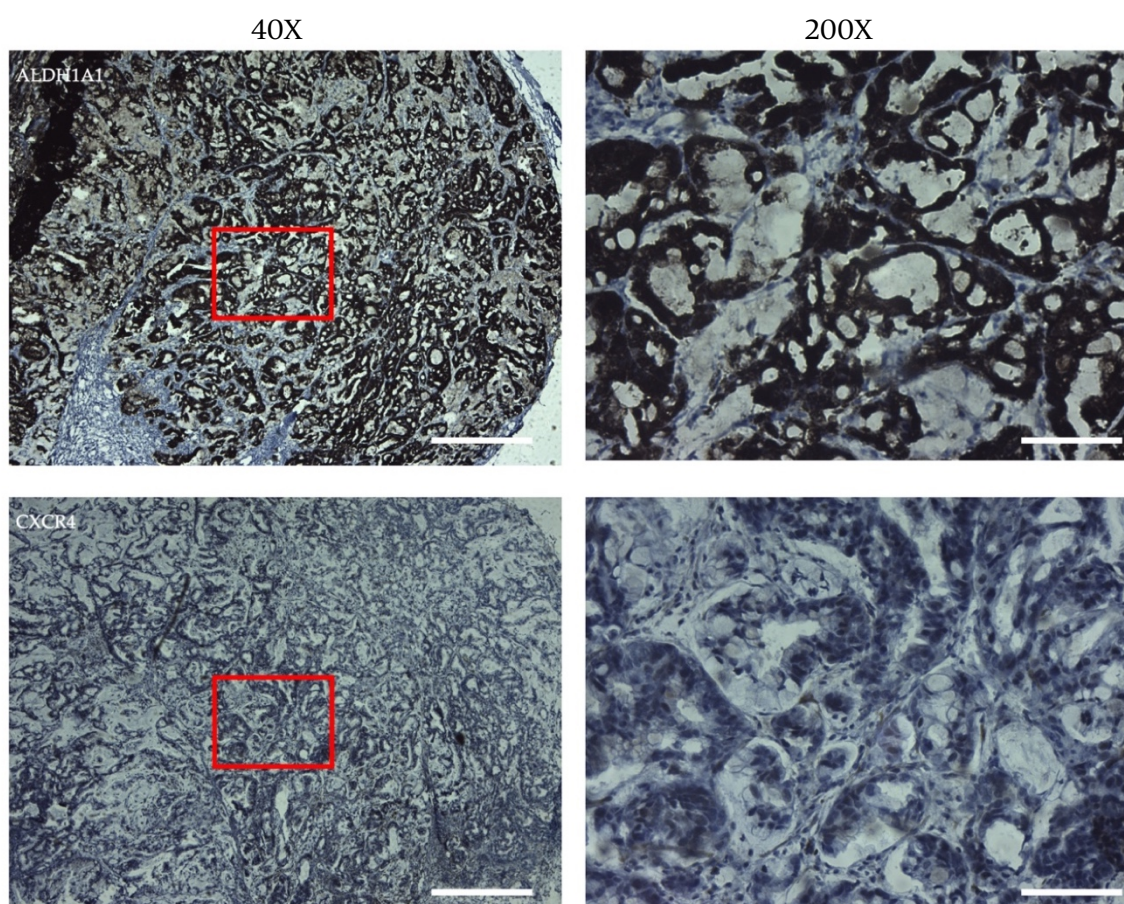


Figure 3.48: IHC of CLO tumour subcutaneously implanted in SCID mouse stained with ALDH1A1 and CXCR4. Red box represents area of 200X image. 40X scale bar 500 μ m, 200X scale bar 100 μ m.

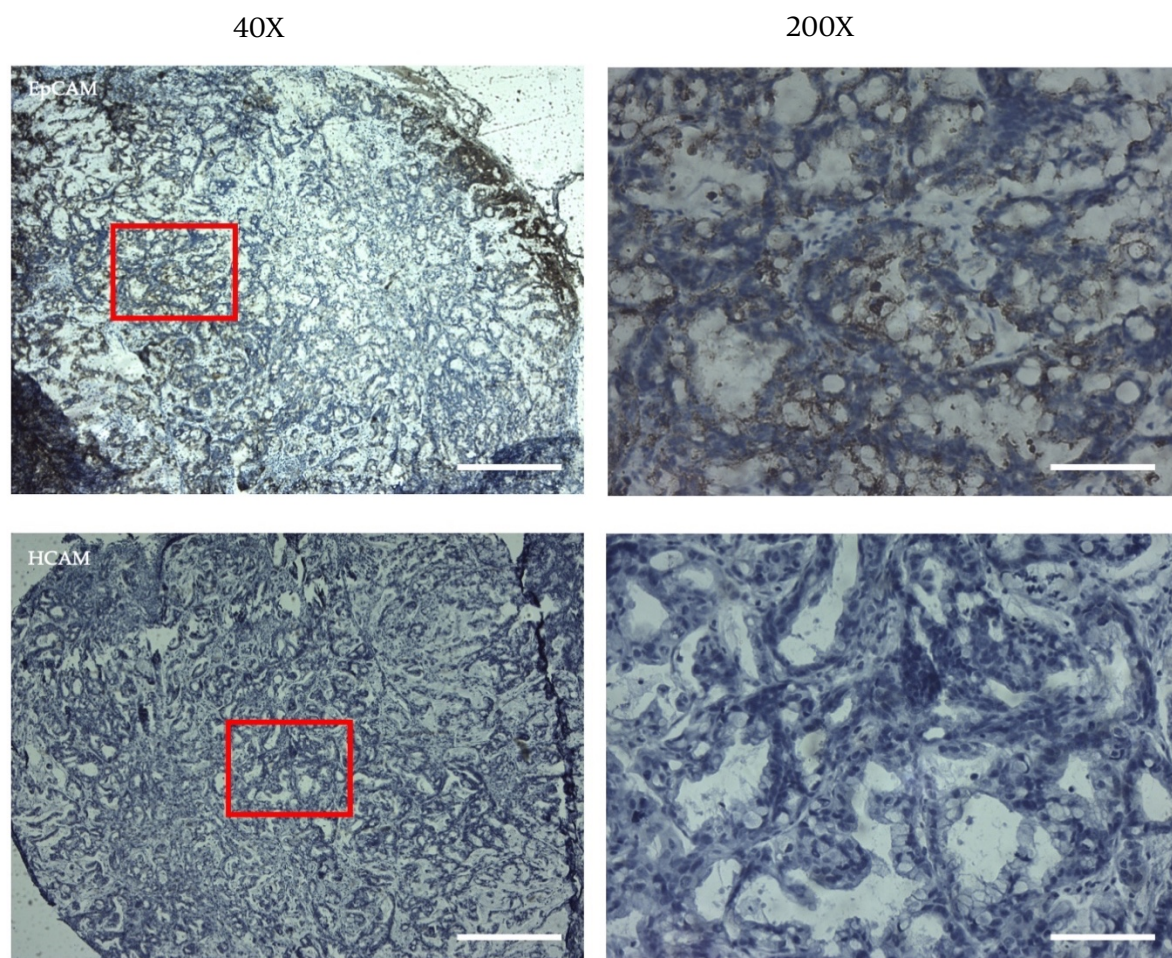


Figure 3.49: IHC of CLO tumour subcutaneously implanted in SCID mouse stained with EpCAM and HCAM. Red box represents area of 200X image. 40X scale bar 500 μ m, 200X scale bar 100 μ m.

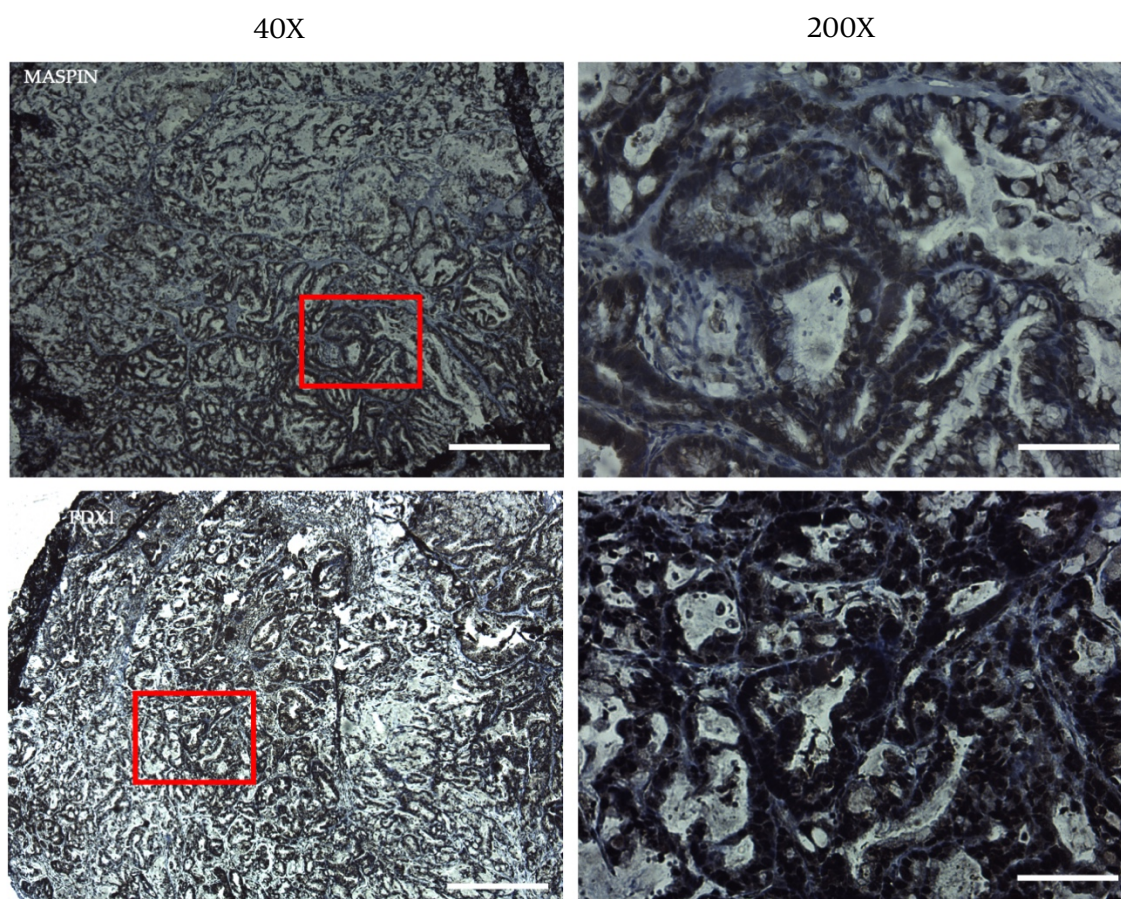


Figure 3.50: IHC of CLO tumour subcutaneously implanted in SCID mouse stained with MASPIN and PDX1. Red box represents area of 200X image. 40X scale bar 500 μm , 200X scale bar 100 μm .

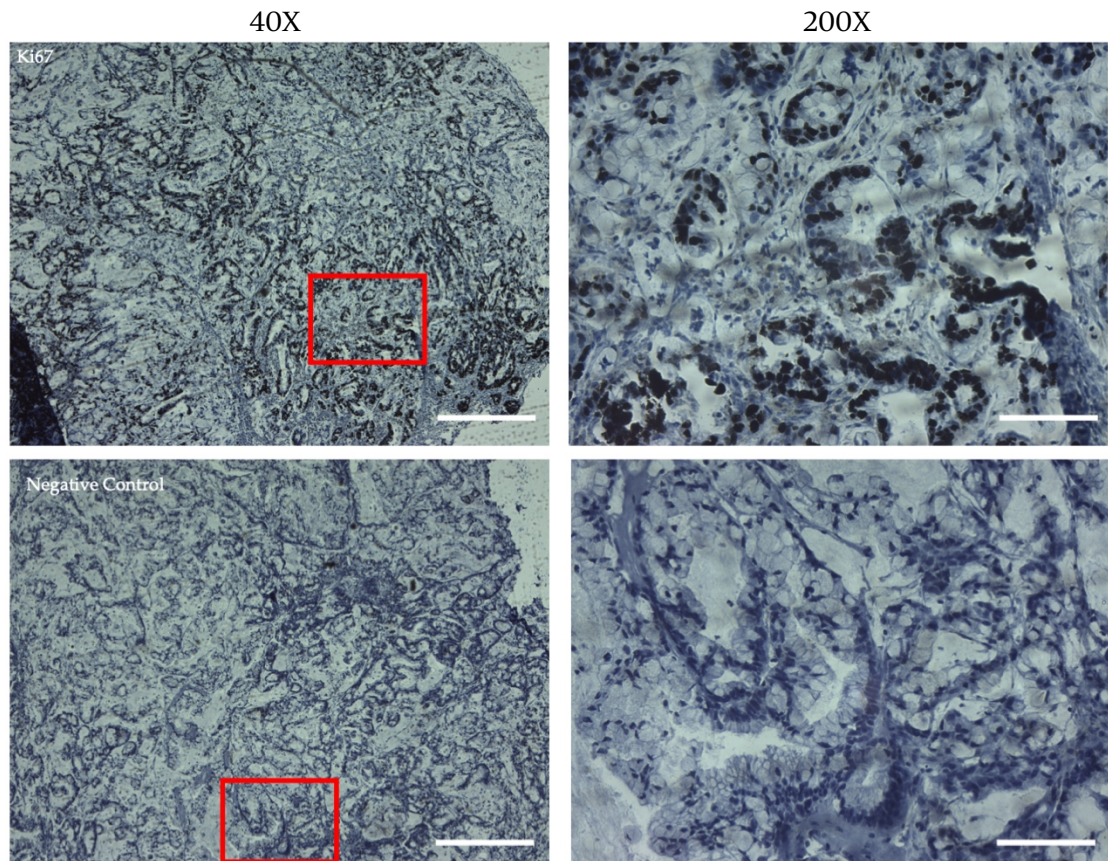


Figure 3.51: IHC of CLO tumour subcutaneously implanted in SCID mouse stained with Ki67 and negative control. Red box represents area of 200X image. 40X scale bar 500 μ m, 200X scale bar 100 μ m.

3.3. Discussion

Two-dimensional, cell-based assays have been an important tool for cancer research for over 50 years, and while the usefulness of cell lines in cancer research is certain, their use as reliable clinical models is debatable [287]. During passaging, cell lines undergo genomic changes including the introduction of point mutations and copy number variation. The use of early passage primary cell lines allows researchers to recapitulate the heterogenous nature of PDAC, however, these also fail to recapitulate the complex 3D structure of the tumour. Organoids are cell models which develop from single cancer stem cells and form complex 3D *in vitro* tumour avatars. The use of organoids in clinical research can overcome many of the limitations of traditional cancer research models, as they contain a more appropriate morphology and signalling pathways, and can be used to study cell-to-cell interaction, drug penetration, response and resistance [207,208].

This chapter describes a method which can be used for the simultaneous generation of organoids and isogenic cell lines from PDAC PDX tumours. In addition, it presents a novel model of PDAC – organoids generated from early passage primary cell lines, which has been termed CLOs. These methods were developed from the PDAC organoid protocol by Boj *et al.* [226] however, this protocol was modified in a number of ways. Firstly, this method uses conditioned media from the L-WRN cell line (ATCC, 3276). This cell line was generated by the Stappenbeck lab and secretes the factors Wnt3A, R-spondin3 and noggin into the medium [288]. The L-WRN cell line produces R-spondin3 rather than R-spondin1, which is typically used in PDAC organoid culture [224,226,232,233,296]. However, R-spondin1 and R-spondin3 act in a functionally redundant manner to permit Wnt/ β -catenin signalling and are used interchangeably in the culture of intestinal organoids [297–299]. While these three growth factors are commercially available in recombinant form, it is costly to maintain the high volume of media for large-scale cultures that are required for standard assays using organoids. The use of L-WRN is a cost-effective way of providing the cells with the supplements that they require for growth.

In addition, this method deviates from the Boj *et al.* [226] protocol through the source of tumour for the generation of PDAC organoids. This work utilises an existing resource which was available within the NICB – a PDAC PDX tumour biobank. While the use of PDX tumours for the generation of organoids (termed PDxO) and cell lines have been described previously for pancreatic [300], liver [301], prostate [302], breast [303,304], and lung and ovarian [305] cancers, this protocol is the first to simultaneously generate organoids and primary cell lines from a PDX tumour sample [306]. Although PDX tumours remain stable during passaging, tumour-associated stroma is replaced with murine stromal cells, which account for up to 90% of the cells present in a PDX tumour [205,307]. When establishing primary cell lines from PDX samples, these murine fibroblasts can often quickly out-grow and outnumber the PDAC cancer cells *in vitro*, resulting in a completely unrepresentative model for PDAC [308]. Similar to the work by Borodovsky *et al.* [305] this protocol uses a mouse cell depletion kit to ensure the removal of infiltrating mouse cells from the tumour sample. In addition this method also uses a mouse cell specific PCR screening protocol to validate that all mouse cells have been removed from the sample [290].

One of the key components of organoid culture is the 3D environment in which the cells require grow [309]. To overcome the need for the use of expensive Ultra-Low Attachment Multiple Well Plates in organoid culture, these methods use polyHEMA-coated plates to prevent the cells from attaching to the culture plate [310]. However, this method did utilise the cell's penchant for attaching to the plate in the development of isogenic primary PDAC cell lines. This method allows for the rapid development of primary cell lines, which can be quickly, and cost effectively scaled up and manipulated.

These early passage primary cell lines were cultured using organoid conditions for two passages, which resulted in the generation of CLOs. To validate CLOs as representative organoid models, this study assessed if the organoids and CLOs were phenotypically similar. It was identified that the CLOs and organoids had similar morphologies, proliferation rates and responses to gemcitabine. To test this novel method, and to validate CLOs as a novel model of pancreatic cancer, also the

establishment and phenotypic properties of CLOs using commercially available pancreatic cancer organoid lines (PDM37, PDM41 and PDM106) was assessed. PDM37, which was developed from squamous cell carcinoma subtype of pancreatic cancer, resulted in the successful generation of a primary cell line and CLOs. However, these cells grew very slowly, as demonstrated in the proliferation assay where the cell line reached a maximum confluence $17 \pm 1.5\%$ after 10 days, and therefore were excluded from further analysis.

Molecular profiling of cancer stem cell markers ALDH1A1, CXCR4, EpCAM and HSCAM was performed in our models. Expression of ALDH1A1 was low in both the PDM41 and PT291 primary cell lines, and CLOs, but had high levels of expression in their matched organoid models. ALDH activity selectively defines enhanced tumour-initiating cells [311]. ALDH1A1 is involved in retinoic acid metabolism, and has a role in proliferation and differentiation [312]. CXCR4, a cancer stem cell marker, had consistent low expression in all primary cell lines, with increased expression in matched organoid and CLO models. It has been shown previously that hypoxia promotes the expression of CXCR4 in PDAC [313]. This further validates our use of organoids and CLOs as *in vitro* models of pancreatic cancer, as PDAC tumours are significantly more hypoxic than other clinical tumours [314]. This can be achieved with cell lines however, to induce hypoxia, monolayers of cells, flasks are placed in a gas-controlled chamber [315]. The expression of the transmembrane glycoprotein, EpCAM was retained in all models, and is known to be highly expressed in PDAC and established cell lines [316,317]. Expression of HSCAM was high in organoid and CLO models compared to cell lines, with the exception of PT291, which had similar expression levels in cell line and organoid models.

Interestingly, the different models generated from the same sample show a consistent pattern of expression of pancreatic specific markers MASPIN and PDX1. Increased expression of PDX1 is observed in the cell line and CLO compared to the matched organoids in PDM106, PT127 and PT291. PDX1 is a transcription factor which is essential for pancreatic development and differentiation of all pancreatic cell lineages [318]. The over expression of PDX1 in PDAC highlights its role as an

oncogene [150]. However, PDX1 expression may be linked to the changing stages during cancer development and progression. PDX1 loss or downregulation is associated with a more aggressive phenotype, and the loss of PDX1 observed in subsets of cells undergoing EMT [68,319]. Furthermore, upregulation of key stem cell markers, transcription factors *NANOG*, *OCT4* and *SOX2* in CLOs compared to 2D primary cultures indicates a pluripotent stem cell state and suppression of lineage-specific genes in these models [320].

An unexpected result in this chapter was the difference in response to treatment with gemcitabine of matched 2D cell lines and 3D organoids and CLOs. In our study, the organoids and matched CLOs had increased sensitivity to treatment with gemcitabine compared to the 2D matched cell line. Organoids and CLOs, in all samples achieved an IC_{50} with concentrations as low as 95 nM. In comparison, all the 2D cell lines had reduced sensitivity to gemcitabine, with only one line (PDM41-CL) achieving an IC_{50} at a concentration of 3 μ M. Usually, cells are more sensitive to drug treatment in 2D cultures than in 3D cultures [321]. Wen *et al.* [322] found 3D cultured PANC1 and MiaPaCa-2 cells demonstrated significantly reduced sensitivities to gemcitabine and 5FU compared to 2D monolayer cultures. However, Miyamoto *et al.* [323] observed that when the MiaPaCa-2 cell line was grown in ECM proteins, the 3D cells were sensitive to treatment with gemcitabine at concentrations of 0.002-0.1 μ M. This study found that poorly differentiated tumour cells were less sensitive to the cytotoxicity of anticancer drugs compared to well differentiated tumour cells when grown in 3D embedded in ECM. This correlates with the responses seen in our models as the organotypic 3D models developed from grade 3 tumours (PDM37 and PDM106) showed an increased sensitivity to treatment with gemcitabine compared to the matched 2D cell line, in comparison with models developed from PT127, a grade 2 tumour.

To further validate these novel models, RNA-sequencing was performed on PT127 cell line, CLO and PDX samples, and PT291 cell line, organoid and CLO samples. This analysis highlighted by PCA that the PT291 organoids and CLOs clustered together as did the PT127 PDX tumour and CLOs compared to their respective 2D primary cell lines. Additionally, hierarchical clustering comparing the

dysregulated genes between PT291 and PT127 CLOs versus the 2D primary cell lines, revealed a clear pattern as shown by heat map. The three top up and downregulated genes in the CLOs vs primary cell lines were *TFF3*, *PLEKHBI*, *LTB* and *FOSB*, *CAVI*, *AHNAK2*, respectively. Interestingly, *FOSB* is a proto-oncogene, and *CAVI* expression has been linked with mammary tumorigenesis and increased lung metastasis [324,325]. *AHNAK2* has also previously been found to be upregulated in PDAC [326]. The upregulated signature also contains increased expression of genes associated with cilia (*KIF12*, *RSPH1* and *CCDC114*) [327–329]. Cilia have been found to play an important role in modulating signalling cascades including the Wnt signalling pathway, and are essential in pancreatic tissue organisation [330,331]. The upregulation of stem cell markers *ALDH1A1*, *ALDH1B1* and *LRIG1* may indicate restored embryonic programs and dedifferentiation pattern among our CLOs [312,332,333]. This dedifferentiation theory is supported by a previous study which showed that when implanted orthotopically, organoids recapitulate tumour progression from precursor PanIN lesions to metastatic carcinoma [226]. Interestingly, genes downregulated in PDX tumour/organoids and CLOs include genes which have increased expression in cancers, including potential PDAC biomarkers (*CYR61*, *PLAU*, *FOSB*, *ECM1*, *LIF*, *GJB3* and *CAVI*) [325,326,334–339]. There is also an increased expression of tumour suppressor genes *LRIG1* and *SEMA3F* [332,340]. Molecular subtyping of the newly derived models based on the classification method from Moffitt *et al.* [65] shows the PT291 cell line models have a higher probability of being of the more aggressive, basal subtype of the disease (OR=15.47) than the matched organoids (OR=6.89). This, in addition to the gene signature may highlight the “phenotypic switch” of the 2D primary cell lines as the cells invade through ECM, attach and proliferate, resulting in increased motility, invasion and proliferation of cells by upregulation of oncogenic pathways.

Finally, to assess if the organoids, CLOs and primary cell line retained their original tumour characteristics after being propagated long term as organoids, PT291 cells were re-implanted as a subcutaneous xenografts. All cell models resulted in tumour growth, with measurable tumours for the organoid and CLO tumours forming after 23 days. In comparison, the original PDX tumour was measurable

160 days post implantation. Interestingly, organoid and cell line tumours were faster growing than the matched CLO tumours, and both the organoid and cell line had a similar growth rate to that of the original tumour. H&E staining of the tumours showed clear morphological differences between the tumours developed from the organoids, cell lines and CLOs. Where the organoids and cell lines formed as large cell clusters, with no tumour-like structures forming, however, the CLO tumours formed PanIN-like lesions, and are more representative of an *in vivo* tumour. Other work involving the use of organoids as PDX tumours involved the orthotopic implantation of cells, this resulted in the organoid progressing through all stages of tumour development from PanIN to a PDAC tumour [226]. As there are a number of obstacles to the use of orthotopic implantation in the pancreas, including the difficulty of the implantation surgery, and the requirement for specialist imaging equipment, subcutaneously implanted CLOs are a potential model for organotypic PDX tumours [341].

While this study illustrates the successful generation, characterisation and validation of a novel PDAC *in vitro* cell model, it has several limitations. The limited availability of matched samples in our RNA-sequencing study limits the conclusions we can derive from this work. Our work is also highly reliant on the use of ECM to mimic the laminin-rich pancreas, however, ECM may not fully recapitulate the specific nature of the pancreas microenvironment [342]. ECM is derived from Engelbreth-Holm-Swarm mouse tumours, therefore lot-to-lot variability in protein concentration and molecular composition can affect the growth of organoids and responses to drug treatment [343]. Bulk purchase of ECM was required to maintain reproducibility in experiments.

Further limitations of our study and others studying PDAC organoids [226,232,320] include the lack of a relevant microenvironment integrating the immune and stromal components, as tumour derived organoids are exclusively ductal cells, but not of other pancreatic cell lineages [320]. Tsai *et al.* [238] constructed complex organotypic models containing tumour, stromal and immune components of the tumour microenvironment. This study found that cells cultured as 3D spheroids, such as the established PDAC cell line, PANC-1, a mouse

xenograft cell line, and the A549 lung cancer cell line are phenotypically distinct from primary organoids, as the spheroids formed homogenous, non-lumen forming spheres. However, the cell line models used by Tsai *et al.* [238] are not primary cultures and were not compared to their isogenic matched organoid or PDX tumour, therefore are not directly comparable and tumour to tumour bias may exist.

This study has built on the scientific advancements of PDX and cancer organoid models to bridge the *in vivo/ in vitro* gap in PDAC research [202,205,217,218,224,226,344]. Given the difficulties surrounding obtaining patient samples, this study results in developed a novel protocol to establish organoids from the PDX tumour biobank available in the NICB in DCU. This chapter also presents a novel method for the establishment of 2D primary cell lines from PDAC PDX tumour samples and recapitulation of the cell lines to CLOs. RNA-sequencing has shown that the organoid and CLO models are transcriptomically similar and are more representative of the PDX tumour than the isogenic cell line. Using these organoids and CLOs generated from PDX tumours, as well as commercially available organoid lines, this shows that the novel model, CLOs are morphologically and phenotypically similar to their parental organoids. These CLOs can also be implanted in mice for the generation of tumours representative of patient tumours. This method can be utilised for all cancer types, allowing laboratories to utilise PDX samples already available for the development of organoids and primary cell lines.

Chapter 4. Experimental and *in silico* assessment of the role of MODY pathway genes in PDAC

4.1. Introduction

An association between PDAC and diabetes has been previously identified, approximately 80% of PDAC patients have glucose intolerance, or diabetes at the time of their diagnosis [123]. The majority of diabetes associated with pancreatic cancer is diagnosed with the disease, or within two years of diagnosis [7].

In a pathway analysis study previously performed in our lab by Walsh *et al.* [122] the MODY pathway was identified as the most significant pathway associated with risk of developing PDAC. Regulation of beta-cell development was the second top pathway associated with risk of PDAC development. Both pathways had SNPs in the genes *HNFLA*, *HNFB*, *HNF4G*, and *HNF4A*.

MODY is a rare form of diabetes which accounts for 1-2% of cases [133]. MODY typically develops before the patient is 25 years. Unlike T1DM and T2DM which are multifactorial disorders, involving several genes with a strong environmental component, MODY is a monogenic disease. MODY patients typically have a family history of diabetes in two or more generations [134]. Fourteen autosomal dominant forms of MODY, as described in Table 1.4 (Chapter 1) have been identified, manifest through mutations in genes which disrupt insulin production or release. It is usually caused by mutations in genes which act as transcription factors. The four most common forms of MODY – 1, 2, 3 and 5 account for 90% of the cases. Diagnosis of MODY and the identification of the genetic subtype is important, as age of onset, response to treatment and hyperglycaemia levels differ [135].

4.1.1. Aims

Aim 1: Identification and validation of regulatory SNPs capable of modulating genes involved in the MODY pathway in PDAC, through *in silico* SNP analysis and dual luciferase reporter assays.

Aim 2: Investigate the role of genes associated with SNPs of interest in PDAC, through using publicly available online databases and RNA-sequencing data.

Aim 3: Assess the role of genes of interest in PDAC by developing knockout organoids using CRISPR. Analysing the role of genes of interest by performing CUT&RUN to profile protein-DNA interactions.

4.2. Results

4.2.1. Functional analysis of SNPs and genes associated with the MODY pathway identified through GWAS pathway analysis

All candidate SNPs identified in PDAC GWAS pathway analysis previously performed in our group (Figure 4.1) and SNPs in LD with a $R^2 > 0.8$ and a regulomeDB score ≤ 3 were assessed for prioritisation for functional analysis 2.12. A $R^2 > 0.8$ was selected to limit the number of SNPs to be assessed, a regulomeDB score ≤ 3 was selected as scores ≥ 4 have minimal binding evidence.

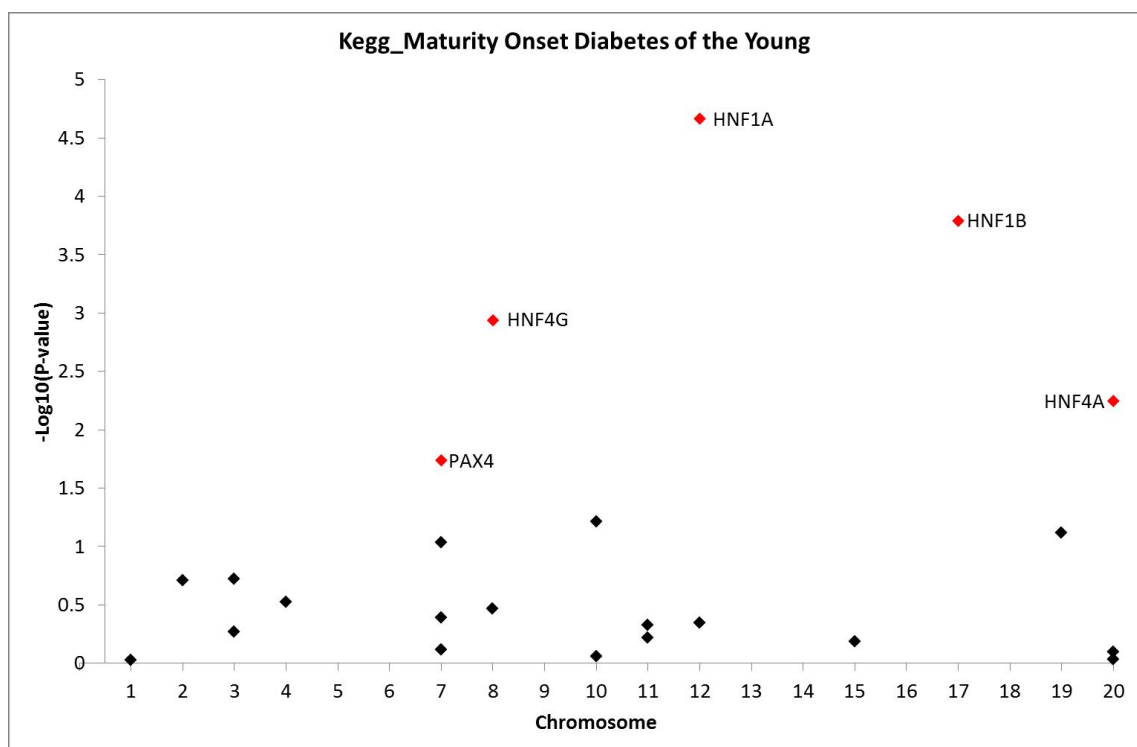


Figure 4.1: Genes associated with PDAC in KEGG maturity onset diabetes of the young (MODY) pathway ($P = 5.10 \times 10^{-7}$). A total of five genes (red marker, $P < 0.02$) were identified from the gene set as contributing the most to PDAC association in the GWAS pathway analysis. All of these genes were selected for further analysis.

Figure taken from Walsh et al. [122].

4.2.1.1. SNPs identified in MODY genes associated with PDAC

GWAS SNPs, GWAS pathway SNPs and SNPs identified through literature searches resulted in the identification of 27 SNPs in 8 genes from the MODY pathway associated with PDAC. The genes identified, and their role in the normal pancreas

are outlined in Table 4.1. The SNPs identified are listed in Table 4.2, SNPs in LD with these were identified using LD link. Duplicate SNPs were removed, LD pairwise analysis tables were generated using R^2 values from LD-matrix (<https://ldlink.nci.nih.gov/?tab=ldmatrix>) and heatmaps were generated using Prism GraphPad v9. The R^2 values represent the non-random association of two SNPs [345]. To refine the SNPs identified for post-GWAS functional validation, the best candidate from the SNP blocks was identified using the methods shown in Figure 4.2 and brought forward for functional analysis.

Table 4.1: MODY pathway genes associated with SNPs identified in literature searches.

<i>Gene Name</i>	<i>Symbol</i>	<i>Function</i>
Carboxyl ester lipase	<i>CEL</i>	Enzyme
Hepatocyte nuclear factor 1-alpha	<i>HNFI A</i>	Transcription factor
Hepatocyte nuclear factor 1-Beta	<i>HNFI B</i>	Transcription factor
Hepatocyte nuclear factor 4 - alpha	<i>HNF4 A</i>	Transcription factor
Hepatocyte nuclear factor 4 - gamma	<i>HNF4 G</i>	Transcription factor
Nuclear receptor subfamily 5 group a member 2	<i>NR5A2</i>	Transcription factor
Paired box gene 4	<i>PAX4</i>	Transcription factor
Pancreatic and duodenal homeobox 1	<i>PDX1</i>	Transcription factor

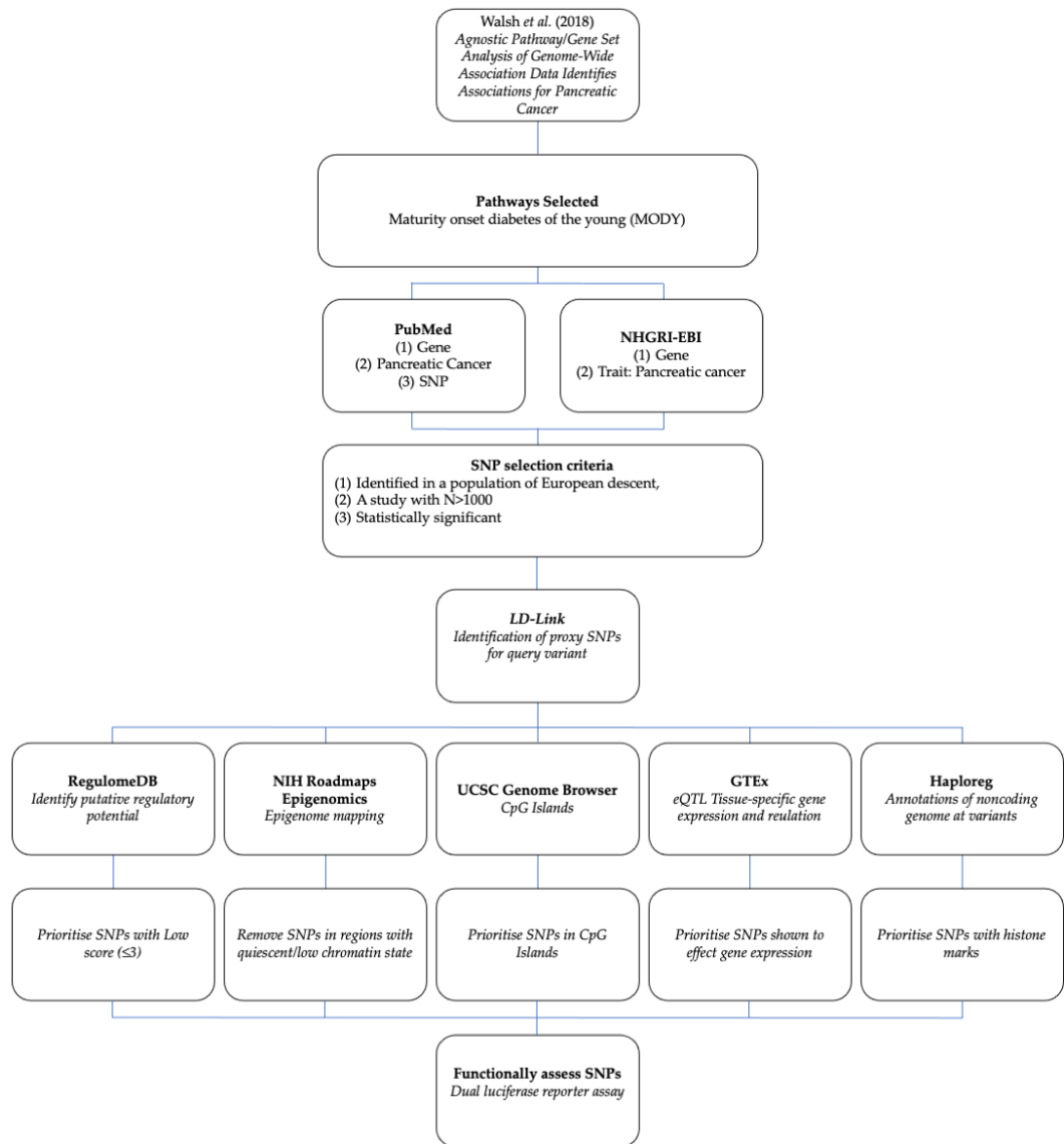


Figure 4.2: The criteria used for SNP identification, in silico assessment and prioritisation of SNPs from the MODY pathway from the study by Walsh et al. [122].

Table 4.2: SNPs located in MODY pathway genes predicted to have a role in the development of PDAC.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	P-value	Cases Controls	Allelic OR (95% CI)	Location/ Function	Ref
CEL\$	rs488087	9	135946599	C T	5.00E-04	36 78	4.72 (1.87-11.91)	Exon II CEL	[346]
HNF1A	rs2244608#	12	121416988	A G	1.54E-06	9,040 12,496	1.11 (1.07-1.16)	Intronic	[122]
HNF1A	rs1169296#	12	121428407	A G	2.63E-07	9,040 12,496	1.12 (1.07-1.17)	Intronic	[122]
HNF1A	rs1182933*	12	121454621	T C	3.49E-07	9,040 12,496	1.11 (1.07 - 1.15)	316 bp 5' of C12orf43	[116]
HNF1A	rs2464196	12	121435426	C T	6.30E-05	3,851 3,934	1.15 (1.07-1.24)	Missense	[121]
HNF1A	rs1169300	12	121431225	G A	7.90E-05	3,851 3,934	1.15 (1.07-1.23)	Intronic	[121]
HNF1A	rs7953249	12	121403724	A G	2.30E-04	3,851 3,934	1.13 (1.06-1.21)	2.7 kb 3' of HNF1A-AS1	[121]
HNF1B	rs12951345#	17	36077863	A C	1.38E-06	9,040 12,496	0.88 (0.84-0.93)	Intronic	[122]
HNF1B	rs7223387#	17	36082473	T G	3.46E-06	9,040 12,496	0.89 (0.85-0.94)	Intronic	[122]
HNF1B	rs4795218*	17	36078510	A G	1.32E-08	9,040 12,496	0.88 (0.84-0.92)	Intronic	[116]
HNF1B	rs4794758	17	36080428	C T	7.30E-04	3,851 3,934	0.88 (0.82-0.95)	Intronic	[121]
HNF4A	rs1853150#	20	43081164	G T	1.64E-05	9,040 12,496	1.10 (1.05-1.15)	Upstream transcript variant	[122]
HNF4G	rs2943547#	8	76451098	A G	2.31E-05	9,040 12,496	1.09 (1.05-1.14)	Intronic	[122]

<i>HNF4G</i>	rs1913641#	8	76483239	T	G	1.55E-05	9,040 12,496	1.10 (1.05-1.14)	4.2 kb 3' of HNF4G	[122]
<i>HNF4G</i>	rs2941471*	8	76470404	G	A	6.60E-10	9,040 12,496	0.89 (0.85-0.93)	Intronic	[116]
<i>HNF4G</i>	rs2977926	8	76448847	T	G	3.40E-04	3,851 3,934	0.88 (0.82-0.94)	Intronic	[121]
<i>HNF4G</i>	rs1805100	8	76476396	G	A	3.20E-05	3,851 3,934	1.15 (1.08-1.23)	3' prime UTR	[121]
<i>NR5A2</i>	rs2737621	1	199997403	T	C	6.30E-05	3,851 3,934	1.22 (1.11-1.34)	Intronic	[121]
<i>NR5A2</i>	rs2821347	1	199985792	G	A	2.50E-05	3,851 3,934	1.21 (1.11-1.33)	Intronic	[121]
<i>NR5A2</i>	rs2816939	1	199985700	T	C	2.00E-05	3,851 3,934	1.22 (1.11-1.33)	Intronic	[121]
<i>NR5A2</i>	rs2821367	1	200016146	T	C	1.60E-06	3,851 3,934	1.18 (1.10-1.27)	Intronic	[121]
<i>NR5A2</i>	rs3790844*	1	200007432	T	C	2.45E-10	3,851 3,934	0.77 (0.71-0.84)	Intronic	[113]
<i>NR5A2</i>	rs3790843	1	200010824	G	A	1.34E-03	3,851 3,934	0.83 (0.74-0.93)	Intronic	[113]
<i>NR5A2</i>	rs2816938	1	199985368	T	A	4.88E-15	5,107 8,845	1.2	1.4kb 5' of U6	[117]
<i>PAX4</i>	rs62483175#	7	127247114	C	T	2.21E-03	9,040 12,496	0.78 (0.67 - 0.91)	3.2kb 3' of PAX4	[122]
<i>PAX4</i>	rs118117270#	7	127263589	T	C	4.23E-04	9,040 12,496	0.67 (0.54 - 0.84)	7.6kb 5' of PAX4	[122]
<i>PDX1</i>	rs9581943*	13	28493997	G	A	2.40E-09	7,683 14,397	1.15 (1.10-1.20)	2kb upstream	[114]

*This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, allelic odds ratio (CI 95%), and reference paper in which SNP was identified. *Highlights GWAS SNPs. #Highlights GWAS SNPs. \$Study with low number of cases/controls.*

4.2.1.2. Analysis of *CEL*-associated SNPs

Carboxyl ester lipase (*CEL*) is a gene located at 9q34.13. It is 10 kb in length and contains 11 exons. *CEL* is the causative gene of MODY8. *CEL* has high levels of expression in the pancreas and lactating mammary glands. It is secreted by the pancreas to the duodenum where its role is to hydrolyse fats, cholesteryl esters and fat-soluble vitamins. A full *in silico* analysis of the potential functionality of rs488087 was performed as outlined in Table 4.3. University of California Santa Cruz (UCSC) genome browser view of the region surrounding the SNP are presented in Figure 4.3, no SNPs in LD were within the selection criteria set out ($R^2 > 0.8$, regulomeDB score ≤ 3).

Table 4.3: Full analysis of SNP located in the CEL gene. Underlined SNP selected for functional analysis.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State	Location/ Function	Histone Markers	CpG Island	eQTL (NES)
<u>CEL</u>	<u>rs488087</u>	<u>2</u>	<u>135946599</u>	<u>C</u> <u>T</u>	<u>3a</u>	<u>Strong</u> <u>transcription</u>	<u>Exon 11</u> <u>CEL</u>	<u>H3K4me1</u> <u>H3K4me3</u> <u>H3K27ac</u>	<u>Yes</u>	<u>↓ -0.35</u> <u>Oesophagus</u>

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG island. eQTL includes NES, gene and tissue.

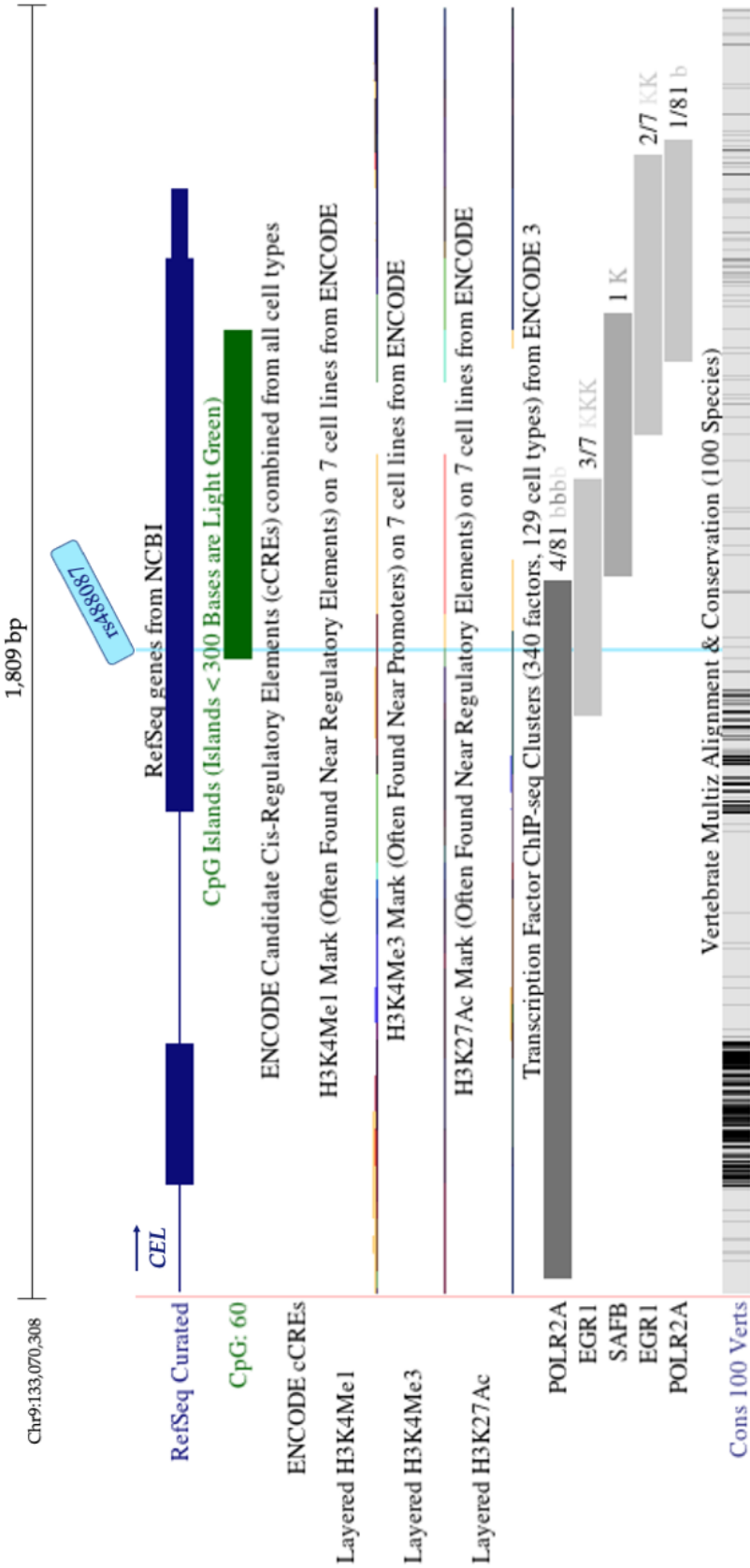


Figure 4.3: University of California Santa Cruz (UCSC) Genome Browser view of CEL gene regions surrounding rs488087 (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factor from ENCODE.

4.2.1.3. Analysis of *HNFLA* associated SNPs

Hepatocyte Nuclear Factor-1-Homeobox-A (*HNFLA*) is a gene located at 12q24.31 and contains 9 exons. *HNFLA* has high levels of expression in the kidney, duodenum, liver and pancreas, where it acts as a transcription factor with *HNFB*. The GWAS SNP (rs1182933), two GWAS pathway SNPs (rs1169269 and rs2244608), three SNPs identified in literature searches (Table 4.2) and eight SNPs in LD are outlined in a LD block ($R^2 > 0.8$; $n=14$ SNPs) and highlighted in a heatmap (Figure 4.4). The full *in silico* analysis of the 14 SNPs associated with *HNFLA* is outlined in Table 4.4. SNPs were then prioritised for functional analysis based on the criteria outlined in Figure 4.2. Rs2258287 was selected for functional analysis as it is located at an active transcription start site, within a CpG island and has histone marks H3Kme1 and H3K4me3 in the pancreas. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 4.5.

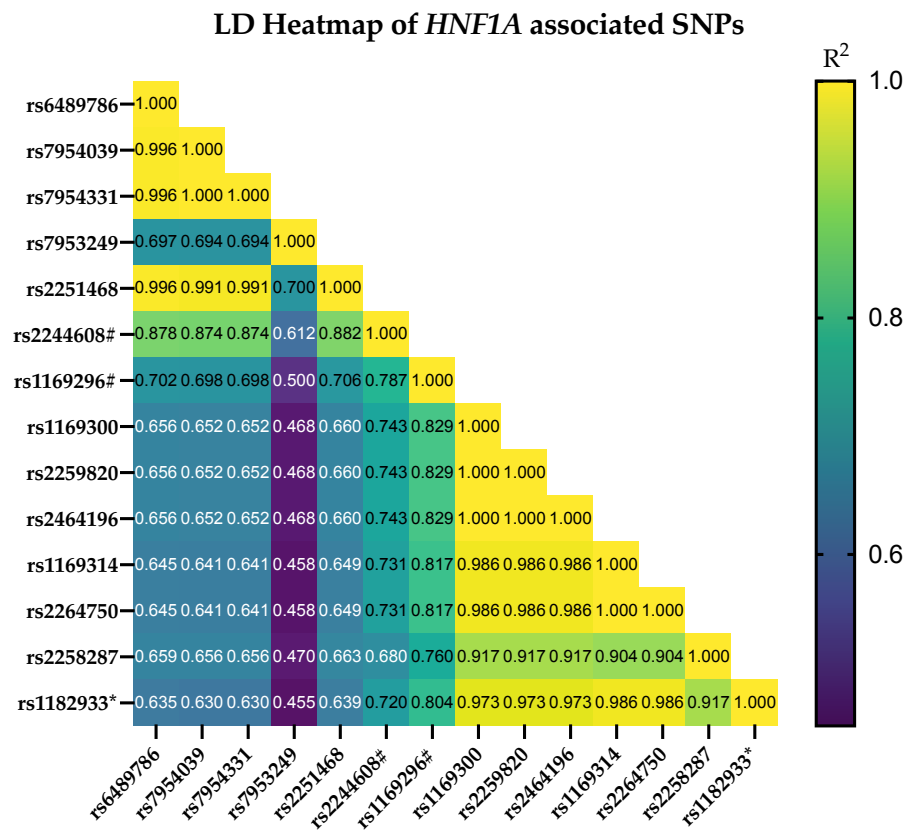


Figure 4.4: LD heatmap of *HNFLA* SNPs of interest based on R^2 values in European populations (LDLink). The R^2 values represent the non-random association of two SNPs. *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

Table 4.4: Full in silico analysis of SNPs associated with the gene *HNFA*. Underlined SNPs selected for functional analysis.

Gene	Marker	Chr	Position	Alleles Ref Alt	RegDB	Chromatin State <i>Pancreas</i>	Location/ Function	Histone Markers <i>Pancreas</i>	CpG Islands	eQTL (NES)
<i>HNFA</i>	rs1169300	12	121431225	G A	4	Strong transcription	Intronic	—	—	↓ -0.28, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs2259820	12	121435342	C T	2b	Strong transcription	Synonymous	H3K4me1	—	↓ -0.28, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs1169296#	12	121428407	A G	3a	Strong transcription	Intronic	—	—	—
<i>HNFA</i>	rs2464196	12	121435426	G A	4	Strong transcription	Missense	H3K4me1	—	↓ -0.28, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs2251468	12	121405126	C A	1f	Strong transcription	1.3 kb 3' of HNFA-ASI	H3K4me1	—	↑ 0.26, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs1182933*	12	121454621	C T	3a	Quiescent/ Low	2KB upstream	H3K4me1	—	↓ -0.28, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs2244608#	12	121416988	A G	2b	Active TSS	Intronic	H3K4me3 H3K27ac	—	↓ -0.32, CI2orf43, Heart
<i>HNFA</i>	rs6489786	12	121397875	A G	2a	Strong transcription	8.5 kb 3' of HNFA-ASI	—	—	↑ 0.26, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs7954039	12	121398654	A C	3a	Strong transcription	7.8 kb 3' of HNFA-ASI	—	—	↑ 0.26, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs7954331	12	121398657	G T	3a	Strong transcription	7.8 kb 3' of HNFA-ASI	—	—	↑ 0.26, CI2orf43, Artery - Tibial
<i>HNFA</i>	rs2264750	12	121450165	C T	2b	Strong transcription	Intronic	—	—	↓ -0.28, CI2orf43, Artery - Tibial

MODY Pathway

<u>HNF1A</u>	<u>rs2258287</u>	<u>12</u>	<u>121454313</u>	<u>C</u>	<u>A</u>	<u>1f</u>	<u>Active TSS</u>	<u>7 bp 5' of</u> <u>C12orf43</u>	<u>H3Kmel</u> <u>H3K4me3</u>	<u>Yes</u>	<u>↓ -0.28, C12orf43,</u> <u>Artery - Tibial</u>
HNF1A	rs1169314	12	121443116	A	G	2b	Strong transcription	Intronic	—	—	—
HNF1A	rs7953249	12	121403724	G	A	1f	Strong transcription	Splice region	H3K4mel	—	<u>↑ 0.23, C12orf43,</u> <u>Adipose</u>

*This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG island. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.*

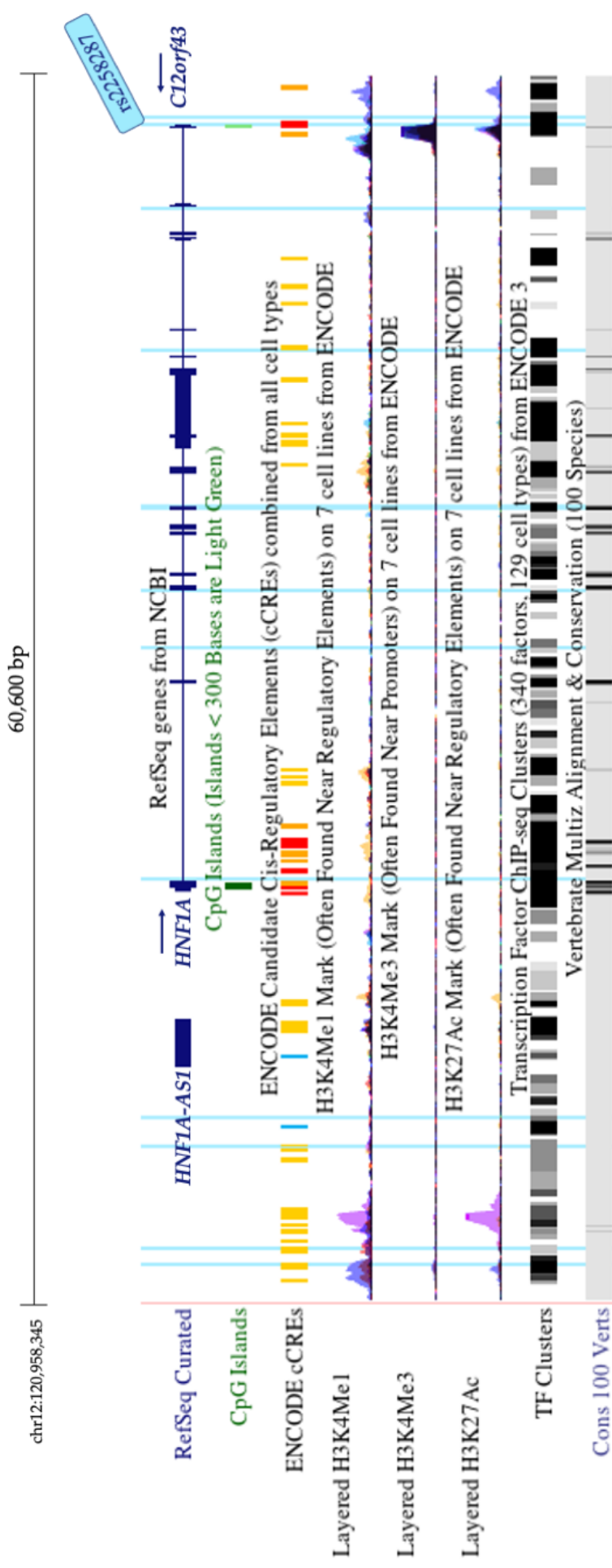


Figure 4.5: UCSC Genome Browser view of HNF1A gene regions surrounding SNPs of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

4.2.1.4. Analysis of *HNFI*B associated SNPs

Hepatocyte Nuclear Factor-1-Homeobox-B (*HNFI*B) is a gene located at 17q12 and contains 11 exons. *HNFI*B has high levels of expression in the kidney, gallbladder, colon and pancreas, where it acts as a transcription factor with *HNFI*A. The GWAS SNP (rs4795218), two GWAS pathway SNPs (rs12951345 and rs7223387), one SNP identified in literature searches and one SNP in LD are outlined in a LD block ($R^2 > 0.8$; $n=5$ SNPs) and highlighted in a heatmap (Figure 4.6). The full *in silico* analysis of the SNPs associated with *HNFI*B is outlined in Table 4.5, and SNPs were then prioritised for functional analysis based on the schematic outlined in (Figure 4.2). SNPs rs718961 and rs4795218 were selected for functional analysis, as both were in enhancer regions, had a high regulomeDB score, and had histone marks in pancreatic tissue. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 4.7.

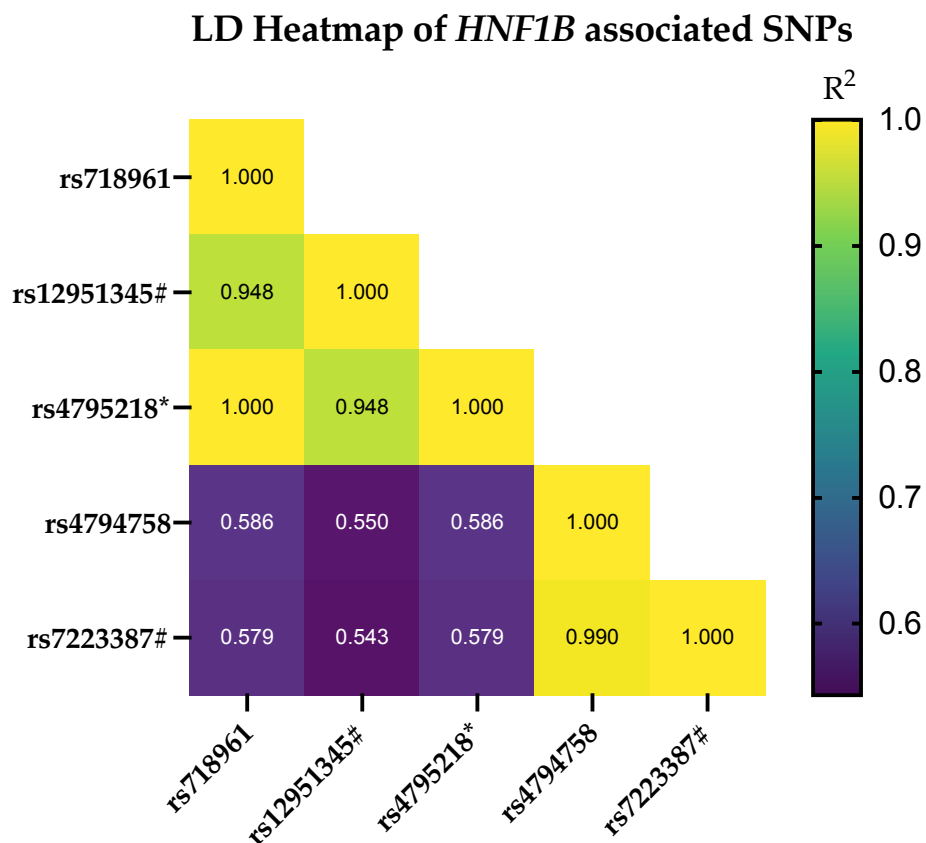


Figure 4.6: LD heatmap of *HNFI*B SNPs of interest based on R^2 values in European populations (LDLink). *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs

Table 4.5: Full analysis of SNPs associated with the gene *HNF1B*. Underlined SNPs selected for functional analysis.

Gene	Marker	Chr	Position	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function	Histone Markers Pancreas	eQTL (NES)
<u>HNF1B</u>	rs718961	<u>17</u>	<u>36077099</u>	<u>A</u> <u>G</u>	<u>2b</u>	<u>Enhancer</u>	<u>Intronic</u>	<u>H3Kme1</u> <u>H3K27ac</u>	—
HNF1B	rs12951345#	17	36077863	C A	4	Strong transcription	intronic	H3Kme1 H3K27ac	—
<u>HNF1B</u>	<u>rs4795218*</u>	<u>17</u>	<u>36078510</u>	<u>A</u> <u>G</u>	<u>3a</u>	<u>Enhancers</u>	<u>intronic</u>	<u>H3Kme1</u> <u>H3K27ac</u>	—
HNF1B	rs4794758	17	36080428	T C	6	Enhancers	intronic	H3K27ac	↓ -0.10, Colon
HNF1B	rs7223387#	17	36082473	G T	5	Enhancers	intronic	—	↓ -0.11, Colon

This table includes gene name, SNP marker, HGI9 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, and histone markers. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

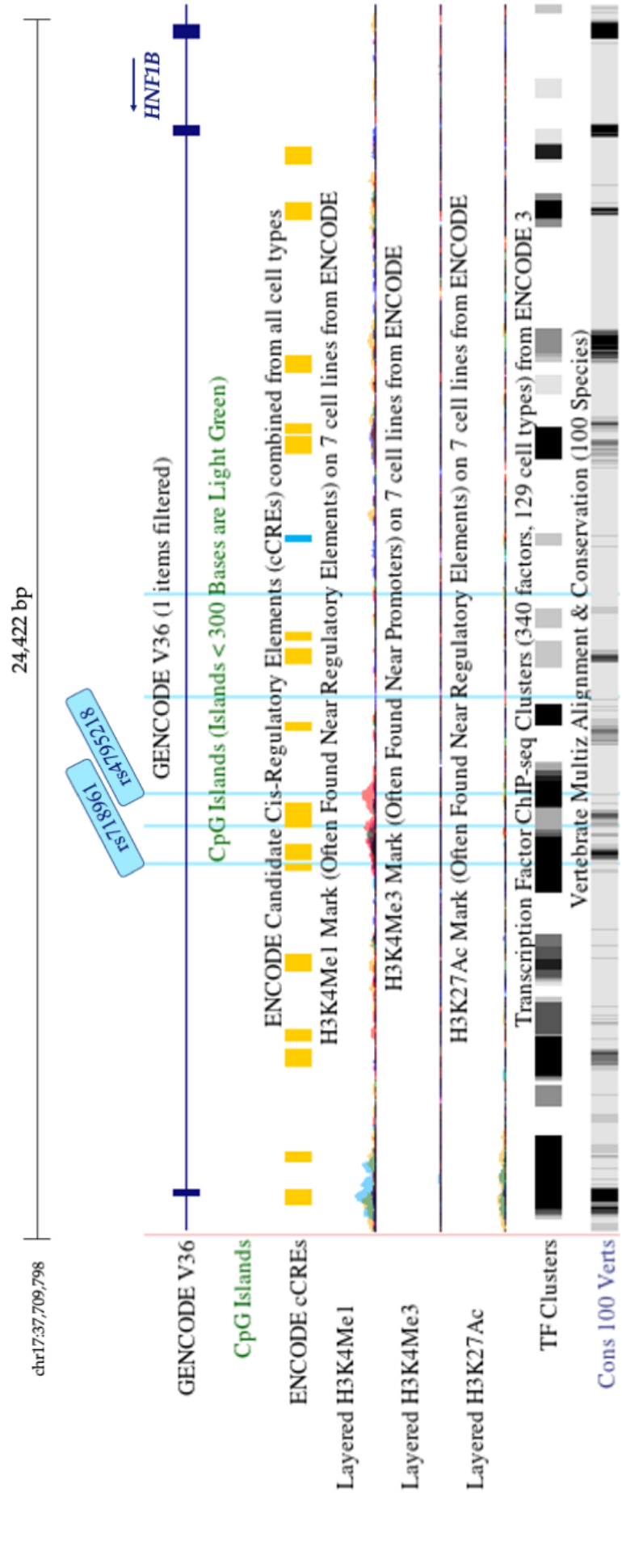


Figure 4.7: UCSC Genome Browser view of HNF1B gene regions surrounding SNPs of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

4.2.1.5. Identification and selection of *HNF4G* SNPs

Hepatocyte Nuclear Factor-4-Homeobox-G (*HNF4G*) is a gene located at 8q21.13 and contains 15 exons. *HNF4G* has high levels of expression in the small intestine, duodenum and colon, where it acts as a transcription factor. The GWAS SNP (rs2941471), two GWAS pathway SNPs (rs1913641 and rs2943547), two SNPs identified through literature searches and the two SNPs in LD ($R^2 > 0.8$; $n=8$ SNPs) are shown in a heatmap (Figure 4.8). The full *in silico* analysis of the SNPs associated with *HNF4G* is outlined in Table 4.6. No SNPs were selected for functional validation. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 4.9.

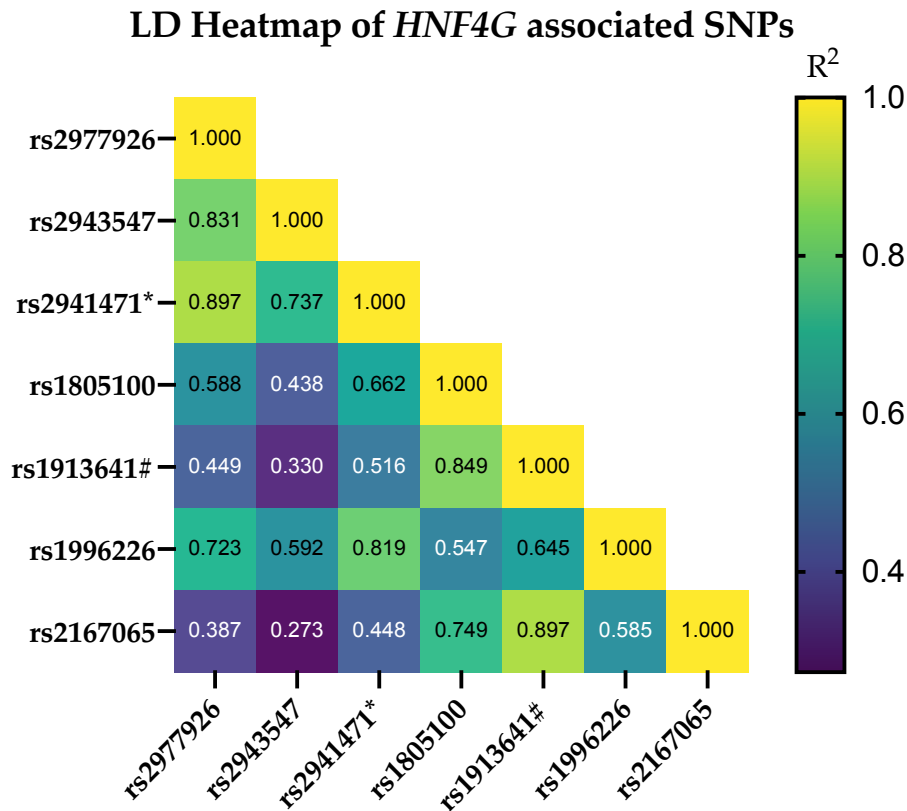


Figure 4.8: LD heatmap of *HNF4G* SNPs of interest based on R^2 values in European populations (LDLink). *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

Table 4.6: Full analysis of SNPs associated with the gene HNF4G.

Gene	Marker	Chr	Position	Alleles Ref Alt	RegDB	Chromatin State	Pancreas	Location/Function	eQTL (NES)
HNF4G	rs2941471*	8	76470404	G A	5	Strong transcription		Intronic	↑ 0.19, Pancreas
HNF4G	rs1805100	8	76476396	G A	4	Strong transcription		3 prime UTR	↓ -0.15, Testis
HNF4G	rs2943547	8	76451098	A G	5	Strong transcription		Intronic	↑ 0.19, Pancreas
HNF4G	rs2977926	8	76448847	G T	6	Quiescent/Low		Intronic	↑ 0.20, Pancreas
HNF4G	rs1913641#	8	76483239	T G	7	Quiescent/Low		—	↓ -0.14, Testis
HNF4G	rs1996226	8	76507184	G T	3b	Low/quiescent		—	↓ -0.14, Testis
HNF4G	rs2167065	8	76560454	C T	3a	Low/quiescent		—	—

This table includes gene name, SNP marker, HGI9 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

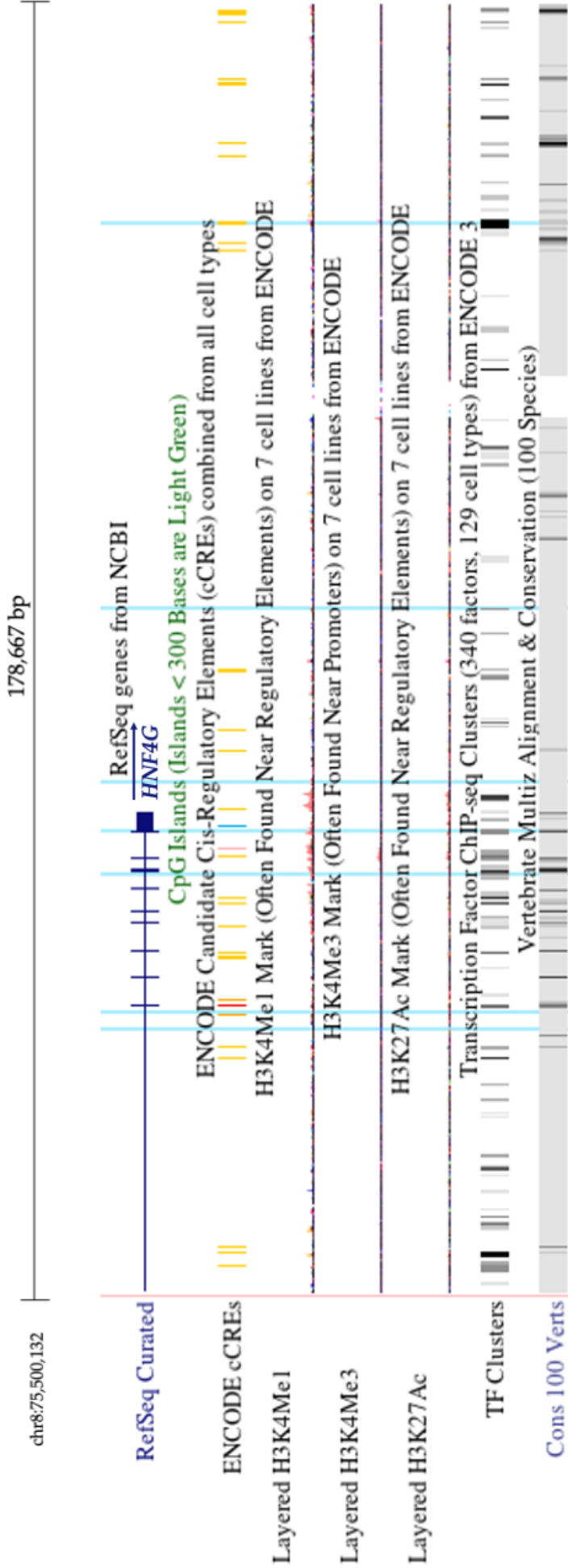


Figure 4.9: UCSC Genome Browser view of HNF4G gene regions surrounding SNPs of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

4.2.1.6. Analysis of NR5A2 associated SNPs

Nuclear Receptor Subfamily 5 Group A Member 2 (*NR5A2*) is a gene located at 1q32.1 and contains 13 exons. *NR5A2* has high levels of expression in the liver, small intestine, and pancreas, where it acts as a DNA-binding zinc finger transcription factor. The GWAS SNP (rs3790844) and six SNPs identified in literature searches (Table 4.2), and the additional two SNPs in LD are shown in a heatmap (Figure 4.10). The full *in silico* analysis of the nine SNPs associated with *NR5A2* is outlined in Table 4.7. SNPs rs3790843 and rs3790844 were selected for functional analysis, both are present in active transcription start sites, and have histone marks in the pancreas, rs3790843 also occurs in an CpG island. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 4.11.

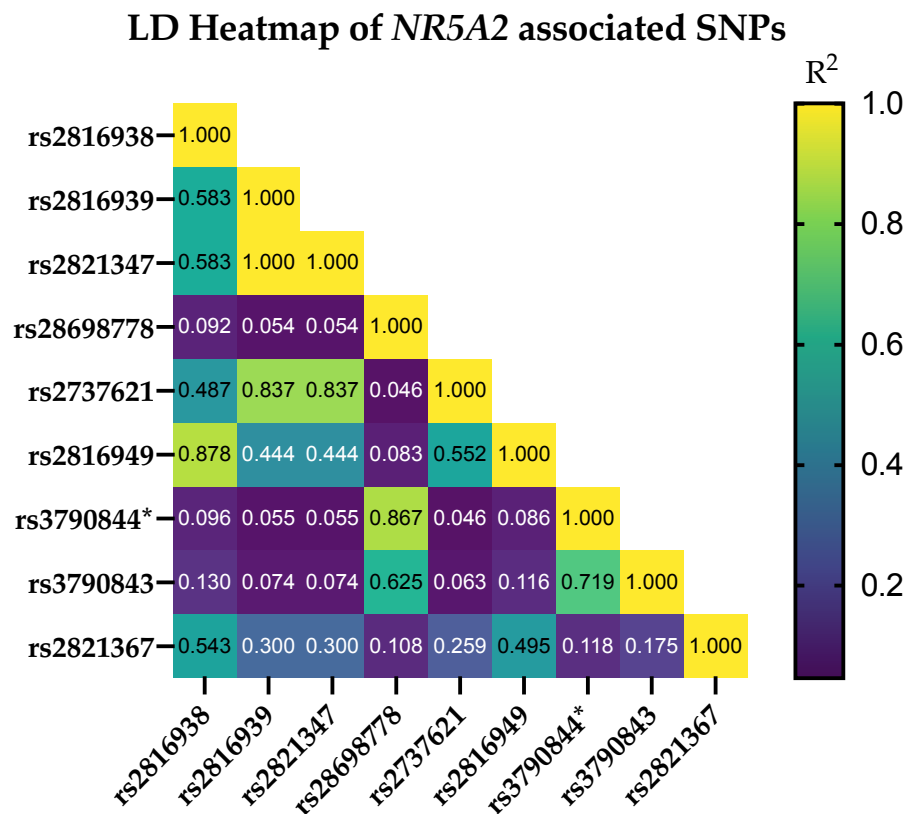


Figure 4.10: LD heatmap of NR5A2 SNPs of interest based on R^2 values in European populations (LDLink). *Highlights GWAS SNPs.

Table 4.7: Full analysis of SNPs associated with the gene NR5A2. Underlined SNPs selected for functional analysis.

Gene	Marker	Chr	Position	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Region	Histone Markers Pancreas	CpG Islands	eQTL (NES)
NR5A2	rs2737621	1	199997403	T C	5	Active TSS	Intronic	—	—	—
<u>NR5A2</u>	<u>rs3790844*</u>	1	<u>200007432</u>	<u>A G</u>	<u>4</u>	<u>Active TSS</u>	<u>Intronic</u>	<u>H3K4me1</u> <u>H3K4me3</u> <u>H3K27ac</u>	—	—
<u>NR5A2</u>	<u>rs3790843</u>	1	<u>200010824</u>	<u>C T</u>	<u>4</u>	<u>Active TSS</u>	<u>Intronic</u>	<u>H3K4me3</u> <u>H3K27ac</u>	<u>Yes</u>	—
NR5A2	rs2816949	1	199997778	A G	3a	Active TSS	Intronic	H3K4me3 H3K27ac	—	—
NR5A2	rs2821367	1	200016146	T C	5	Strong transcription	Intronic	—	—	↑ 0.20, Muscle
NR5A2	rs28698778	1	199990687	T G	2b	Strong transcription	6kb 5' of NR5A2	—	—	—
NR5A2	rs2821347	1	199985792	C T	6	Quiescent/Low	Intronic	—	—	—
NR5A2	rs2816939	1	199985700	T C	5	Quiescent/Low	Intronic	—	—	—
NR5A2	rs2816938	1	199985368	T A	2b	Quiescent/Low	Intergenic	—	—	—

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG island. eQTL includes NES and tissue *Highlights GWAS SNPs.

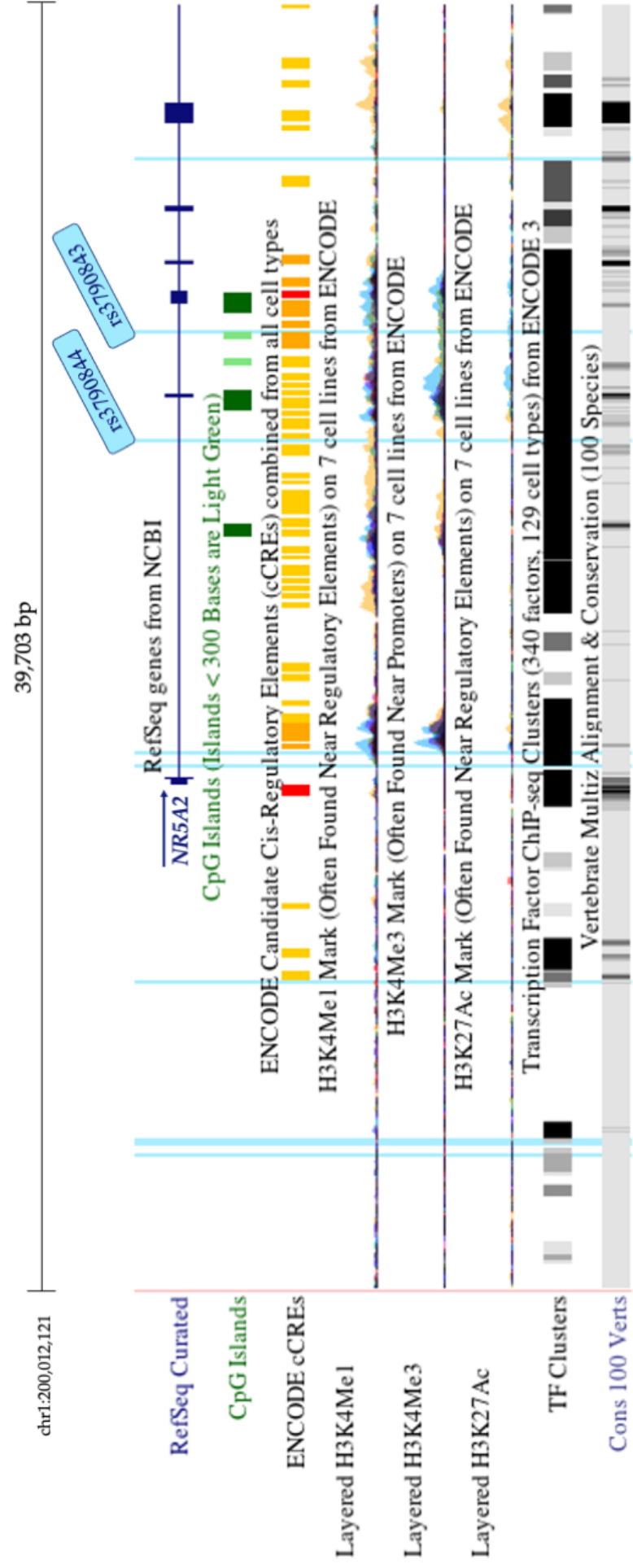


Figure 4.11: UCSC Genome Browser view of NR5A2 gene regions surrounding SNPs of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

4.2.1.7. Identification and selection of *PDXI* SNPs

Pancreatic and duodenal homeobox 1 (*PDXI*) is a gene located at 12q12.2 and contains 2 exons. *PDXI* has high levels of expression in the duodenum, small intestine and pancreas, where it acts as a transcription factor and plays a major role in glucose-dependent regulation of insulin gene expression. Full *in silico* analysis of the GWAS SNP associated with *PDXI* is outlined in Table 4.8, UCSC genome browser view of the region surrounding the SNPs is presented in Figure 4.12, no SNPs in LD were within the selection criteria set out ($R^2 > 0.8$, regulomeDB score ≤ 3).

Table 4.8: Full analysis of SNP located in the PDX1 gene.

Gene	Marker	Chr	Position (HG19)	Alleles Ref/Alt	RegDB	Chromatin State Pancreas	Location/ Function	Histone Markers Pancreas	eQTL (NES)
PDX1	rs9581943*	13	28493997	G A	2b	Strong transcription	2KB Upstream	H3K27ac	↓ -0.27, Pancreas

*This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG island. eQTL includes NES and tissue *Highlights GWAS SNPs.*

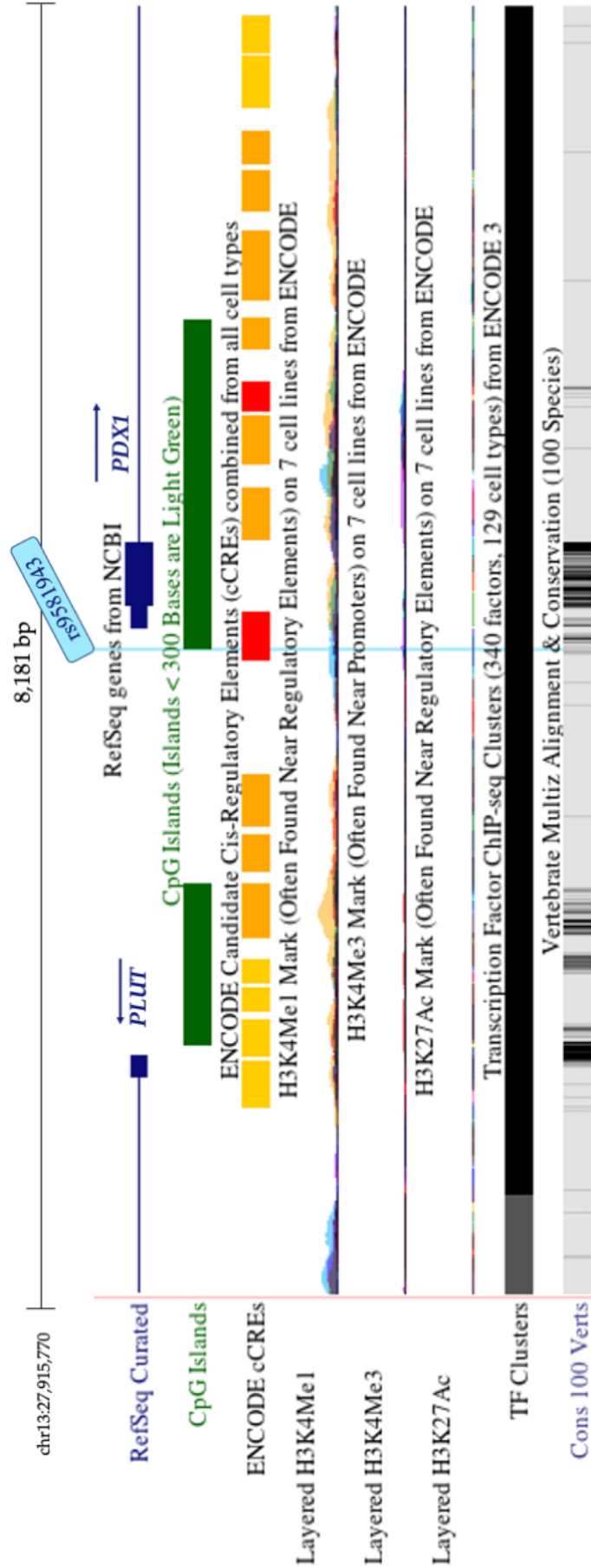


Figure 4.12: UCSC Genome Browser view of PDX1 gene regions surrounding rs9581943 (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factor from ENCODE.

4.2.1.8. SNPs identified in GWAS, GWAS pathway analysis and literature searches, but excluded from further analysis

In silico analysis of SNPs rs1853150, rs483175 and rs118117270 found that all SNPs had a RegulomeDB score of 5 and had a Quiescent/Low chromatin state in the pancreas. None of the SNP were present in CpG islands, or had histone marks, indicating that they were not present in an active regulatory, promoter or enhancer region. No SNPs in LD were within the selection criteria set out ($R^2 < 0.8$, regulomeDB score ≤ 3).

Table 4.9: Full analysis of excluded SNPs located in the HNF4A and PAX4 genes.

Gene	Marker	Chr	Position	Alleles Ref/Alt	RegDB	Chromatin State Pancreas	Location/ Function
HNF4A	rs1853150#	20	43081164	G T	5	Quiescent/Low	Upstream transcript variant
PAX4	rs62483175#	7	127247114	C T	5	Quiescent/Low	N/A
PAX4	rs118117270#	7	127263589	T C	5	Quiescent/Low	N/A

This table includes gene name, SNP marker, HGI9 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP. #Highlights GWAS pathway analysis SNPs.

4.2.1.9. eQTL analysis of SNPs with an effect in the normal pancreas

To further characterise SNP functionality in MODY pathway genes, a cis-eQTL analysis was performed using normal pancreas data from GTEx. Seven SNPs from our analysis were found to act as eQTLs in pancreas tissue (Table 4.10) [347].

Table 4.10: Expression quantitative trait loci (eQTL) for pathway single nucleotide polymorphisms (SNPs) in normal pancreatic tissue from GTEx. Normalised effect size (NES) is the effect of the alternative allele relative to the reference allele.

<i>SNP</i>	<i>Gene</i>	<i>P-value</i>	<i>NES</i>	<i>Organ</i>
rs488087	<i>CELP</i>	1.3E-06	-0.27	Pancreas
rs6489786	<i>SPPL3</i>	5.3E-06	0.22	Pancreas
rs7954039	<i>SPPL3</i>	5.2E-06	0.22	Pancreas
rs7954331	<i>SPPL3</i>	5.2E-06	0.22	Pancreas
rs2941471	<i>HNF4G</i>	9.8E-05	0.19	Pancreas
rs2943547	<i>HNF4G</i>	2.3E-05	0.19	Pancreas
rs9581943	<i>PDX1</i>	2.5E-06	-0.27	Pancreas

Homozygous SNP rs488087 (TT) results in decreased expression of the *CEL* pseudogene *CELP* in the normal pancreas ($P=1.3E-06$), with a normalised effect size (NES) of -0.27 (Figure 4.13). Rs488087 (TT) also acts as an eQTL decreasing the expression of *CELP* in brain, colon and skin-sun exposed tissues; as well as decreasing *CEL* expression in oesophagus – mucosa.

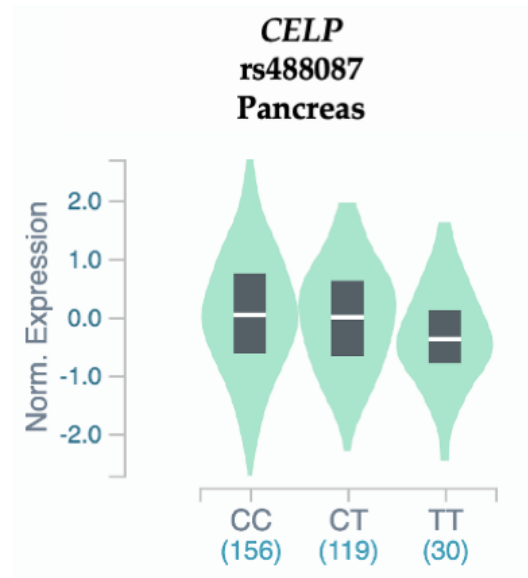


Figure 4.13: eQTL violin plots the effects of homozygous and heterozygous expression of SNPs rs488087 on the gene expression of *CELP* in the pancreas.

GTEx eQTL analysis found the presence of SNPs rs7954331 (TT), rs7954039 (CC) and rs6489786 (GG) results in an increased expression of *SPPL3* in the normal pancreas with an NES of 0.22 (Figure 4.14).

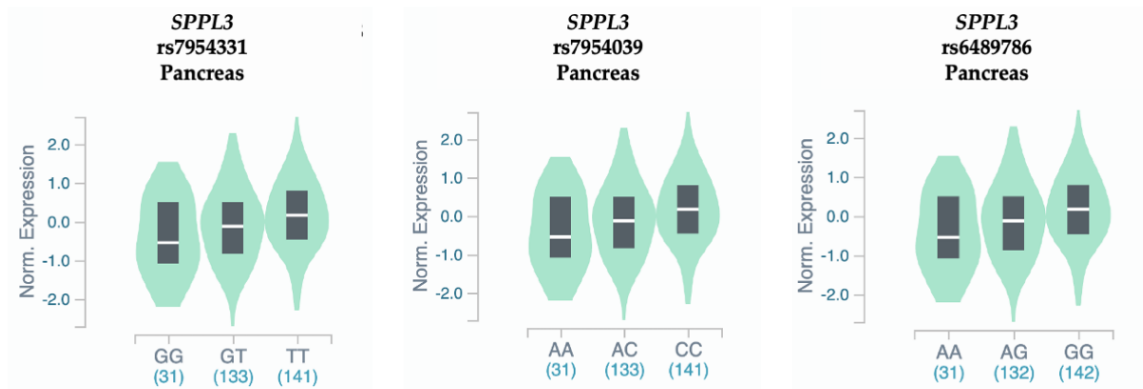


Figure 4.14: eQTL violin plots the effects of homozygous and heterozygous expression of SNPs rs7954331, rs7954039 and rs6489786 on the gene expression of *SPPL3* in the pancreas.

SNPs rs2943547 (GG) and rs2941471 (AA) results in an increased expression of *HNF4G* in the pancreas with an NES of 0.19 (Figure 4.15).

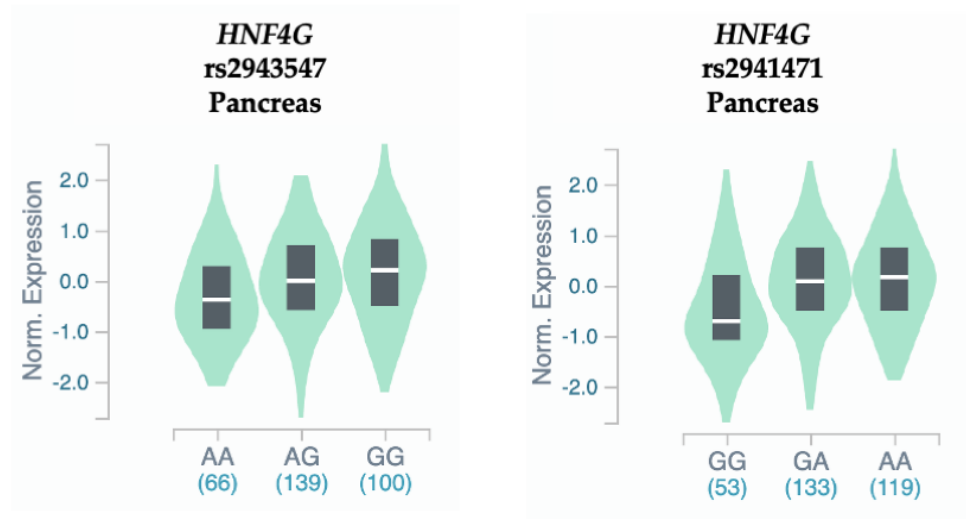


Figure 4.15: eQTL violin plots the effects of homozygous and heterozygous expression of SNPs rs2943547 and rs2941471 on the gene expression of HNF4G in the pancreas.

SNP rs9581943 (AA) acts an eQTL resulting in a decreased expression of PDX1 in normal pancreas with a NES of -0.27 (Figure 4.16).

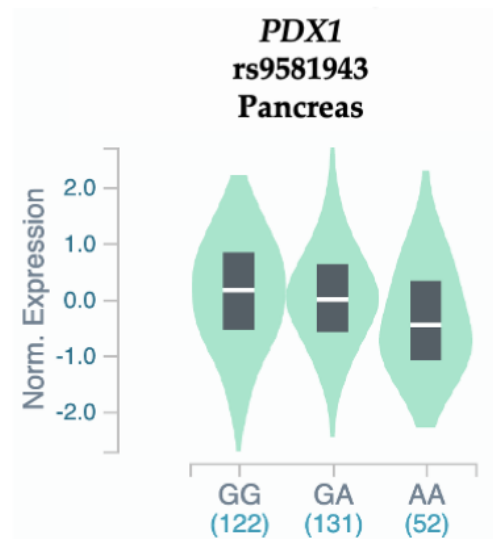


Figure 4.16: eQTL violin plots the effects of homozygous and heterozygous expression of SNPs rs9581943 on the gene expression of PDX1 in the pancreas.

4.2.1.10. SNPs selected for functional validation

In total, six SNPs from four genes were selected for functional validation. Two SNPs from genes *CEL* (rs488087) and *HNFLA* (rs2258287) were successfully cloned and edited using site directed mutagenesis. *HNF1B* SNP (rs4795218) and *NR5A2* SNP (rs3790843) were not brought forward for dual luciferase reporter assays as attempts to optimise PCR cloning of the region of interest were unsuccessful. Optimisation included the use of three different primer pairs, and extension times of 30 seconds, one minute and two minutes. No colonies were observed following three attempts to perform site directed mutagenesis of *HNF1B* SNPs (rs718961) and *NR5A2* SNP (rs3790844) therefore these SNPs not brought forward for dual luciferase reporter assays (Table 4.11).

Table 4.II: Final list of SNPs selected for functional validation, including nearest gene, marker, HGI9 position, reference and alternative SNP and reasons for exclusion.

Gene	Marker	Chr	Position (HGI9)	Ref	Alt	Reason for exclusion
<i>CEL</i>	rs488087	9	135946599	C	T	Included
<i>HNFA</i>	rs2258287	12	121454313	C	A	Included
<i>HNFB</i>	rs718961	17	36077099	A	G	Site directed mutagenesis unsuccessful
<i>HNFB</i>	rs4795218	17	36078510	A	G	Unable to clone region of interest
<i>NR5A2</i>	rs3790844	1	200007432	A	G	Site directed mutagenesis unsuccessful
<i>NR5A2</i>	rs3790843	1	200010824	C	T	Unable to clone region of interest

4.2.1.II. Dual luciferase reports assays

SNPs rs2258287 (*HNFI*A), and rs488087 (*CEL*(P)), were selected for functional validation using a dual luciferase reporter assay was performed, refer to Section 1.9.3 for an outline of dual luciferase reporter assay. The dual luciferase reporter assay was performed using two established PDAC cell lines (ASPC1 and PANC1), and HEK293T cells, as per methodology described in Section 2.13. HOmo sapiens COmprehensive MOdel COllection (HOCOMOCO) v11 is a database of TF binding models for 680 human and 453 mouse TFs [285]. HOCOMOCO-v11 was used to generate position weight matrixes for changes in transcription factor binding due to the presence of the alternative SNP.

Rs2258287 alternative allele (A) resulted in no change in expression in any cell line (Figure 4.17). The presence of the alternative allele (A) is predicted to result in a loss of binding of TFs TYY2 (22.72-fold, $p=7.42E-03$), OVOL2 (19.72-fold, $3.76E-03$) and ZSC22 (10.4-fold, $p=4.34E-03$). Rs2258287-A results in increased binding potential of ZN260 (7.76-fold, $p=5.60E-05$) and RUNX3 (9-fold, $p=4.16E-04$).

Rs488087 alternative allele (T) resulted in increased luciferase expression compared to reference allele in ASPC1 cell line (Figure 4.18) but resulted in a slight decrease in PANC1. The presence of the alternative allele (T) is predicted to result in a gain of binding of ZEB1 and PRGR by 47.5-fold ($P=1.35E-04$) and 14.83-fold ($P=2.08E-04$) respectively. It also is predicted to result in a loss of binding potential of KLF1, EGR4 and GABPA of 23.85-fold ($P=1.48E-03$), 12.31-fold ($P=2.19E-03$), 11.47-fold ($P=5.54E-03$) respectively.

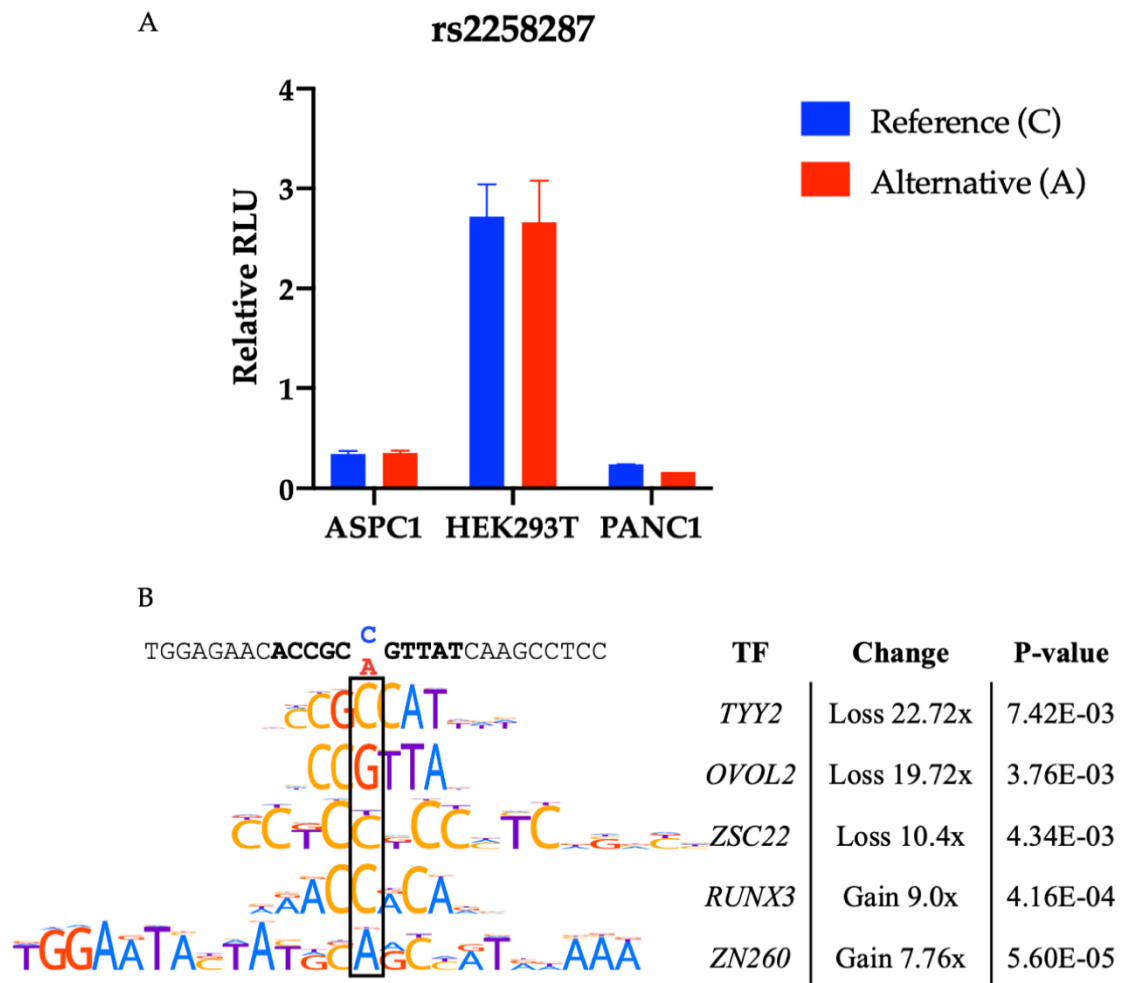


Figure 4.17: (A) Effects of the candidate functional SNP rs2258287 on gene expression in a dual luciferase reporter assay. Two allelic variants of 5'UTR containing each allele variant were cloned upstream of the Firefly luciferase gene. A panel of cell lines – HEK293T, ASPC1 and PANC1 – were transfected with reporter constructs together with Renilla. Luminescence was measured 24 hours after transfection using Dual-Glo Luciferase Assay System (Promega). Luciferase expression was normalised to Renilla and made relative to p2Luc plasmid with no insertion. Blue boxes – reference allele, red boxes – alternative allele. (B) DNA sequence surrounding rs2258287 and ChIP-seq based motif logos for TTY2, OVOL2, ZSC22, RUNX3 and ZN260 according to HOCOMOCO-v11 [285]. C to A substitution alters motif binding (fold-change is given according to PERFECTOS-APE [286]).

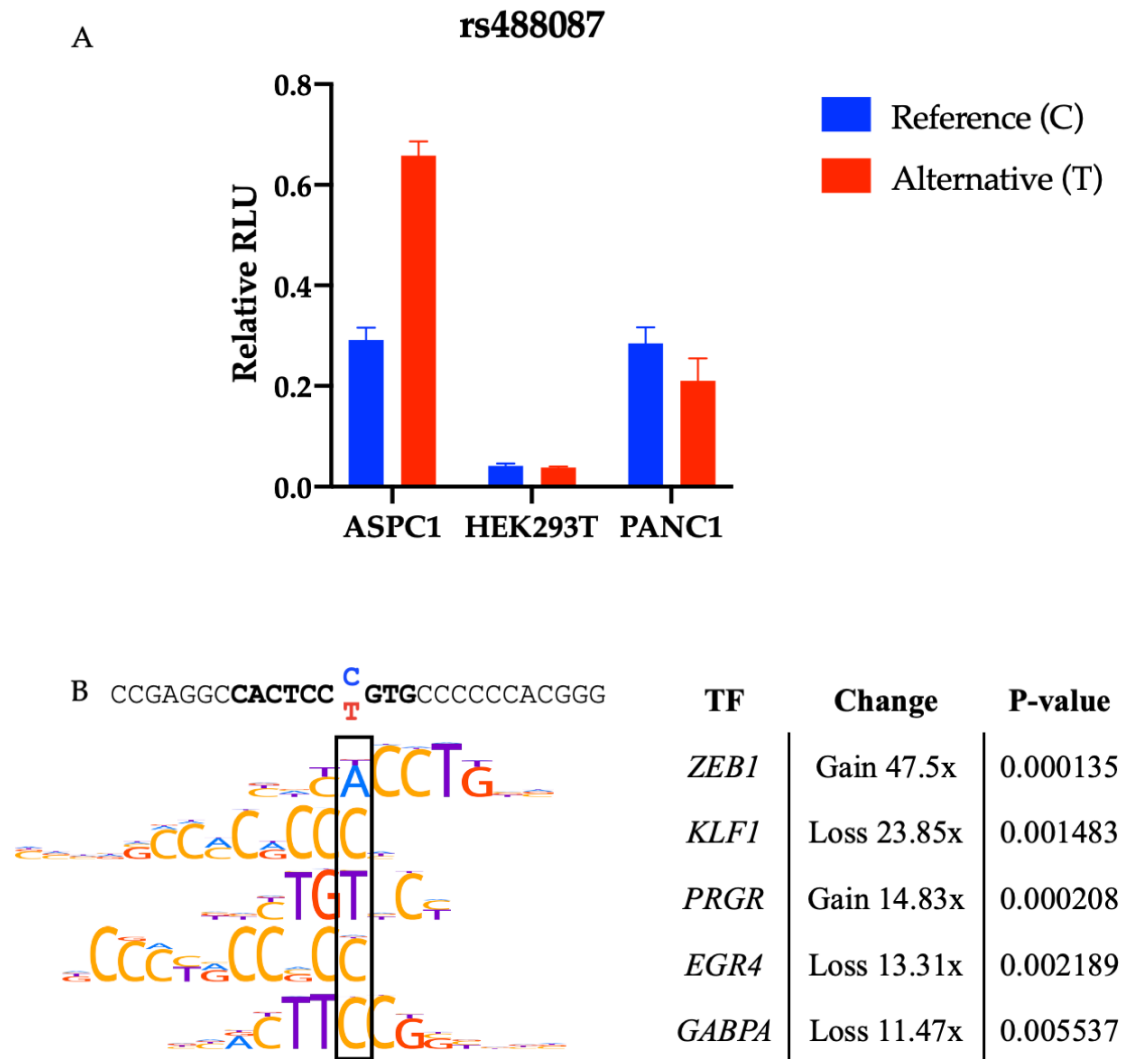


Figure 4.18: Effects of the candidate function SNPs on gene expression in a dual luciferase reporter assay. Two allelic variants of 5' UTR containing each allele variant were cloned upstream of the Firefly luciferase gene. A panel of cell lines – HEK293T, ASPC1 and PANC1 – were transfected with reporter constructs together with Renilla. Luminescence was measured 24 hours after transfection using Dual-Glo Luciferase Assay System (Promega). Luciferase expression was normalised to Renilla and made relative to p2Luc plasmid with no insertion. Blue boxes – protective allele, red boxes – potential risk allele. DNA sequence surrounding rs488087 and ChIP-seq based motif logos for ZEB1, KLF1, PRGR, EGR4 and GABPA according to HOCOMOCO-vII [285]. C to T substitution alters motif binding (fold-change is given according to PERFECTOS-APE [286]).

4.2.2. Role of genes in the MODY pathway in the development and progression of PDAC

The role of MODY genes in PDAC was assessed using *in silico* analysis of genes using publicly available datasets from Gene Expression Profiling Interactive Analysis 2 (GEPIA2) and RNA-sequencing data from PDAC organoids and normal samples from a study by Tiriach *et al.* [236] (dbGAP: phs001611).

GEPIA2 is an enhanced web server for large-scale expression profiling and interactive analysis using datasets from TCGA and GTEx [348]. GEPIA2 was used to assess differences in expression of genes in PDAC versus normal pancreas samples, and to perform Kaplan-Meier survival analysis on patients with high ($\geq 80^{\text{th}}$ percentile) and low ($\leq 20^{\text{th}}$ percentile) gene expression. In addition, PCC analysis was performed to identify genes with a similar expression pattern of the gene of interest in PDAC and normal pancreas, to identify aberrant gene expression patterns.

Data from a study by Tiriach *et al.* [236] was used in RNA-sequencing data analysis. In the study, the group established a patient derived organoid library from 159 PDAC tumour and metastases in 138 patients. Organoid cultures were successfully established for 75% of samples taken. In addition, 11 normal pancreatic ductal organoid cultures were established from healthy normal pancreases. RNA-sequencing was performed on 44 PDAC-confirmed and 11 normal pancreas organoid samples.

In RNA-sequencing data analysis, patient samples were manually divided into groups of high and low expression of the gene of interest using normalised gene counts. RNA-seq data was analysed looking at differences in expression of downstream genes controlled by transcription factors in the MODY pathway, with lists obtained from <https://maayanlab.cloud/Harmonizome/> (Appendix C). Using this data, PCA of normal pancreas organoid samples and organoid samples with high and low expression of the gene of interest was performed. Volcano plots of differentially expressed genes in the samples with high and low expression of the

gene of interest were generated, and significantly upregulated and downregulated pathways were also identified. As the literature has shown that the combination of a loss of *GATA6*, *HNFLA* and *HNF4A* results in a phenotypic switch to the aggressive basal subtype of the disease, RNA-seq data analysis was also performed on downstream genes of *GATA6* and *HNF4A*.

Genes were also assessed using experimental functional techniques such as CRISPR knockouts of genes *HNFLA*, *HNF1B*, *HNF4G* and *NR5A2*, and CUT&RUN to identify putative binding sites of *HNFLA* and *HNF1B*.

4.2.2.1. Assessment of the role and expression of *HNFLA*

The overall survival of PDAC patients with low *HNFLA* expression (cut-off $\leq 20\%$) was significantly increased ($P=0.02$) compared to patients with high *HNFLA* expression (cut-off $\geq 80\%$). The expression of *HNFLA* is increased in PDAC tumour samples compared to normal pancreas. PCC analysis was performed to identify genes with a similar expression pattern in PDAC and normal pancreas, and *FOXP4* was the only gene occurring in both tumour and normal samples (Figure 4.19).

RNA-seq analysis of genes controlled by transcription factor *HNFLA* (1131 genes) found samples with high *HNFLA* expression clustered together and clustered separately to samples with low *HNFLA* expression (Figure 4.20 A). The most upregulated gene in samples with low *HNFLA* expression was *RORB* (log2 fold-change 6.79; p-adj=6.7E-05). The most significantly downregulated gene was *CDH17* (log2 fold-change -7.93; p-adj=2.22E-15) (Figure 4.20 B).

Differentially expressed genes (DEG) were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in *HNFLA*-low and *HNFLA*-high samples (Figure 4.21). Thirteen pathways were found to be downregulated in *HNFLA*-low samples, with cell differentiation 31 genes in the pathway the most significant (p-adj=1.0E-03). Fifteen pathways were upregulated in *HNFLA*-low samples. The most significantly upregulated pathway was anatomical structure morphogenesis, with 32 genes (p-adj=3.40E-08) (Table 4.12).

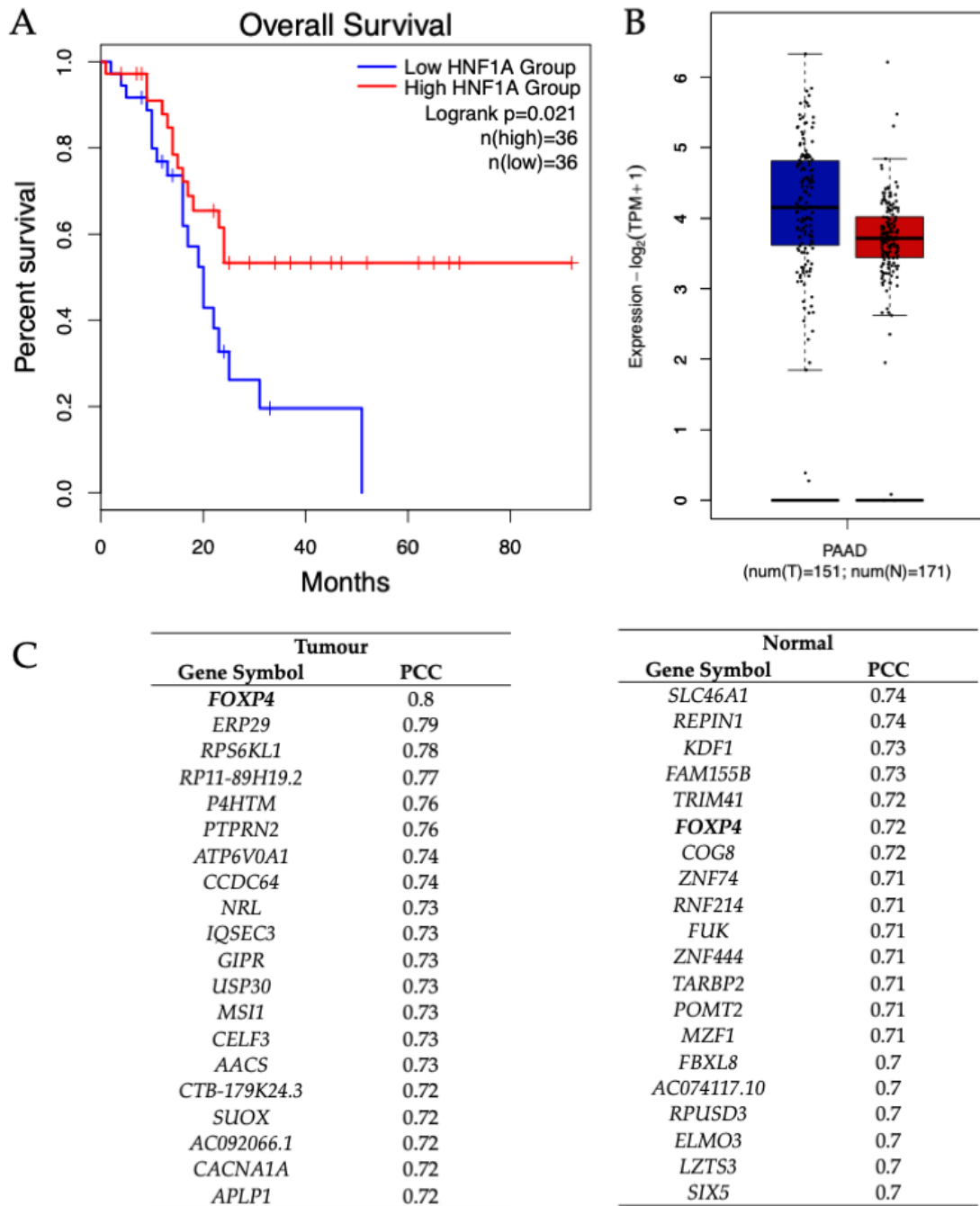


Figure 4.19: Analysis of HNF1A using TCGA and GTEx datasets on GEPIA2, including (A) survival analysis of overall survival of pancreatic adenocarcinoma (PAAD) patients with high (≥ 80 th percentile) (red) and low (≤ 20 th percentile) (blue) HNF1A expression in PDAC. (B) Comparison of HNF1A mRNA expression in tumour (blue) and normal (red) samples. (C) Pearson correlation coefficient (PCC) analysis identified genes with a similar expression pattern to HNF1A in PDAC and normal pancreas.

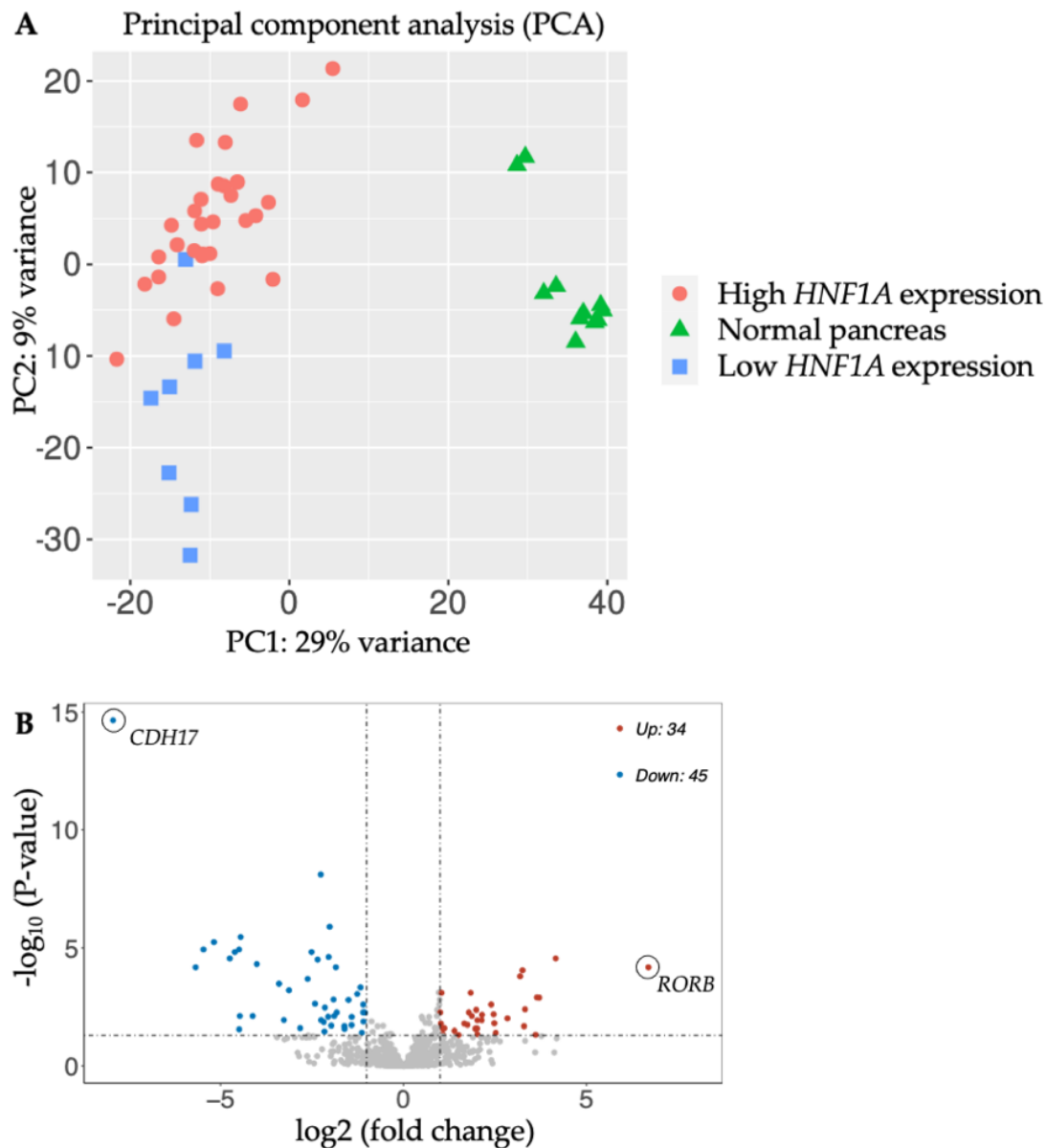


Figure 4.20: RNA-seq analysis of samples with high and low *HNF1A* expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high expression of *HNF1A* mRNA (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low *HNF1A* vs. high *HNF1A* mRNA expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ P-value from logistic regression analysis.

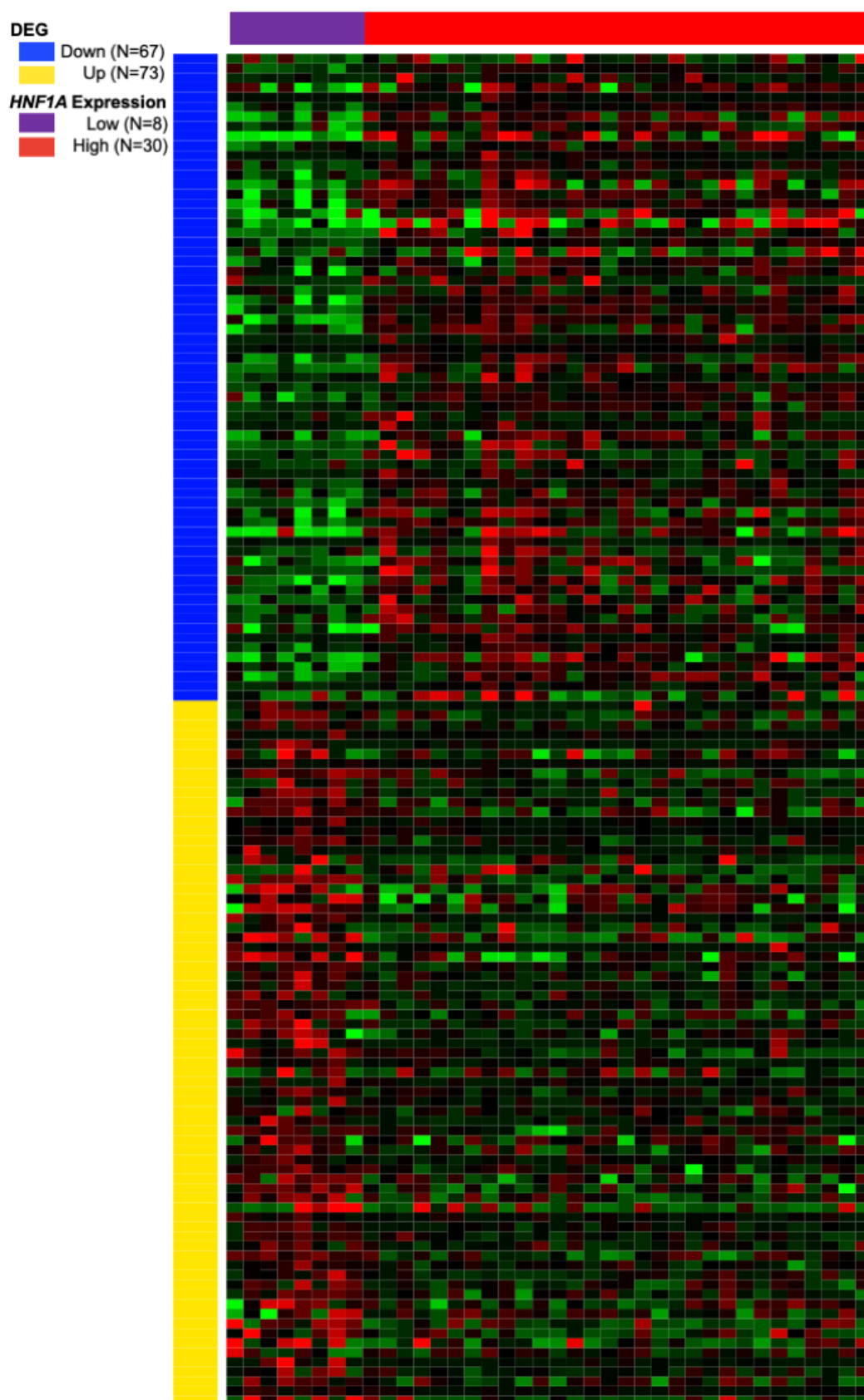


Figure 4.21: A heatmap shows upregulated (blue) and downregulated (yellow) differentially expressed genes (DEG) in HNF1A-low (purple) and HNF1A-high (red) samples.

Table 4.12: Differentially expressed genes were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in samples with low and high HNF1A mRNA expression. The table lists the upregulated and downregulated pathways, including adjusted P-value (p-adj), number of genes, and pathway name.

<i>P-adj</i>	<i>Number of Genes</i>	<i>Pathway</i>
Downregulated		
1.0E-03	31	Cell differentiation
1.2E-03	7	Epithelial cell development
1.2E-03	20	Cell development
2.4E-03	22	Anatomical structure morphogenesis
3.0E-03	4	Glandular epithelial cell differentiation
3.3E-03	7	Neuron projection guidance
6.8E-03	3	Glandular epithelial cell development
6.8E-03	3	Type B pancreatic cell differentiation
6.8E-03	18	Movement of cell or subcellular component
6.8E-03	5	Digestion
7.9E-03	3	Sperm capacitation
8.0E-03	3	Epithelial cell morphogenesis
9.9E-03	9	Chemotaxis
Upregulated		
3.4E-08	32	Anatomical structure morphogenesis
8.7E-08	20	Animal organ morphogenesis
8.7E-08	31	Regulation of developmental process
1.7E-07	35	Animal organ development
1.7E-07	22	Positive regulation of developmental process
1.7E-07	25	Positive regulation of multicellular organismal process
2.8E-07	25	Regulation of cell differentiation
4.5E-06	26	Nervous system development
4.5E-06	31	Regulation of multicellular organismal process
4.5E-06	13	Sensory organ development
4.5E-06	6	Glomerulus development
4.7E-06	16	Regulation of nervous system development
4.7E-06	24	Regulation of multicellular organismal development
7.6E-06	35	Cell differentiation
7.6E-06	13	Developmental growth

4.2.2.2. Assessment of the role and expression of *HNFB*

There was no significant difference in overall survival of PDAC patients with high ($\geq 80^{\text{th}}$ percentile) and low ($\leq 20^{\text{th}}$ percentile) *HNFB* mRNA expression. The expression of *HNFB* is increased in PDAC tumour samples compared to normal pancreas. PCC analysis was performed to identify genes with a similar expression pattern in PDAC and normal pancreas. Both tumour and normal had *ENRICH5* and *ANXA4* with the highest PCC (Figure 4.22).

RNA-seq analysis of genes controlled by transcription factor *HNFB* (451 genes) found no clustering of samples with high *HNFB* mRNA expression compared to samples with low *HNFB* mRNA expression (Figure 4.23). The most significantly upregulated gene in samples with low *HNFB* mRNA expression was *MAGIX* (log2 fold-change 2.40; p-adj=8.34E-05). The most significantly downregulated gene was *SLC01B1* (log2 fold-change -5.90; p-adj=1.58E-02). No differentially pathways expressed were identified between the two groups.

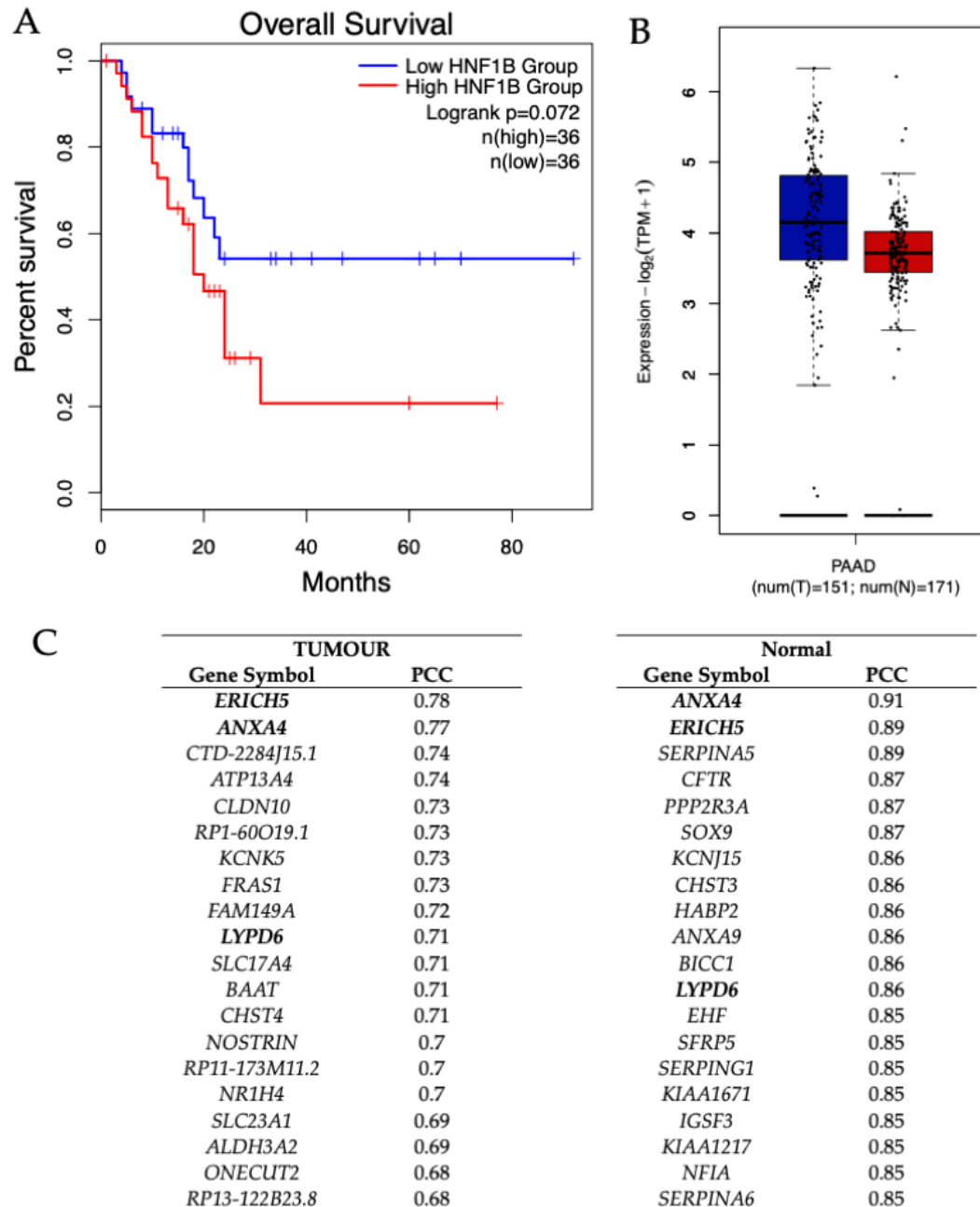


Figure 4.22: Analysis of HNF1B using TCGA and GTEx datasets on GEPIA2 including (A) survival analysis of overall survival of pancreatic adenocarcinoma patients with high ($\geq 80^{\text{th}}$ percentile) (red) and low ($\leq 20^{\text{th}}$ percentile) (blue) HNF1B expression in PDAC (B) Comparison of HNF1B mRNA expression in tumour (blue) and normal (red) samples. (C) Pearson correlation coefficient (PCC) analysis identified genes with a similar expression pattern to HNF1B in PDAC and normal pancreas. Genes present in both groups are highlighted in bold.

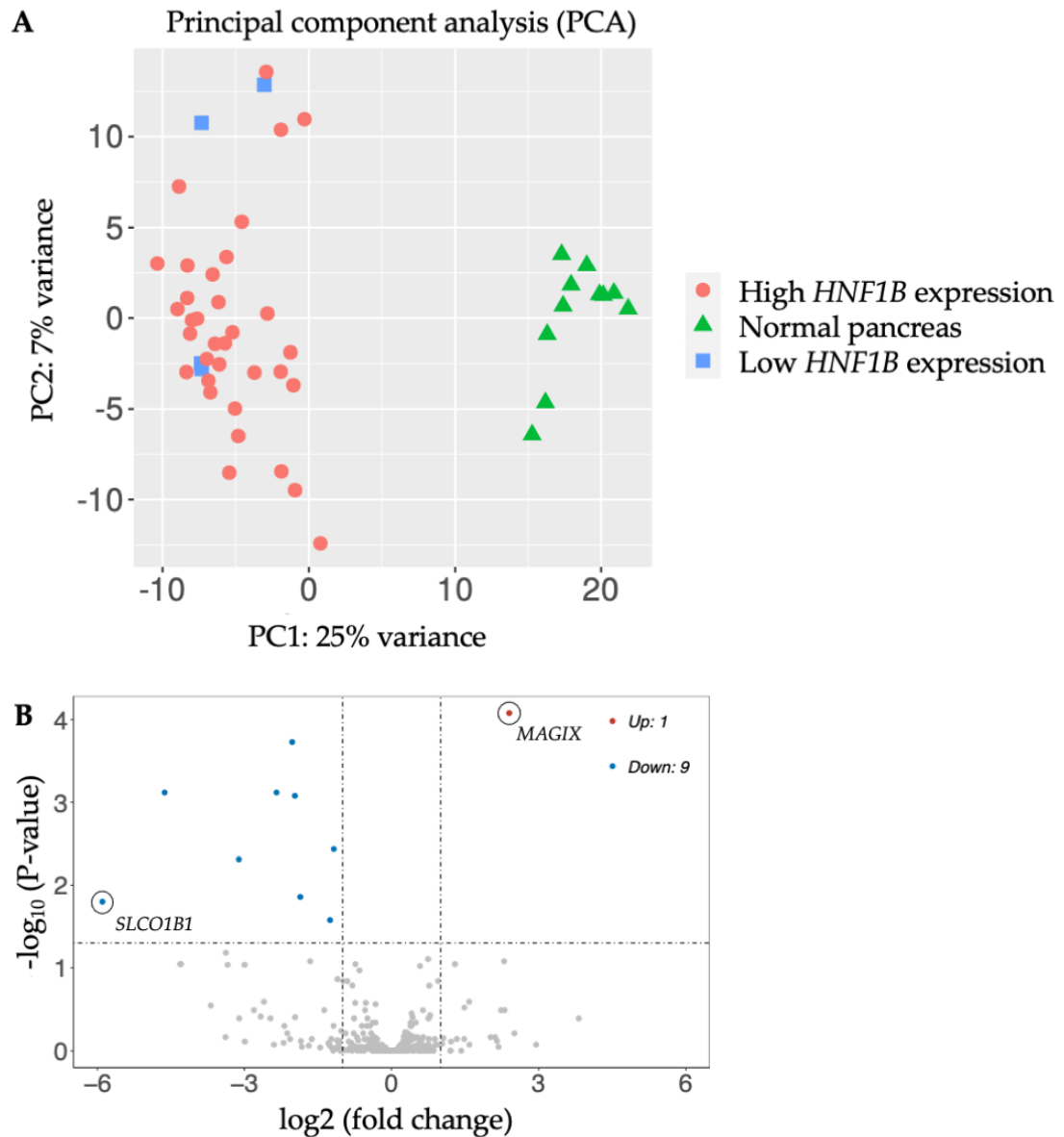


Figure 4.23: RNA-seq analysis of samples with high and low *HNF1B* expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high *HNF1B* mRNA expression. (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low vs. high *HNF1B* mRNA expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ P-value from logistic regression analysis.

4.2.2.3. Assessment of the role and expression of *HNF4G*

The overall survival of PDAC patients with low (≤ 20 th percentile) *HNF4G* expression was significantly longer ($P=0.04$) than patients with high (≥ 80 th percentile) *HNF4G* expression. The expression of *HNF4G* is significantly increased in PDAC tumour samples compared to normal pancreas. PCC analysis was performed to identify genes with a similar expression pattern in PDAC and normal pancreas, *RBM47* and *ATP8B1* were shown to have similar expression pattern to *HNF4G* tumour and normal samples (Figure 4.24).

RNA-seq analysis of genes controlled by transcription factor *HNF4G* (4121 genes) found samples with high *HNF4G* mRNA expression clustered together and compared to samples with low *HNF4G* mRNA expression clustered separately (Figure 4.25 A). The most upregulated gene in samples with low *HNF4G* expression was *GNG4* (\log_2 fold-change 3.62; $p\text{-adj}=2.47\text{E-}02$). The most significantly downregulated gene was *HGD* (\log_2 fold-change -6.51; $p\text{-adj}=4.00\text{E-}22$), MODY genes *HNFLA*, *HNF4A* and *HNFLA-AS1* were also downregulated with a \log_2 fold-change -3.87, -3.89 and -8.63 and $p\text{-adj}$ of $3.90\text{E-}17$, $2.76\text{E-}16$ and $2.76\text{E-}11$, respectively.

DEG were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in *HNF4G*-low and *HNF4G*-high samples (Figure 4.26). Thirteen pathways were found to be downregulated in *HNF4G*-low samples, with the most significant being organic hydroxy compound metabolic process, with 20 genes and a P -value of $1.3\text{E-}06$. Twelve pathways were upregulated in *HNF4G*-low samples. The most significantly upregulated pathway was positive regulation of nitrogen compound metabolic process, with 31 genes ($p\text{-adj}=1.70\text{E-}04$) (Table 4.13).

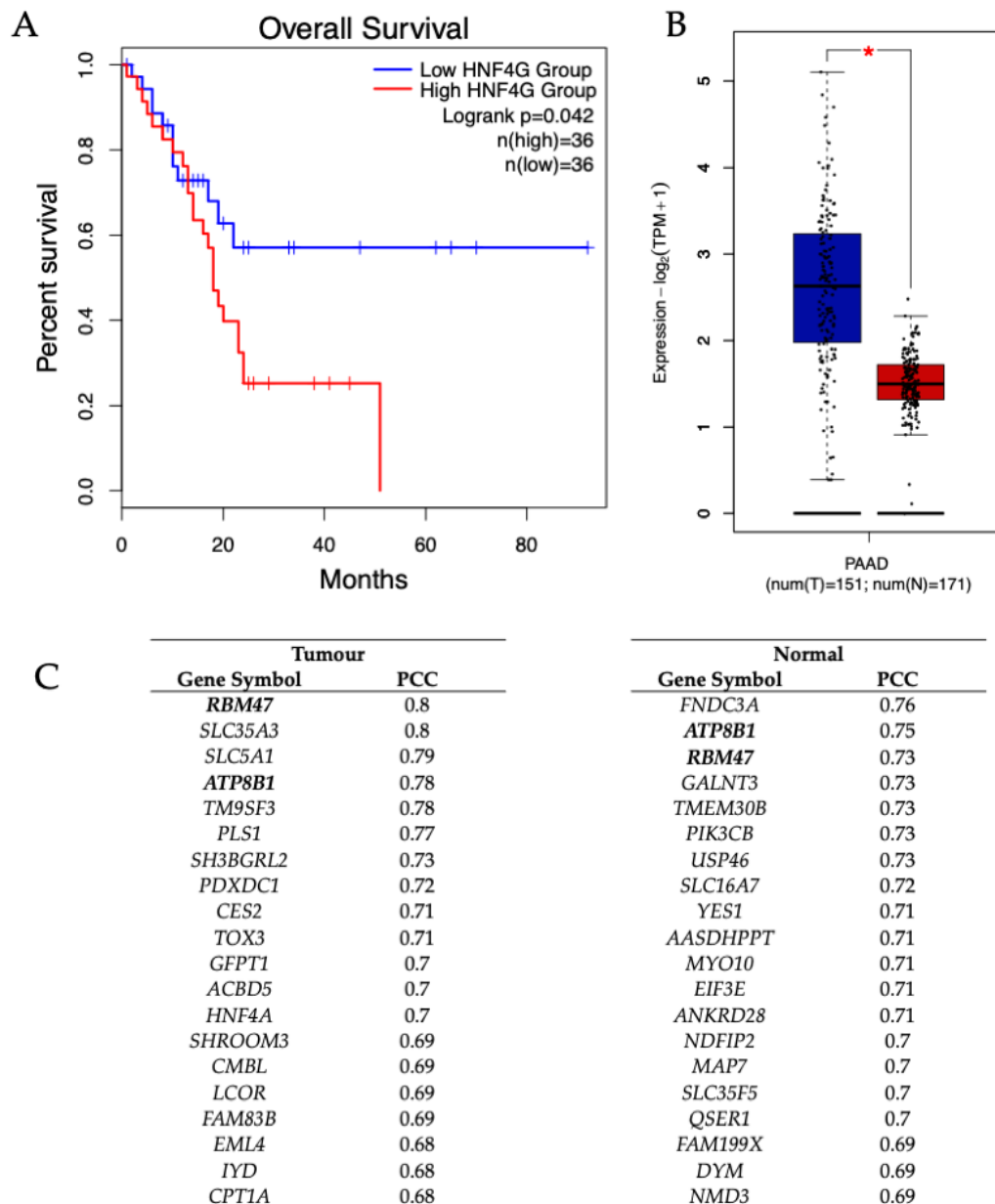


Figure 4.24: Analysis of HNF4G using TCGA and GTEx datasets on GEPIA2 including (A) survival analysis of overall survival of pancreatic adenocarcinoma (PAAD) patients with high (≥ 80 th percentile) (red) and low (≤ 20 th percentile) (blue) HNF4G expression in PDAC. (B) Comparison of HNF4G mRNA expression in tumour (blue) and normal (red) samples. (C) Pearson correlation coefficient (PCC) analysis identified genes with a similar expression pattern to HNF4G in PDAC and normal pancreas.

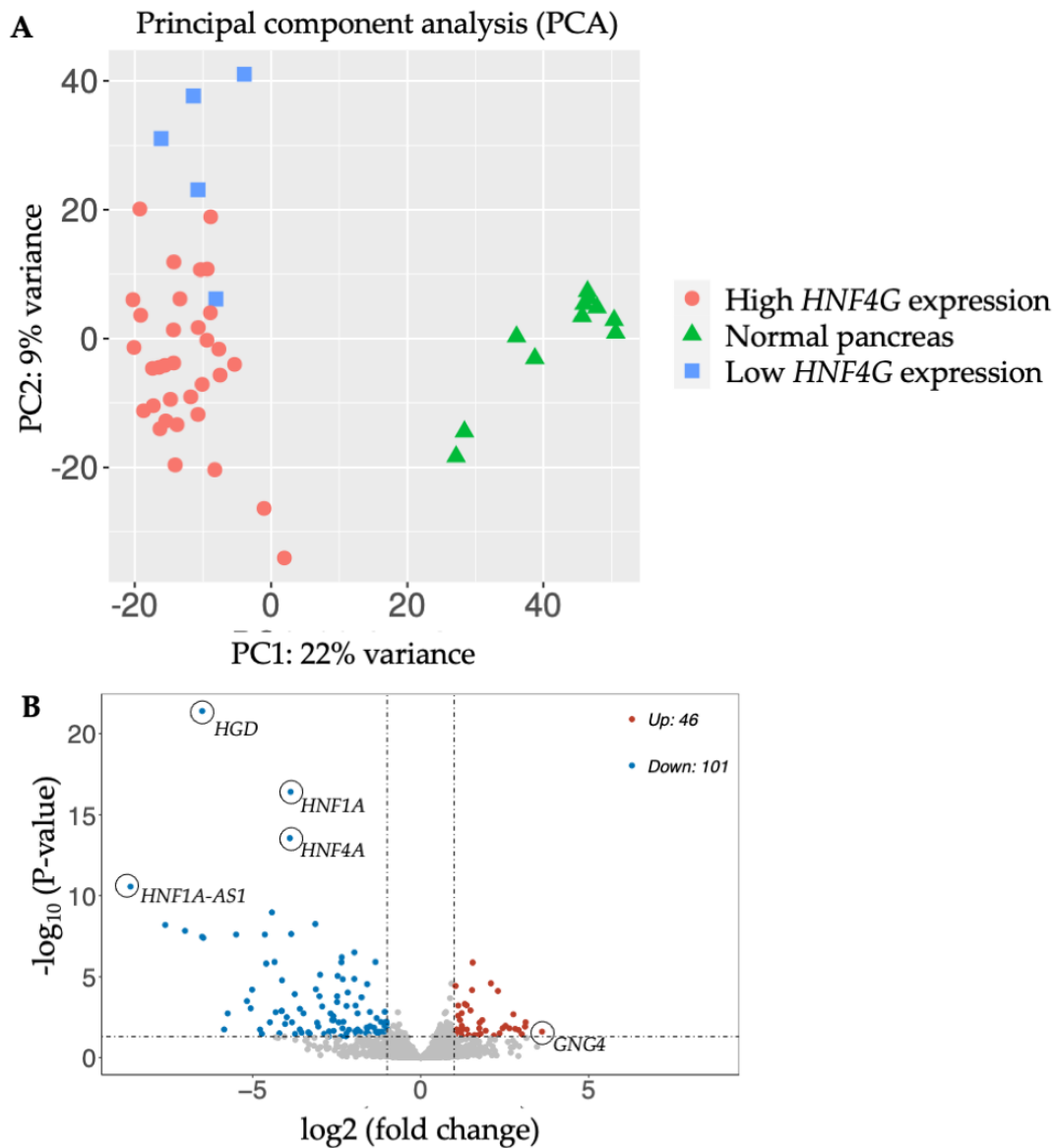


Figure 4.25: RNA-seq analysis of samples with high and low *HNF4G* expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high *HNF4G* mRNA expression. (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low vs. high *HNF1A* expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ P-value from logistic regression analysis.

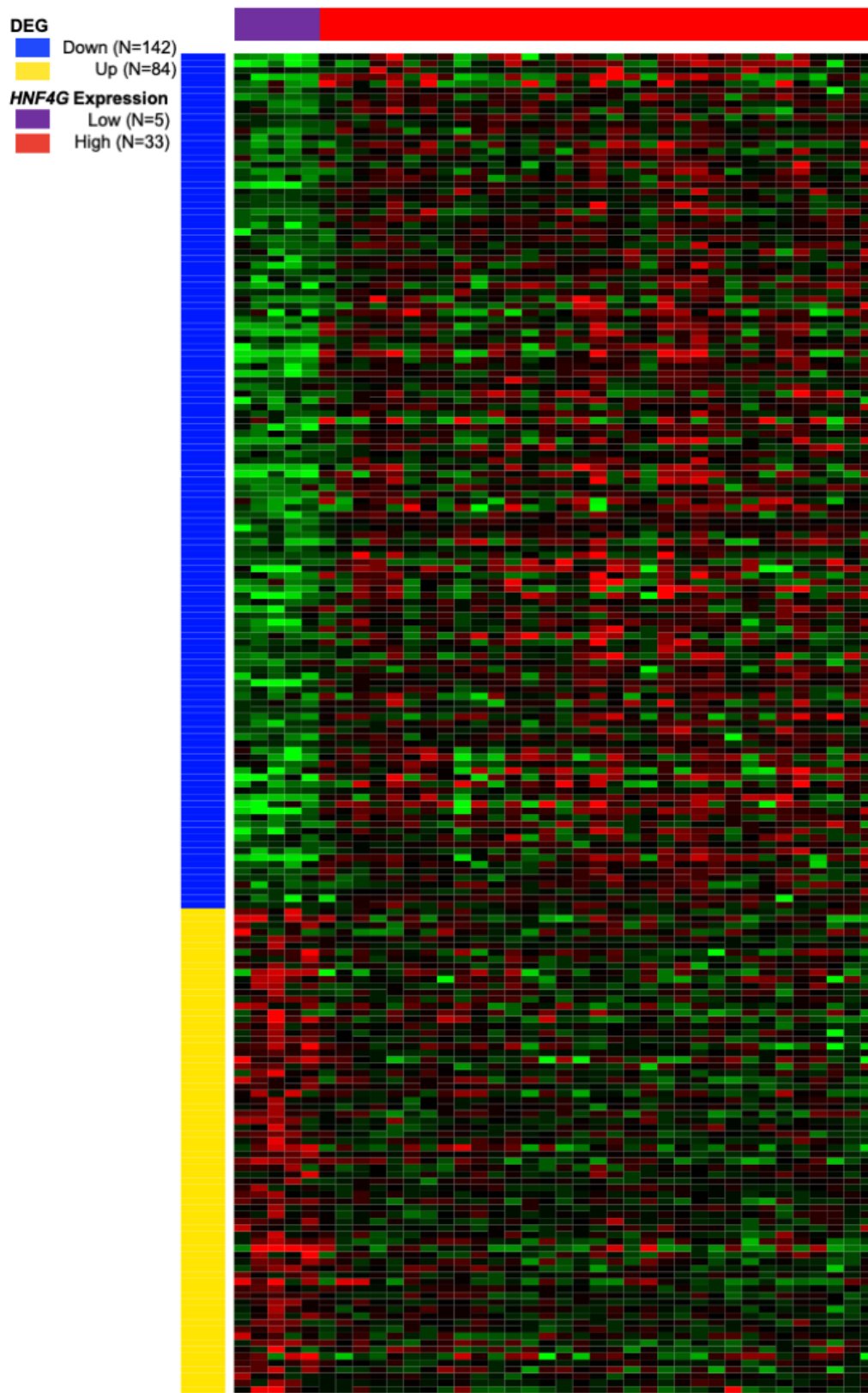


Figure 4.26: A heatmap shows upregulated (blue) and downregulated (yellow) differentially expressed genes (DEG) in HNF4G-low (purple) and HNF4G-high (red) samples.

Table 4.13: Differentially expressed genes were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in samples with low and high HNF4G mRNA expression. The table lists the upregulated and downregulated pathways, including adjusted P-value (p-adj), number of genes, and pathway name.

<i>P-adj</i>	<i>Number of Genes</i>	<i>Pathways</i>
Downregulated		
1.3E-06	20	Organic hydroxy compound metabolic process
3.8E-06	31	Lipid metabolic process
8.0E-06	18	Regulation of hormone levels
9.8E-06	15	Alcohol metabolic process
9.8E-06	26	Organic acid metabolic process
9.8E-06	25	Carboxylic acid metabolic process
9.8E-06	19	Monocarboxylic acid metabolic process
9.8E-06	26	Oxoacid metabolic process
9.8E-06	35	Small molecule metabolic process
1.1E-05	25	Cellular lipid metabolic process
1.1E-05	6	Acylglycerol homeostasis
3.8E-05	13	Steroid metabolic process
4.3E-05	9	Neutral lipid metabolic process
Upregulated		
1.7E-04	31	Positive regulation of nitrogen compound metabolic process
8.3E-04	34	Cell differentiation
8.3E-04	21	Regulation of cell differentiation
8.3E-04	26	Regulation of developmental process
8.3E-04	15	Negative regulation of developmental process
9.1E-04	29	Positive regulation of cellular metabolic process
1.0E-03	30	Positive regulation of macromolecule metabolic process
1.1E-03	31	Positive regulation of metabolic process
1.1E-03	7	Regulation of endocytosis
1.1E-03	5	Regulation of blood coagulation
1.1E-03	19	Positive regulation of protein metabolic process
1.6E-03	24	Anatomical structure morphogenesis

4.2.2.4. Assessment of the role and expression of NR5A2

The expression of *NR5A2* is significantly decreased in PDAC tumour samples compared to normal pancreas. The overall survival of PDAC patients with high ($\geq 80^{\text{th}}$ percentile) *NR5A2* expression compared to low ($\leq 20^{\text{th}}$ percentile) *NR5A2* expression showed no significant difference in patients. PCC analysis was performed to identify genes with a similar expression pattern in PDAC and normal pancreas, and no overlapping genes were identified (Figure 4.27).

RNA-seq analysis of genes controlled by transcription factor *NR5A2* (4968 genes) found no clustering of samples with high *NR5A2* expression compared to samples with low *NR5A2* expression (Figure 4.28). The most significantly upregulated gene in samples with low *NR5A2* expression was *FSTL4* (log2 fold-change 4.71; p-adj=3.85E-06). The most significantly downregulated genes were *CRISP2* (log2 fold-change -4.73; p-adj=1.51E-02) and *NRG3* (log2 fold-change -4.70; p-adj=1.51E-02). No differentially expressed pathways were identified between the two groups.

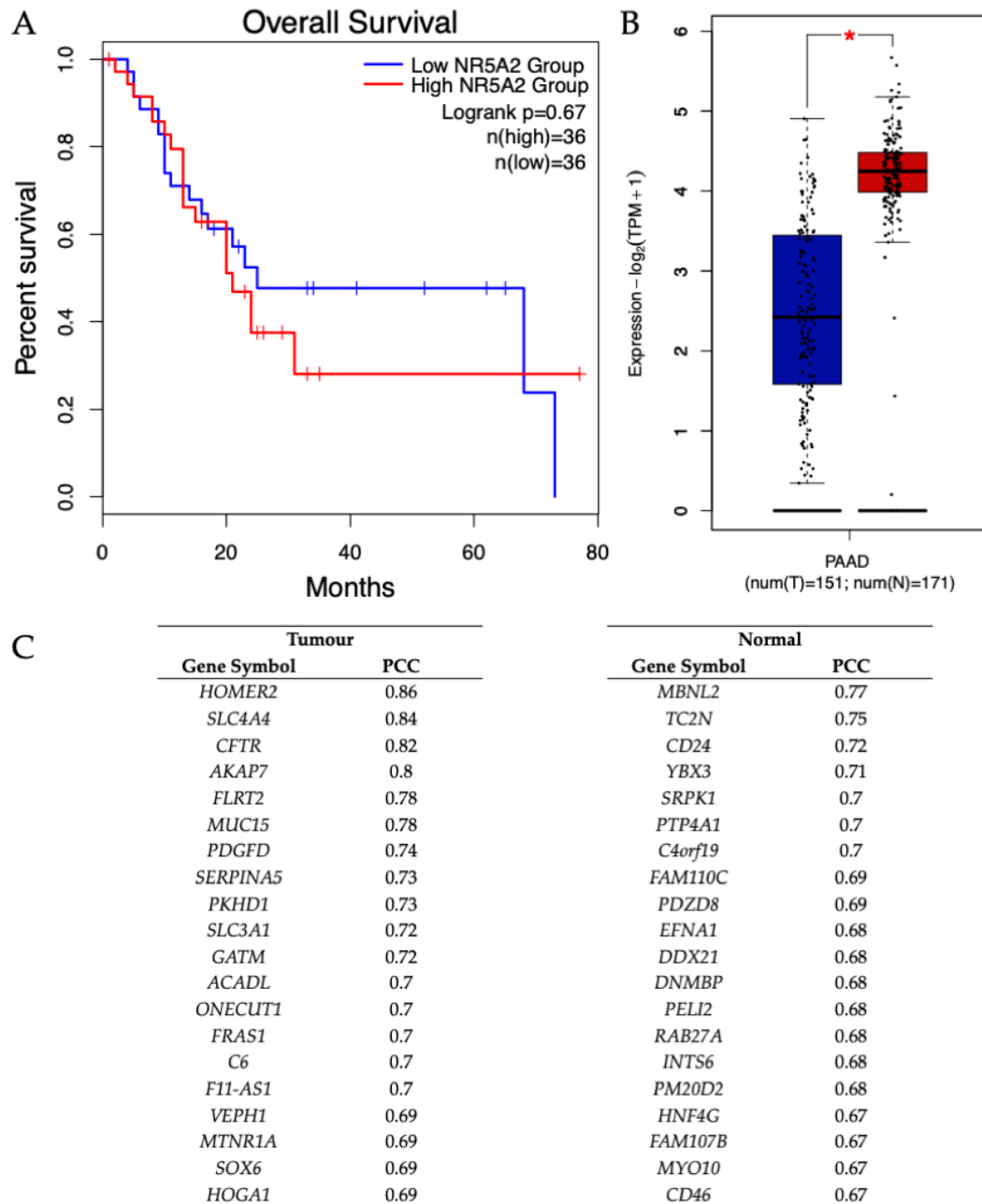


Figure 4.27: Analysis of NR5A2 using TCGA and GTEx datasets on GEPIA2 including (A) survival analysis of overall survival of pancreatic adenocarcinoma (PAAD) patients with high (≥ 80 th percentile) (red) and low (≤ 20 th percentile) (blue) NR5A2 expression. (B) Comparison of NR5A2 mRNA expression in tumour (blue) and normal (red) samples. (C) Pearson correlation coefficient (PCC) analysis identified genes with a similar expression pattern to NR5A2 in PDAC and normal pancreas.

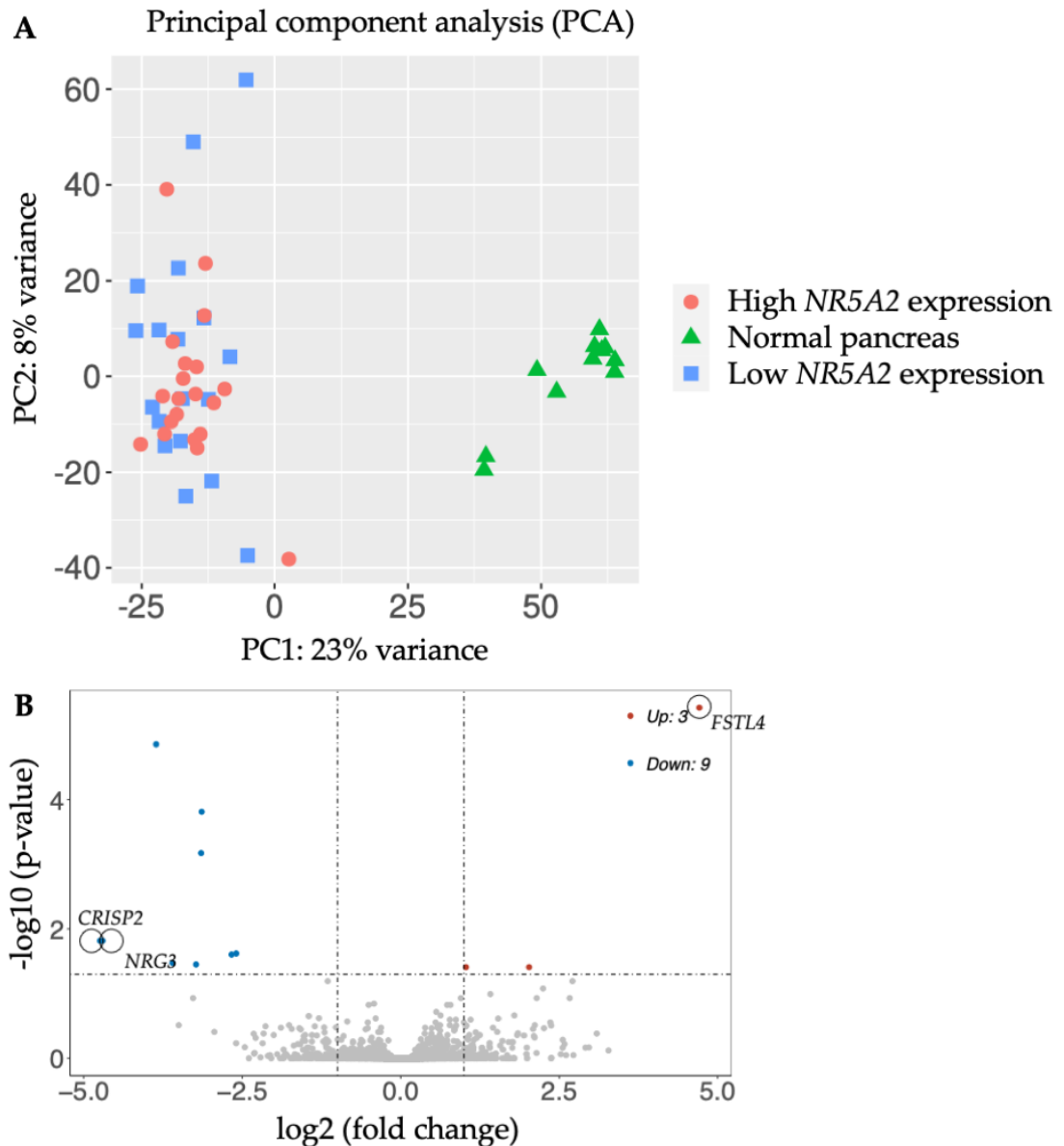


Figure 4.28. RNA-seq analysis of samples with high and low NR5A2 mRNA expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high NR5A2 mRNA expression (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low vs. high NR5A2 expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ P-value from logistic regression analysis.

4.2.2.5. Assessment of the role and expression of *PDXI*

The expression of *PDXI* is increased in PDAC tumour samples compared to normal pancreas. No significant difference was observed in the overall survival of PDAC patients with low *PDXI* expression ($\leq 20^{\text{th}}$ percentile) compared to high *PDXI* expression ($\geq 80^{\text{th}}$ percentile). PCC analysis was performed to identify genes with a similar expression pattern in PDAC and normal pancreas, and no overlapping genes were identified (Figure 4.29).

RNA-seq analysis of genes controlled by transcription factor *PDXI* (638 genes) found no clustering of samples with high *PDXI* expression compared to samples with low *PDXI* expression (Figure 4.30). No genes were found to be significantly up or downregulated.

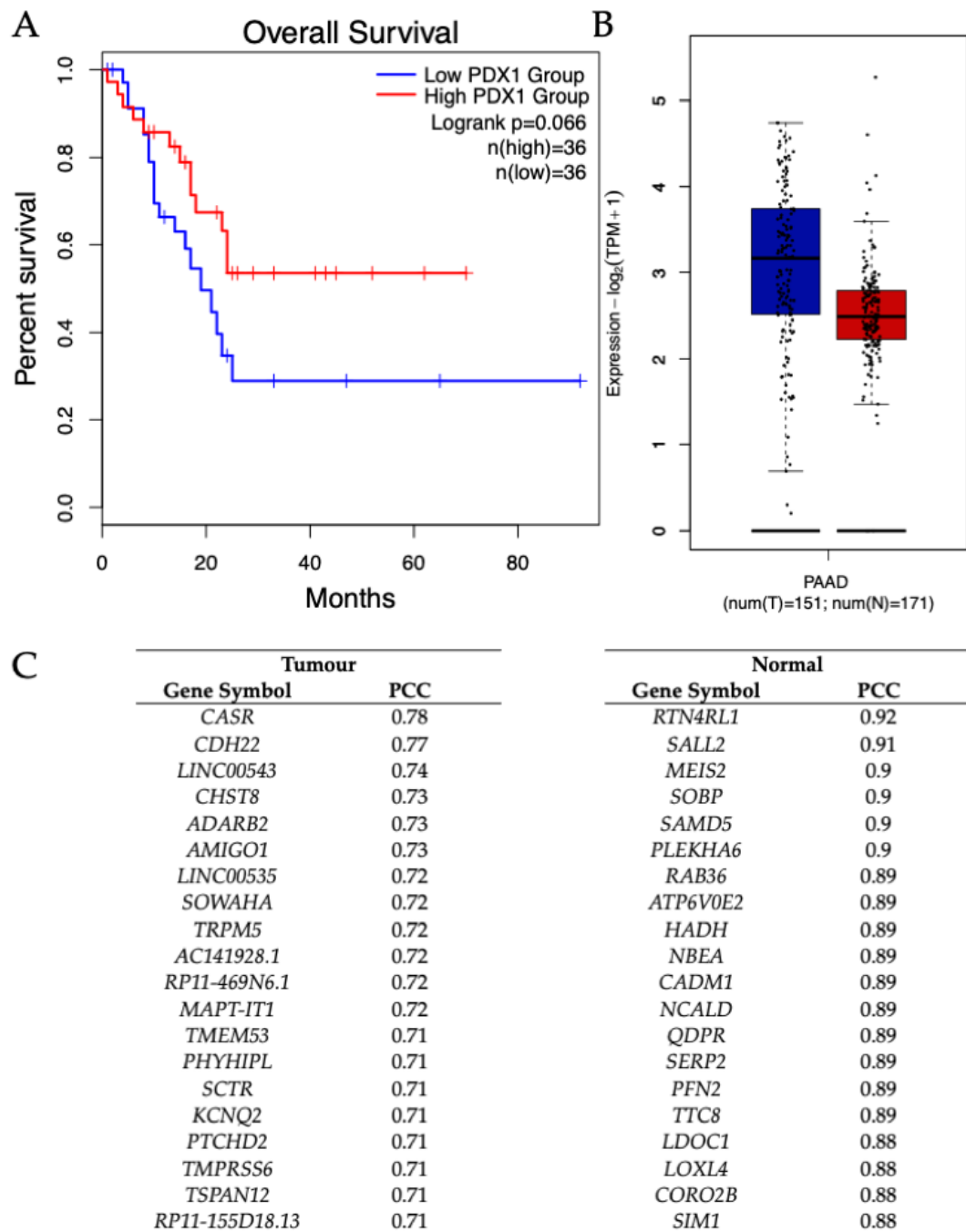


Figure 4.29: Analysis of PDX1 using TCGA and GTEx datasets on GEPIA2 including (A) survival analysis of overall survival of pancreatic adenocarcinoma (PAAD) patients with high ($\geq 80^{\text{th}}$ percentile) (red) and low ($\leq 20^{\text{th}}$ percentile) (blue) PDX1 expression in PDAC. (B) Comparison of PDX1 mRNA expression in tumour (blue) and normal (red) samples. (C) Pearson correlation coefficient (PCC) analysis identified genes with a similar expression pattern to PDX1 in PDAC and normal pancreas.

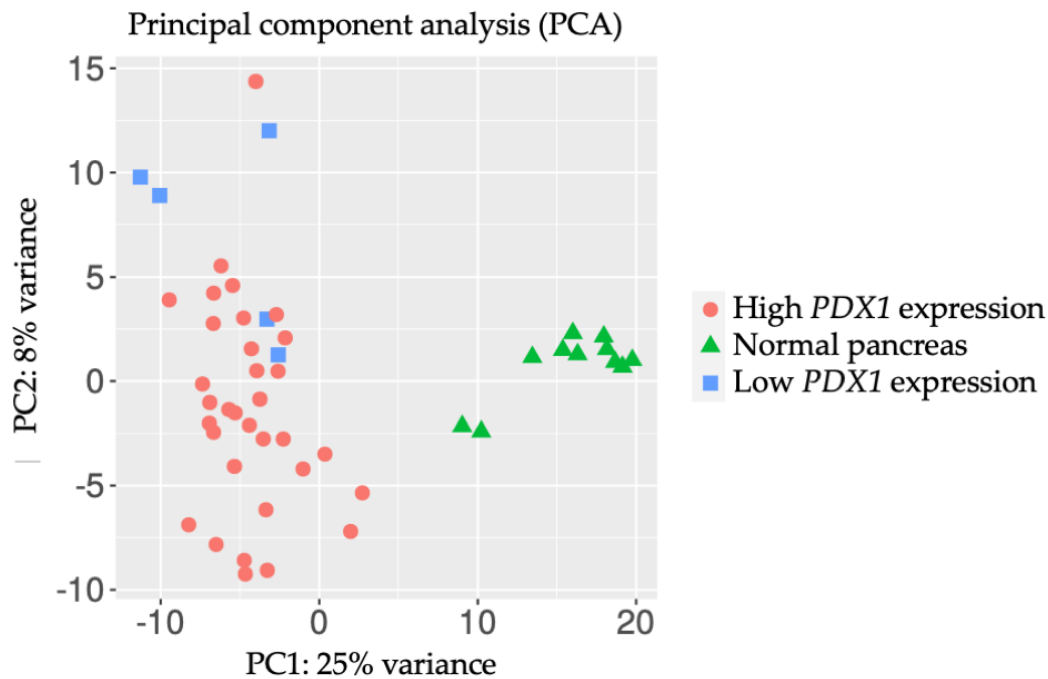


Figure 4.30: RNA-seq analysis of *PDX1* (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low *PDX1* expression and PDAC organoids with high *PDX1* expression.

4.2.2.6. RNA-seq analysis of *GATA6*

RNA-seq analysis of genes controlled by transcription factor *GATA6* (252 genes) found samples with high *GATA6* expression clustered together and clustered separately to samples with low *GATA6* expression (Figure 4.31 A). The most upregulated gene in samples with low *GATA* expression was *HPSE2* (log₂ fold-change 4.79; p-adj=2.03E-04). The most significantly downregulated gene was *REG4* (log₂ fold-change -6.71; p-adj=4.05E-12) (Figure 4.31 B).

DEG were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in *GATA6*-low and *GATA6*-high samples (Figure 4.32). Eleven pathways were significantly upregulated in *GATA6*-low samples. The most significantly upregulated pathway was blood vessel development, with 5 genes upregulated and an adjusted P-value of 4.90E-03 (Figure 4.32).

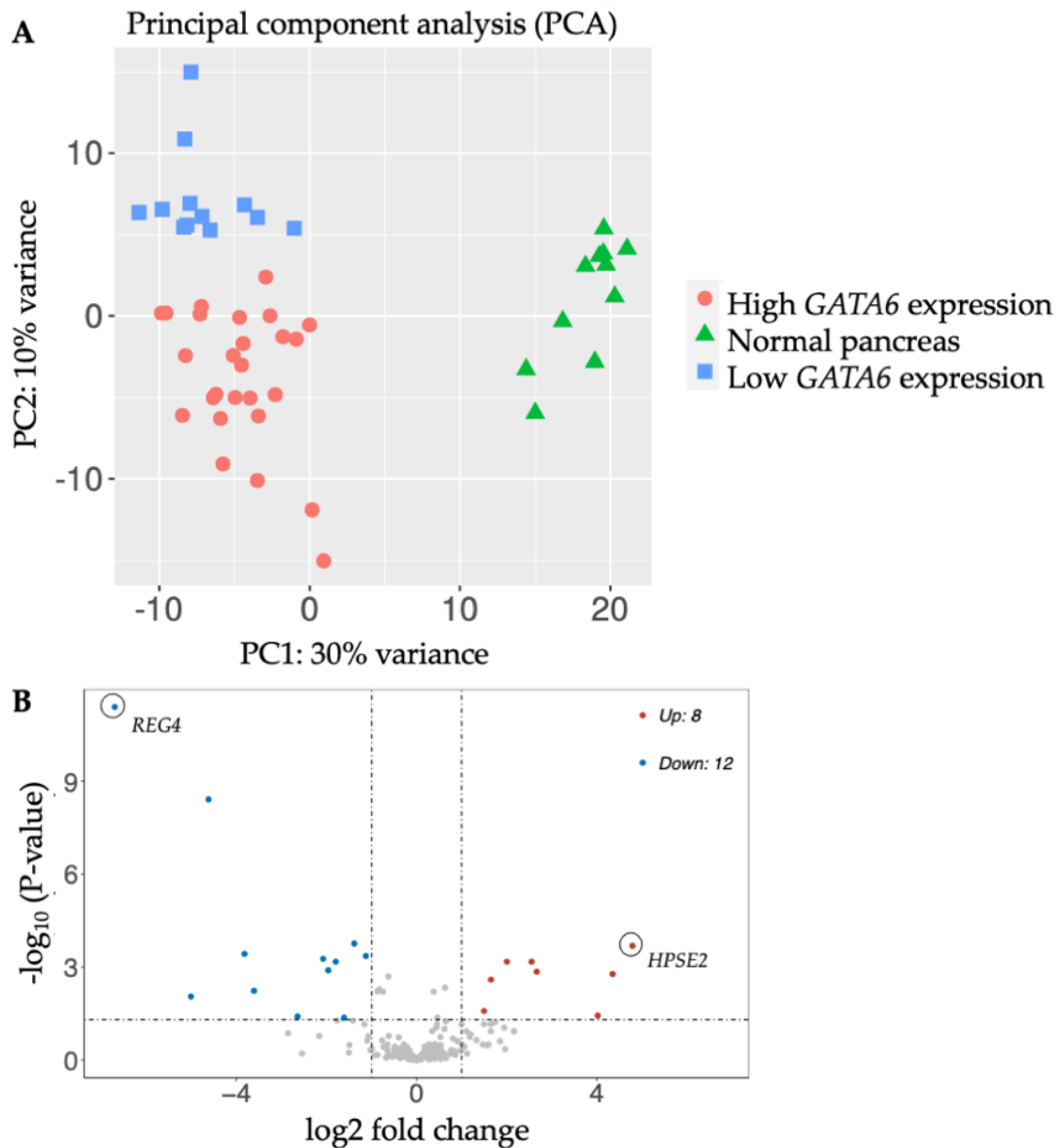


Figure 4.31: RNA-seq analysis of samples with high and low GATA6 expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high GATA6 mRNA expression. (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low vs. high GATA6 expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ FDR from logistic regression analysis.

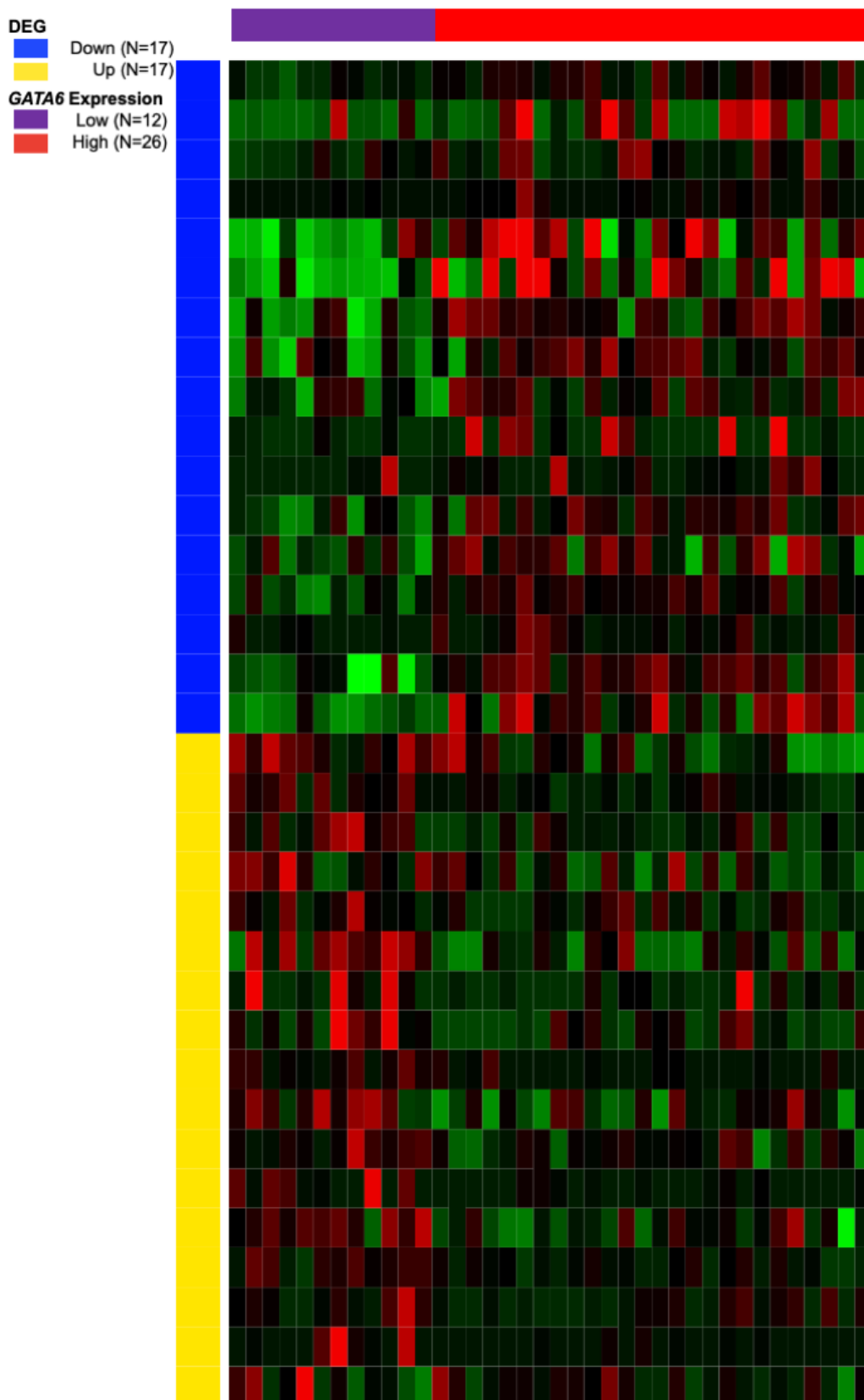


Figure 4.32: A heatmap shows upregulated (blue) and downregulated (yellow) differentially expressed genes (DEG) in GATA6-low (purple) and GATA6-high (red) samples.

Table 4.14: Differentially expressed genes were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in in samples with low and high GATA6 expression. The table lists the upregulated pathways, including adjusted P-value (p-adj), number of genes, and pathway name.

<i>P-adj</i>	<i>Number of Genes</i>	<i>Pathways</i>
Upregulated		
4.9E-03	5	Blood vessel development
4.9E-03	3	Outflow tract morphogenesis
4.9E-03	3	Mesoderm development
4.9E-03	6	Animal organ morphogenesis
4.9E-03	2	Forebrain neuron development
4.9E-03	4	Forebrain development
4.9E-03	3	Response to retinoic acid
4.9E-03	3	Ear morphogenesis
4.9E-03	2	Middle ear morphogenesis
4.9E-03	6	Positive regulation of transcription by RNA polymerase II
4.9E-03	4	Sensory organ morphogenesis

4.2.2.7. RNA-seq analysis of *HNF4A*

RNA-seq analysis of genes controlled by transcription factor *HNF4A* (887 genes) found samples with high *HNF4A* expression clustered together and clustered separately to samples with low *HNF4A* mRNA expression (Figure 4.33 A). The most upregulated gene in samples with low *HNF4A* expression was *COL8A1* (log2 fold-change 5.27; p-adj=8.34E-05). The most significantly downregulated gene was *HEPH* (log2 fold-change -8.95; p-adj=6.57E-06) (Figure 4.33 B).

DEG were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in *HNF4A*-low and *HNF4A*-high samples (Figure 4.34). Three pathways were significantly upregulated in and five pathways were significantly downregulated in *HNF4A*-low samples. The most significantly upregulated pathway was regulation of biological process, with 24 genes upregulated (p-adj=3.50E-03) (Figure 4.34). The most significantly downregulated pathway was negative regulation of developmental process, with 7 genes downregulated (p-adj=6.30E-03) (Table 4.15).

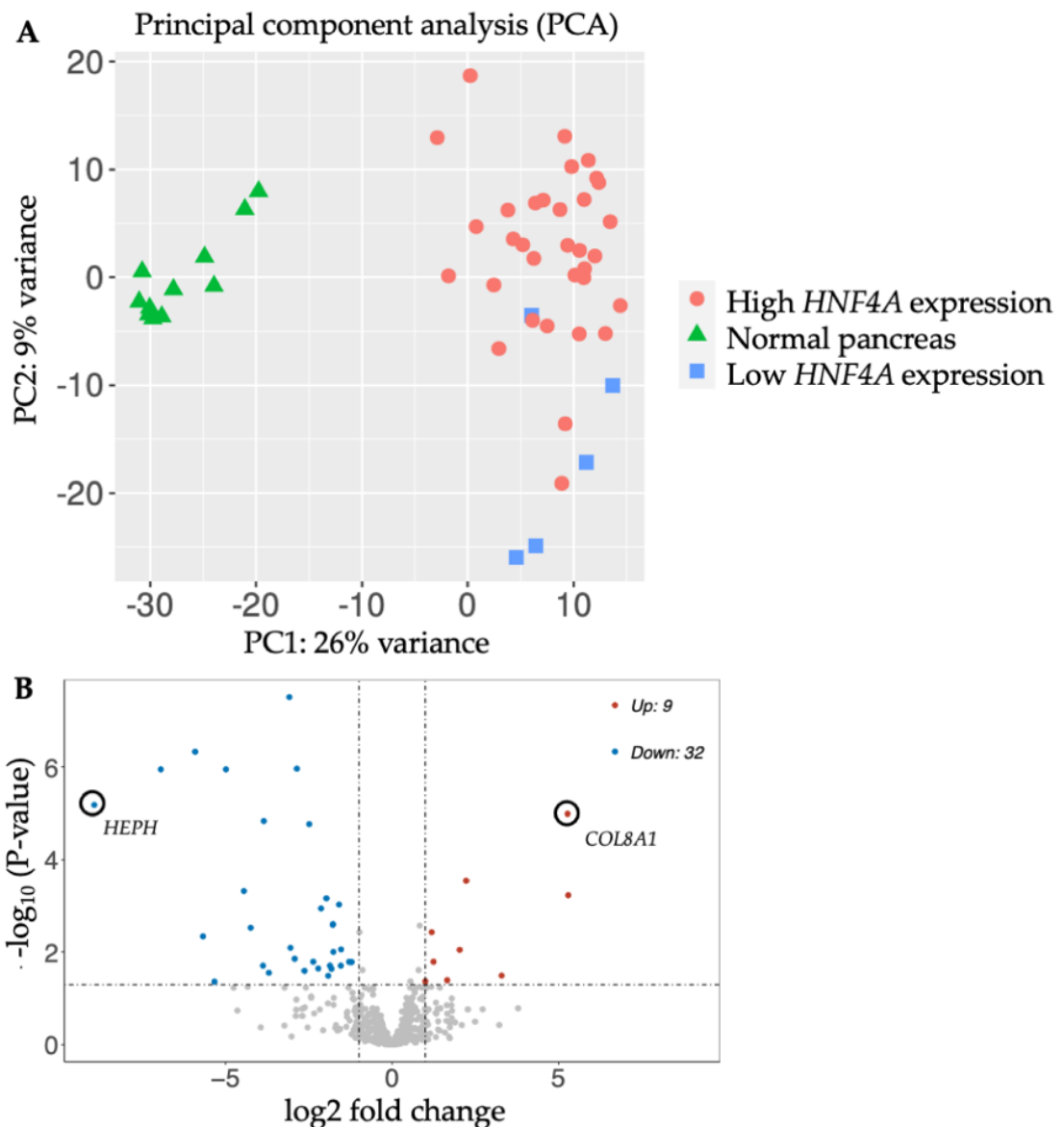


Figure 4.33: RNA-seq analysis of samples with high and low *HNF4A* expression (A) Principal component analysis was performed to compare the expression profiles of normal pancreas organoids, PDAC organoids with low and high *HNF4A* mRNA expression. (B) Volcano plots of the differentially expressed genes downregulated (blue) and upregulated (red) in PDAC organoids with low vs. high *HNF4A* expression. Each dot represents a gene; the X-axis shows the \log_2 transformed fold-change of differentially expressed genes. The Y-axis shows the $-\log_{10}$ FDR from logistic regression analysis.

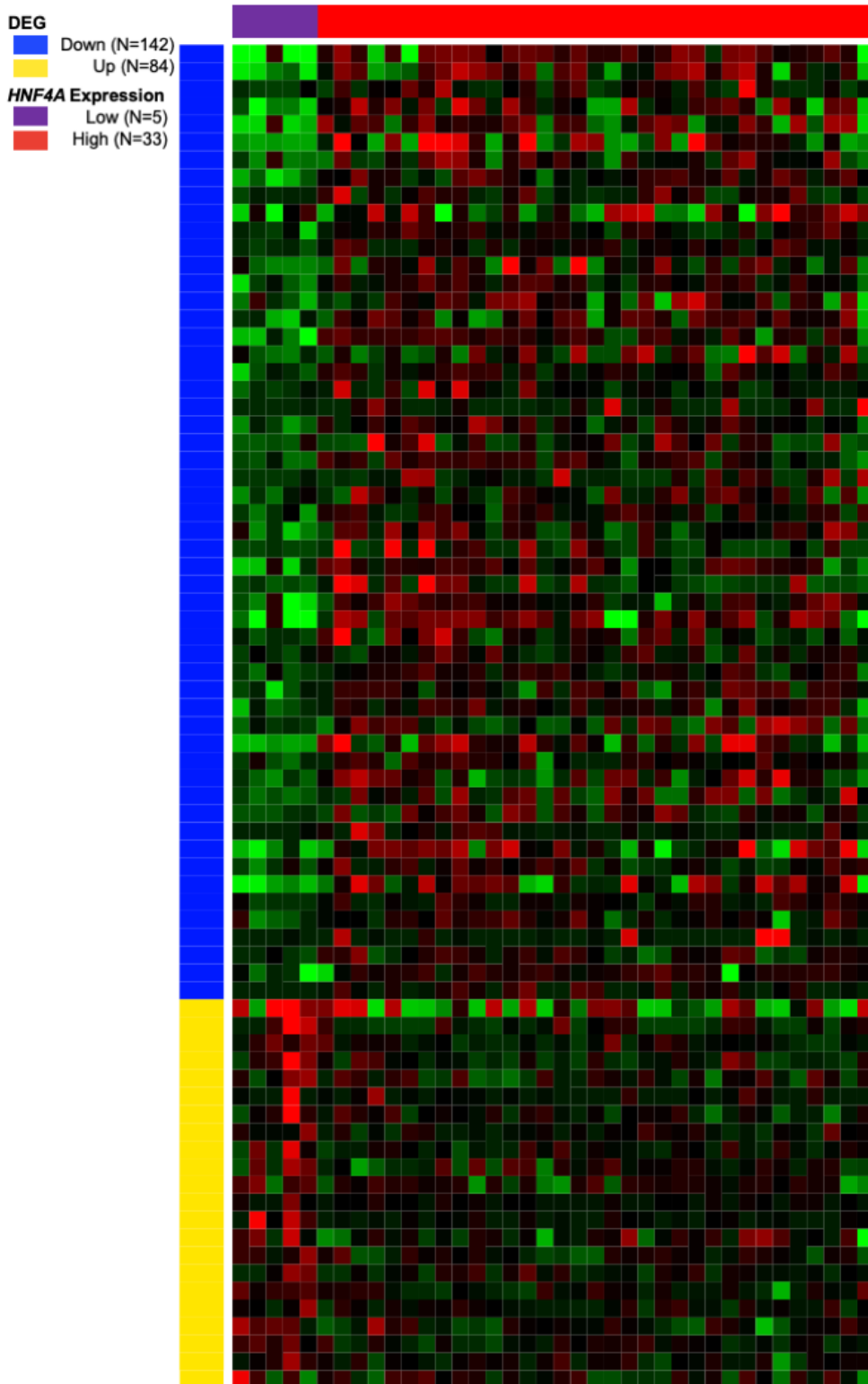


Figure 4.34: A heatmap shows upregulated (blue) and downregulated (yellow) differentially expressed genes (DEG) in HNF4A-low (purple) and HNF1A-high (red) samples.

Table 4.15: Differentially expressed genes were analysed to identify enriched pathways associated with upregulated and downregulated gene sets in low and high HNF4A expression. The table lists the upregulated and downregulated pathways, including adjusted P-value, number of genes, and pathway name.

<i>P-adj</i>	<i>Number of Genes</i>	<i>Pathways</i>
Downregulated		
3.5E-03	5	Columnar/cuboidal epithelial cell differentiation
3.5E-03	16	Homeostatic process
3.5E-03	24	Regulation of biological quality
Upregulated		
6.3E-03	2	Bone marrow development
6.3E-03	7	Negative regulation of developmental process
6.3E-03	2	Negative regulation of stem cell proliferation
7.2E-03	6	Negative regulation of cell differentiation
7.2E-03	2	Negative regulation of transforming growth factor beta production

4.2.3. Optimisation of use of lentiviral CRISPR to generate knockout organoids

To study the effect of the knockout of MODY genes, a method for the use of lentiviral CRISPR in organoids was developed. This work was funded by a EACR Travel Fellowship, and was performed in Aarhus University, Denmark under the supervision of Associate Professor Yonglun Luo. SgRNAs were generated using the Golden Gate Assembly method as described in Section 2.14. Cells were selected for one week using puromycin as a selection reagent. The efficiency of sgRNAs was first assessed by testing the generation of CRISPR knockouts in PT291 2D cell line for genes *HNFLA*, *HNFB*, *HNFB4G*, *NR5A2* and *PDX1* (Table 4.16). sgRNA resulted in successful gene editing of *HNFLA* (54%), *HNFB* (17%), *HNFB4G* (9%), and *NR5A2* (53%). No gene edits were observed in samples for *PDX1*.

Table 4.16: Knockout efficiency of lentiviral CRISPR using sgRNA in PT291 cell line.

<i>Gene</i>	<i>ICE</i>	<i>KO-Score</i>	<i>Fit Score - R²</i>
<i>HNFLA</i>	54	54	0.67
<i>HNFB</i>	28	17	0.98
<i>HNFB4G</i>	10	9	0.99
<i>NR5A2</i>	53	53	0.95
<i>PDX1</i>	0	0	1

Fit score represents inference of CRISPR edits (ICE) linear regression is computed during generation of the ICE Score, the Pearson correlation coefficient (Fit Score - R^2) is also computed and reported. The knockout-score represents the proportion of cells that have a frameshift.

These optimised sgRNAs were used in PT291 organoids, in addition, the non-efficient *PDX1* was used as a transfection control, to ensure the process does not result in cell death. Transduction with *HNFLA*, *HNFB* and *HNFB4G* sgRNAs in PT291 organoids resulted in organoids breaking down from the typical large spheroid morphology into single cells, then dying (Figure 4.35). This was observed to a lesser extent in *NR5A2* and was not observed in the negative control (no transduction), or with use of an inefficient sgRNA - *PDX1*.

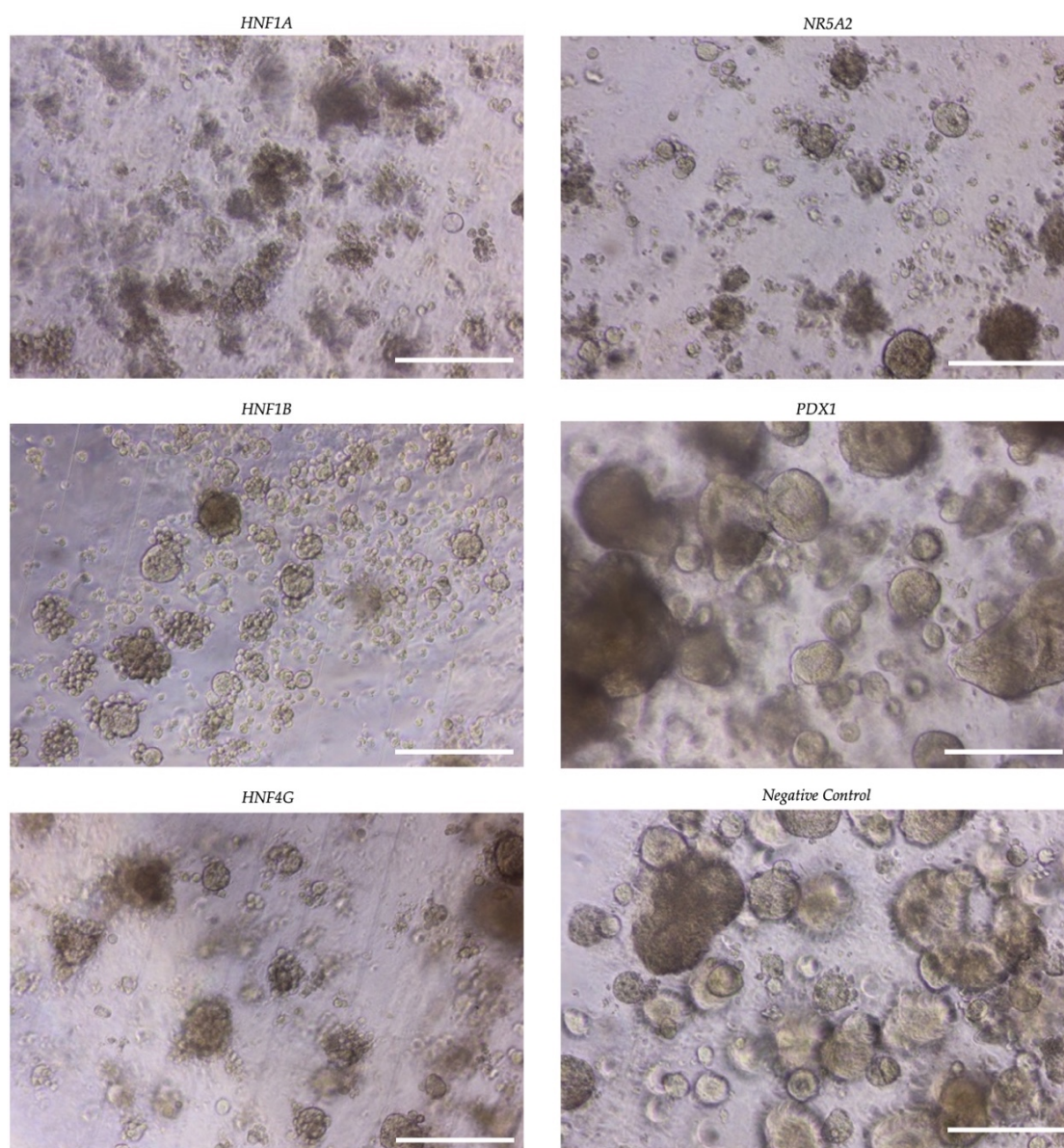


Figure 4.35: PT291 organoids imaged one week after transfection with CRISPR sgRNA targeting genes HNF1A, HNF1B, HNF4G, NR5A2 and PDX1, and negative control (no transduction). Organoids were selected by treatment with 1 μg/mL of puromycin for one week. Images taken at 100X magnification, scale bar 1000 μm.

4.2.4. CUT&RUN of *HNFI1A* and *HNFI1B* in CLO models

As described in Section 1.9.2 CUT&RUN is a novel technique used to profile protein-DNA interactions. This was performed in PT127 and PT291 CLO models to determine the role of transcription factors *HNFI1A* and *HNFI1B* which act as homodimers and heterodimers, as per the methods set out in Section 2.11. While the protocols were followed exactly, the sequencing results indicated that the experiment did not work. IGV software was used to compare the samples at whole genome level, for both PT127 (Figure 4.36) and PT291 (Figure 4.37) it is observed that there are no unique peaks in the sample compared to the negative control. Both samples also had different binding patterns to *HNFI1A* and *HNFI1B* in HepG2 cells, which were visualised as a comparison [283,284].

Additionally, the region upstream of *EIF4EBP2*, a gene known to be transcribed by *HNFI1A* and *HNFI1B*, was viewed. The genome viewer shows differential peaks between triplicate groups for both PT127 (Figure 4.38) and PT291 (Figure 4.39). Additionally, the samples bind regions which were also negative control, and the samples showed different binding patterns to those of *HNFI1A* and *HNFI1B* in HepG2 cells.

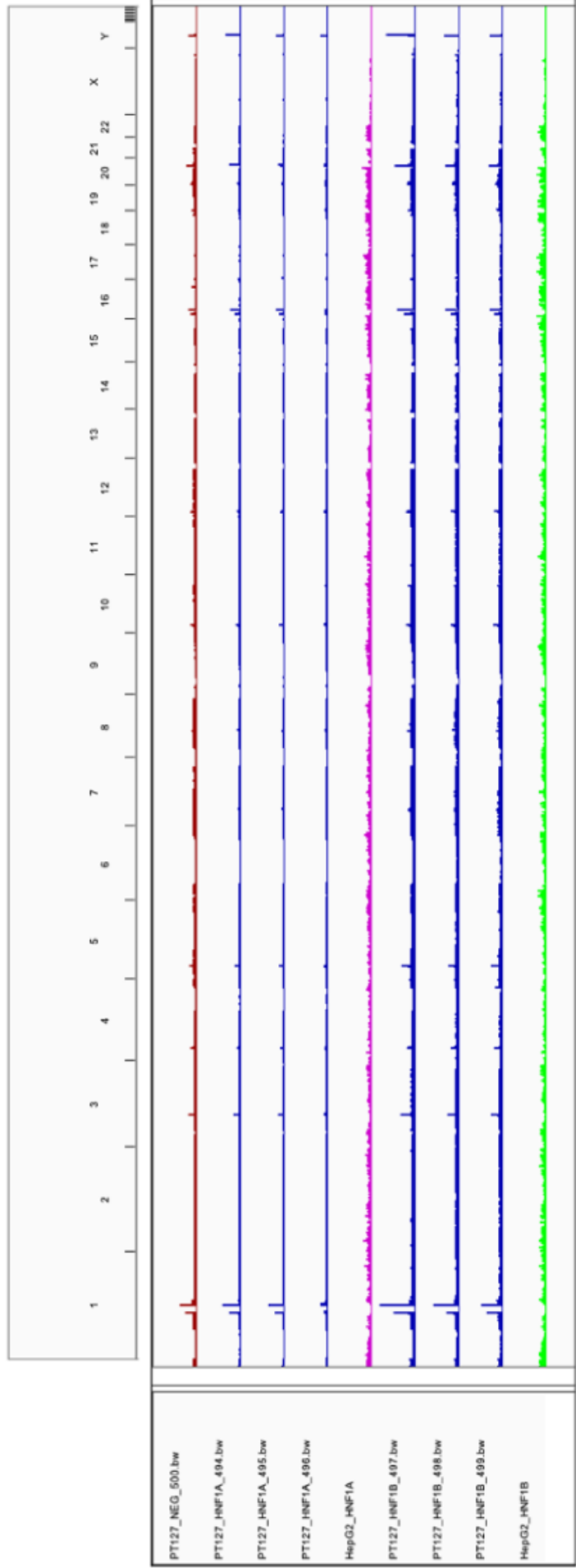


Figure 4.36: IGV whole genome view of sequencing results for CUT&RUN experiment for PT127 CLO. Red peaks represent the negative control, blue peaks represent samples sent for sequencing. Purple and green peaks represent ENCODE ChIP-seq data for HNF1A and HNF1B in Hep2G respectively.

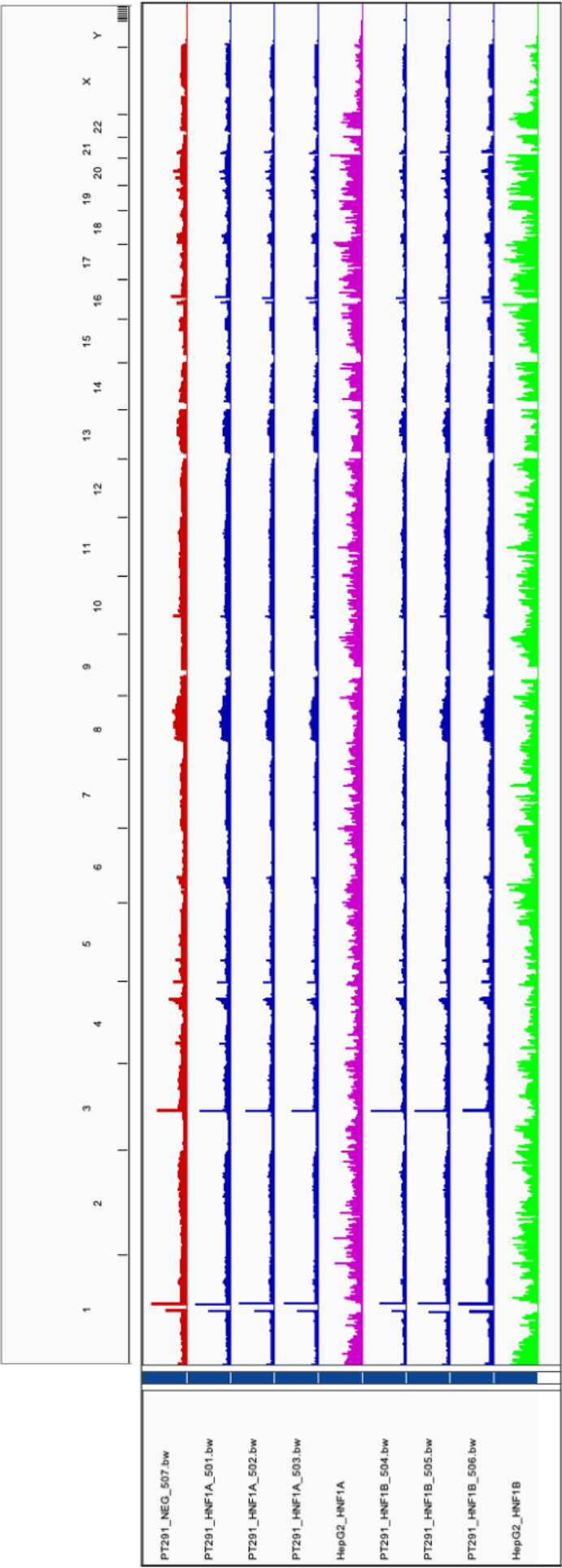


Figure 4.37: IGV whole genome view of sequencing results for CUT&RUN experiment for PT291 CLO. Red peaks represent the negative control, blue peaks represent ENCODE ChIP-seq data for HNF1A and HNF1B in Hep2G respectively. Purple and green peaks represent ENCODE ChIP-seq data for HNF1A and HNF1B respectively.

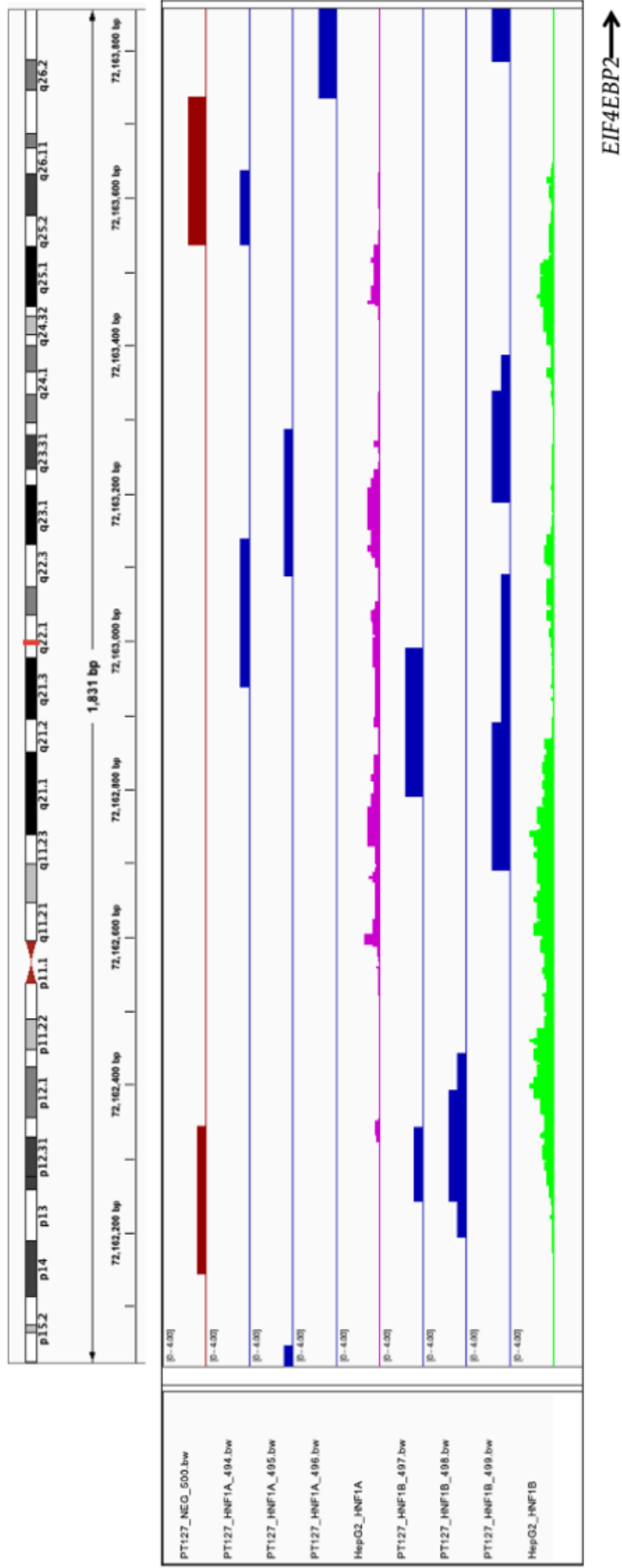


Figure 4.38: IGV view of sequencing results for CUT&RUN experiment for PT127 CLO upstream of EIF4EBP2, a gene transcribed by HNF1A and HNF1B. Red peaks represent the negative control, blue peaks represent samples sent for sequencing. Purple and green peaks represent ENCODE ChIP-seq data for HNF1A and HNF1B in Hep2G respectively.

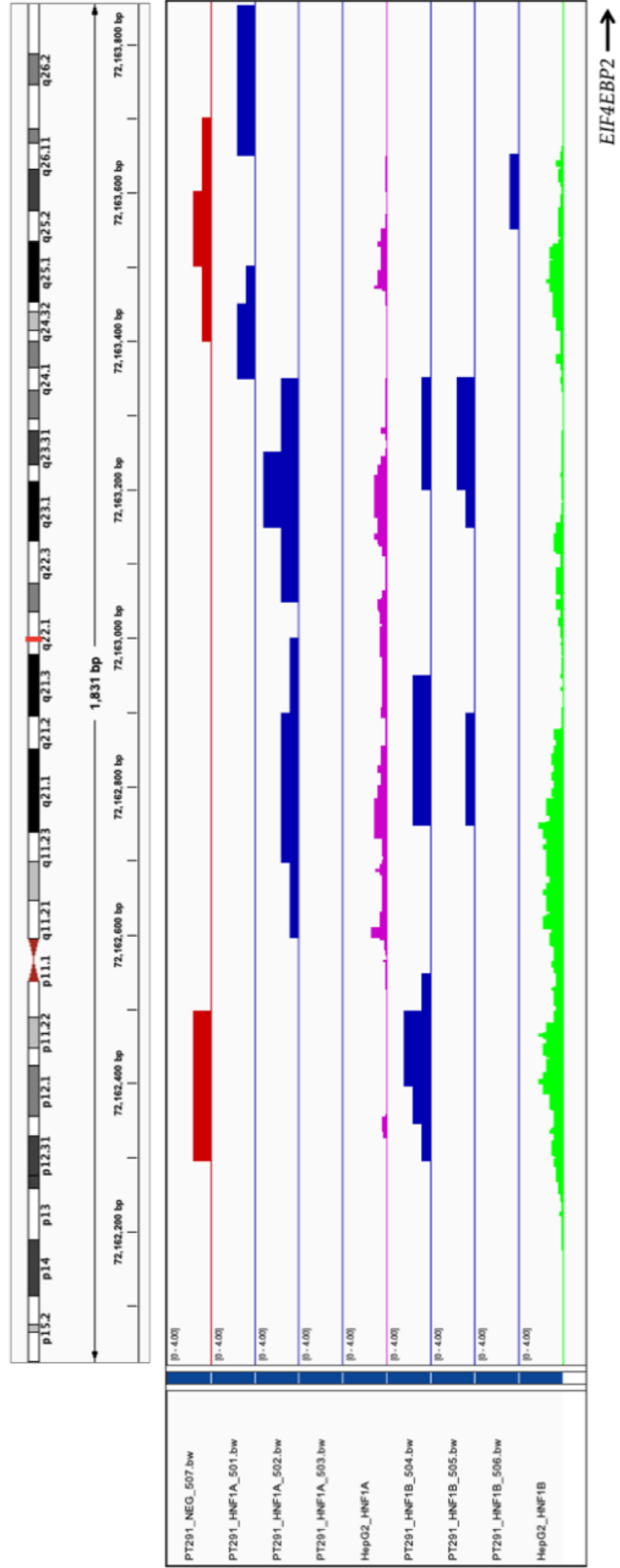


Figure 4.39: IGV view of sequencing results for CUT&RUN experiment for PT291 CLO upstream of EIF4EBP2, a gene transcribed by HNF1A and HNF1B. Red peaks represent the negative control, blue peaks represent samples sent for sequencing. Purple and green peaks represent ENCODE ChIP-seq data for HNF1A and HNF1B in Hep2G respectively.

4.3. Discussion

While GWAS has played a role in uncovering the mutational landscape of PDAC, little is known about how these polymorphisms effect the growth and development of the tumour [349,350]. An important step in improving outcomes for patients with pancreatic cancer includes validation of genomic variants of PDAC identified through GWAS and pathway analysis studies. The work in this chapter aims to identify the best candidate SNPs for functional validation. The GWAS pathway analysis study performed by Walsh *et al.* [122] identified five pathways and gene sets associated with the development of PDAC, including the MODY pathway. SNPs with an association with PDAC from all genes in the MODY pathway were identified through literature searches and brought forward for *in silico* analysis. As GWAS are performed using genotyping arrays – where a single candidate SNPs is used to represent all SNPs in a haplotype block – the index SNP identified in the study may not be the causative SNP [351]. Therefore, in addition to the SNPs identified through these searches, SNPs in LD with the SNP of interest ($R^2 > 0.8$) were included for analysis.

Two SNPs – rs2258287 and rs488087 – were selected for functional validation. These SNPs were selected as *in silico* analysis revealed that both SNPs were located in regions in the genome which were transcribed in the pancreas, including the presence of histone markers, and CpG islands [352,353]. Both SNPs are also within, or close to the nearest gene, rs488087 in exon 11 of *CEL* and rs2258287 is 7bp 5' of *CL2orf43*. In the assay performed to assess the function of SNP rs2258287, the presence of the alternative SNP (A) resulted in no change in luciferase gene expression in any line.

In the dual luciferase reporter assay performed to assess the function of SNP rs488087, the presence of the alternative SNP (T), results in increased luciferase expression in ASPC1, and a slight decrease in luciferase expression in PANC1. eQTL analysis from GTEx, showed that the presence of the alternative allele results in decreased expression of *CELP* in the normal pancreas. *CELP* is a pseudogene of *CEL*, which lacks exons 2-7. Exon 11 of *CEL* contains a variable number of tandem

repeat (VNTR) region consisting of 7 to 23 of nearly identical 33-bp segments [354–356]. Single base pair deletion mutations in the VNTR results in the onset of MODY, illustrating the potential of single mutations in the gene to be causative of disease [357]. A paper by Fjeld *et al.* [356] suggested that a crossover between *CEL* gene and *CELP*-pseudogene causes chronic pancreatitis, a known risk factor for the development of PDAC [358]. While the results showing the effect of rs488087 on gene expression in ASPC1 and PANC1 cell lines in the dual luciferase reporter assay are promising, the results of this study are limited. This SNP was also identified in a small study in a French cohort of 36 PDAC patients, and would need to be validated in a larger cohort [346]. The effect of this SNP in PDAC also requires further validation in a full panel of PDAC cell lines.

Analysis of the function of MODY TFs, and the genes they regulate was performed using RNA-seq data from a previously published PDAC organoid cohort. A correlation was observed between if the level of gene expression has a significant effect on patient survival, and if the high expression and low expression groups clustered in PCA. This was observed in transcription factors *HNF1B*, *PDX1*, and *NR5A2*, where high ($\geq 80^{\text{th}}$ percentile) and low ($\leq 20^{\text{th}}$ percentile) expression did not affect patient's overall survival (OS) and downstream gene PCA did not cluster. However, both differences in expression of *HNFLA* and *HNF4G* significantly affected patients' outcomes, with high *HNFLA* expression resulting in longer OS ($P=0.02$), and low *HNF4G* expression resulting a longer OS ($P=0.04$). A potential explanation is the hypothesis of gene dosage compensation, a mechanism by which the expression of certain genes is modulated to compensate for differences in gene dosage, so changes in gene expression do not result in a phenotypic change [359]. An example of this within these results is *NR5A2* – *NR5A2* and *NR5A1* both recognise the same DNA sequence, and both act with *GATA4* and *GATA6* to activate promoter activity [360,361].

Analysis of RNA-seq data from patients with low levels of expression of *HNF4G* compared to patients with high levels of *HNF4G* showed patients have significantly decreased expression of MODY pathway genes *HNFLA* and *HNF4A*. In addition, *HNFLA-AS1* was the most significantly downregulated gene in these patients. Little

research has been done into the role of *HNFLA-AS1* in PDAC, however it has been found to be downregulated in PDAC [362]. In oesophageal adenocarcinoma, *HNFLA-AS1* was found to regulate cell proliferation and migration [363]. *HNFLA-AS1* has also been identified to act as an oncogene in hepatocellular carcinoma through sponging tumour suppressor hsa-miR-30b-5p [364]. This data analysis found that in *HNFB*-low expression samples, the most downregulated gene is *SLCO1B1*, (solute carrier organic anion transporter family member 1B1) which is involved in drug transport. *SLCO1B1* has been found to be overexpressed in pancreatic cancer when compared with normal pancreatic tissue [365].

In addition to the MODY genes, the *GATA6* and *HNF4A* expression profiles were also assessed, as a loss of *GATA6*-expression with concomitant loss of *HNFLA* and *HNF4A*, results in a full phenotypic switch from classical subtype to basal subtype [366]. Decreased expression of *GATA6* and *HNF4A* results in the upregulation of genes involved in the remodelling of the cellular matrix, *HPSE2* and *COL8A1*, a collagen related gene. Patients with the basal subtype of PDAC have reduced survival times, and ECM deposition creates a physical barrier preventing T-cells from migrating toward tumour cells, preventing effective use of treatment with immunotherapies [367,368].

Preliminary results using CRISPR in PT291 organoids to knockout genes *HNFLA*, *HNFB*, *HNF4G* and *NR5A2* resulted in cell death, however cells in the negative control, and non-efficient PDX1 control survived. Given that these are preliminary results, caution must be taken prior to their use in concluding the function of the genes in PDAC. However, data from the Cancer Dependency Map (<https://score.depmap.sanger.ac.uk/>) shows that *HNFLA* is an essential gene in ASPC1, a cell line subtyped as classical PDAC. Results from Chapter 3 found that PT291 cell lines had a higher probability of being in the basal subtype than PT291 organoid models. Potentially, *HNFLA* is an essential gene in PT291 to maintain the organoid phenotype, so knockdown using CRISPR results in cell death. Additionally, the RNA-seq data analysis also showed that samples with low *HNF4G* expression also had significant downregulation of *HNFLA* compared to samples with high *HNF4G* expression, which may have the same effect. In addition to their

role in MODY, *HNFLA*, *HNFB*, and *HNF4G* are also involved in the regulation of beta-cell development pathway, and in the development of the embryonic pancreas. Further studies into the function of these genes is required to uncover their role in PDAC. It also highlights that it is of particular importance that the appropriate subtype of PDAC is being used to model their role in the disease.

One of the aims of this chapter was to assess the role of *HNFLA* and *HNFB* in PDAC and identify if these transcription factors worked together in the regulation of gene expression in the disease, however results were inconclusive. CUT&RUN was selected to perform this experiment for a number of reasons, including reduced cell numbers required for the experiment. ChIP-seq requires 10 million cells per replicate and obtaining this cell number using organoids or CLO models is challenging [369]. ChIP-seq also requires serum free conditions, which is difficult to model using organoids as the Wnt-3a/Rspondin3/Noggin required for the growth of organoids is derived from serum containing conditioned media from L-WRN cells. ChIP-seq also requires a large number of sequencing reads to differentiate signal from background, increasing the cost of the experiment [369]. Limitations of the CUT&RUN experiment include: a lack of optimised antibodies, at the time of writing, no antibodies have been validated for CUT&RUN for *HNFLA* or *HNFB* and there have been no publications using this method with these transcription factors. Finally, while qPCR should be performed for quality control prior to sequencing, CUT&RUN results in small (50 bp) regions, designing PCR primers to target these amplicons is extremely difficult [269].

In conclusion, here we propose that SNPs in the MODY genes may be involved in PDAC. This work found that rs488087 may have a functional role in gene regulation, eQTL analysis shows this SNP is associated with decreased expression of *CEL(P)*. Analysis of MODY transcription factors using publicly available data, in a PDAC RNA-seq dataset has highlighted differences in expression of *HNFLA* and *HNF4G* results in differences of overall survival of patients. Differential expression of both genes has downstream effects on a number of molecular pathways. These findings further underpinning the potential role of the MODY genes in PDAC.

Chapter 5. Assessing the role of the *ATM*/DNA repair pathway on drug treatment response

5.1. Introduction

DNA damage is a common event in cells and requires immediate repair to ensure that only exact genetic material is passed to daughter cells [370]. To prevent damaged DNA from being passed down to daughter cells, during the cell cycle, cells go through a series of cell cycle checkpoints before cell division [175]. In response to stress, the checkpoints are activated to prevent progression through the cell cycle and initiate repair [176]. In a process called the DDR, if the DNA damage is beyond repair, apoptosis is induced – this is to ensure the elimination of cells with mutations in the DNA [177]. The enzyme PARP acts to repair single stranded DNA breaks. In *BRCA* mutated cancers, inhibiting the enzyme activity of PARP results in synthetic lethality by preventing DNA repair (Figure 5.1) [371].

The GWAS pathway analysis performed by Walsh *et al.* [122] identified the Pujana ATM PCC Network, as one of the top five gene sets for risk of developing PDAC [186]. Within this gene set, a number of SNPs were associated with the development of the disease. SNP rs3124737 (*CASP7*) was identified as having a positive influence on *CASP7* expression in normal pancreas using GTEx and an independent pancreatic cancer dataset from the Laboratory of Translational Genomics in the NCI. Patients who are homozygous for this SNP had a reduced probability of developing PDAC (OR=0.91; 95% CI 0.88-0.95). Rs7859034, associated with *SMC2*, was identified as the top signal in the ATM PCC gene set 0.9 (95% CI 0.86-0.93) for rs7859034. GTEx analysis of rs7859034 did not identify any eQTL in the normal pancreas, however, the SNP did have an eQTL effect on the expression of *SMC2* in multiple tissues, including the heart, thyroid and brain.

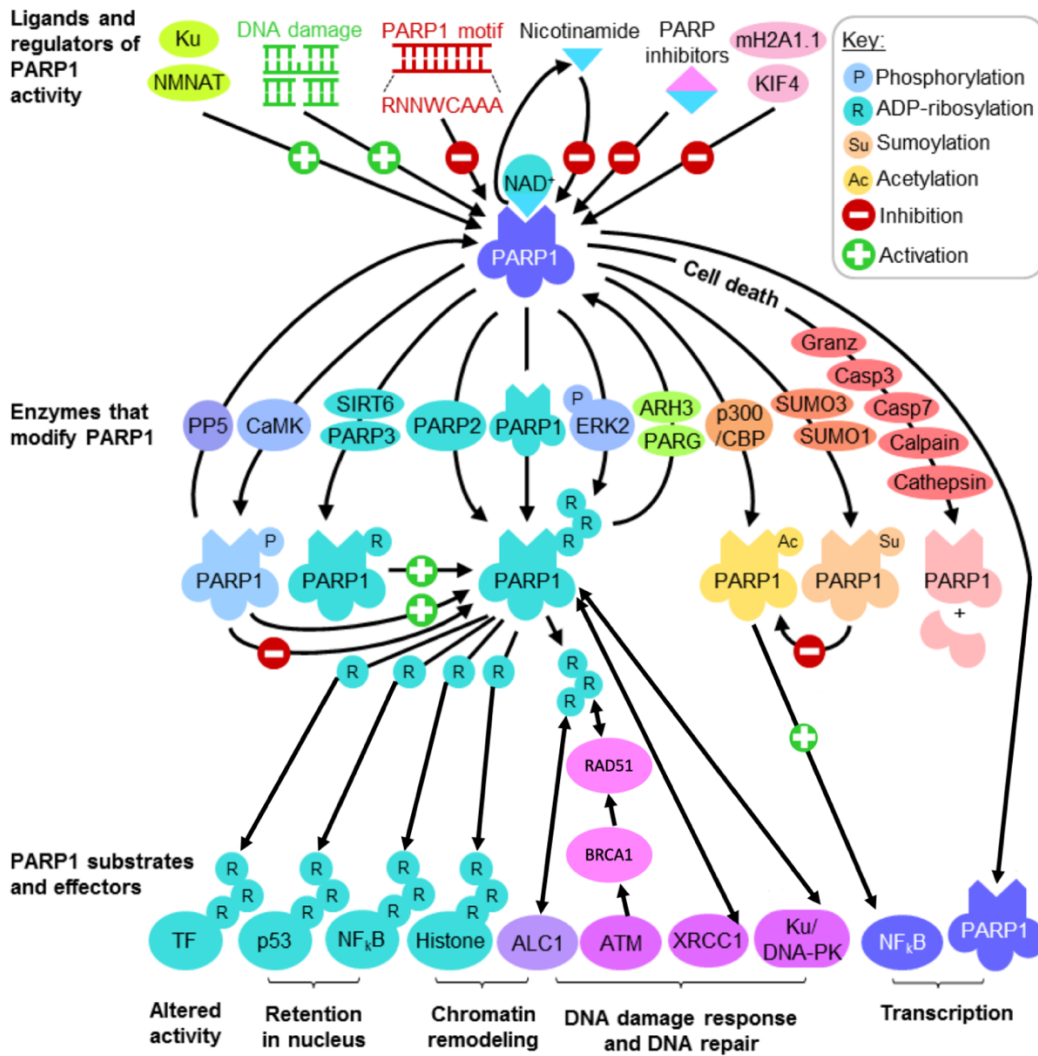


Figure 5.1: PARP1 function and regulation. Abbreviations: Granz: Granzyme; Casp: Caspase; TF: Transcription factor. Modified without permission from Ko et al. [371].

In addition to DDR SNPs altering the risk of PDAC, DDR genes include some of the most frequently mutated genes in cancer. These genes are able to induce the mutator phenotype and act as a driver of more mutations [177]. The unstable genome subtype of PDAC represents 14% of all cases and has a strong correlation with mutations in DDR genes, including *ATM*, *BRCA1/2* and *PALB2* [42]. In a study by Jones *et al.* [40] large-scale sequencing of PDAC has shown that on average, there are 63 mutations in the PDAC genome. In the PDAC classification by Collisson *et al.* [40,66], genes involved in the DDR pathway are some of the most commonly mutated genes in PDAC, with mutations in *ATM* the most frequent. In collaboration with the Cavalli group (Italy), the role of a novel DDR pathway targeting compound (ARN24089) in combination with cisplatin and PARP

inhibitor (PARPi) olaparib in the treatment of PDAC was studied. The chemical structure of these molecules is shown in Figure 5.2. The group synthesised a dihydroquinoline pyrazoline-based molecule (ARN24089) that disrupts the RAD51-BRCA2 protein-protein interaction, thus mimicking the effect of *BRCA2* mutation. ARN24089 inhibits the homologous recombination in a human pancreatic adenocarcinoma cell line. In addition, it synergises with olaparib to trigger synthetic lethality. This strategy aims to widen the use of PARPi in BRCA-competent and olaparib-resistant cancers, making fully small-molecule-induced synthetic lethality an innovative approach toward unmet oncological needs [372].

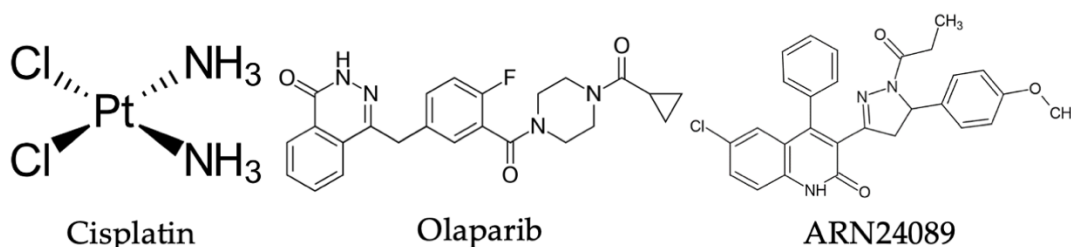


Figure 5.2: Chemical structure of compounds used in this chapter, cisplatin, olaparib and ARN24089. Figure modified without permission from Jamieson and Lippard [373], Chen et al. [374] and Bagnolini et al. [372].

5.1.1. Aims

Aim 1: Identification and validation of regulatory SNPs capable of modulating genes involved in the DDR pathway in PDAC, through *in silico* SNP analysis and dual luciferase reporter assays.

Aim 2: Assess the role of genes associated with SNPs of interest in PDAC, through IHC, RT-qPCR, and using publicly available online databases.

Aim 3: Using novel drug ARN24089 to target the DDR pathway in non-*BRCA* mutated PDAC, in combination with cisplatin and olaparib. Assessing the mechanism of action of the drug through assessing levels of cleaved caspase-3 and -7 and assessing RAD51 expression.

5.2. Results

5.2.1. Functional analysis of SNPs and genes associated with ATM and DDR pathways identified through GWAS pathway analysis

All candidate SNPs identified in PDAC GWAS pathway analysis (Figure 5.3), and SNPs in LD with a $R^2 > 0.8$ were assessed for prioritisation for functional analysis, as per methodology described in Section 2.12.

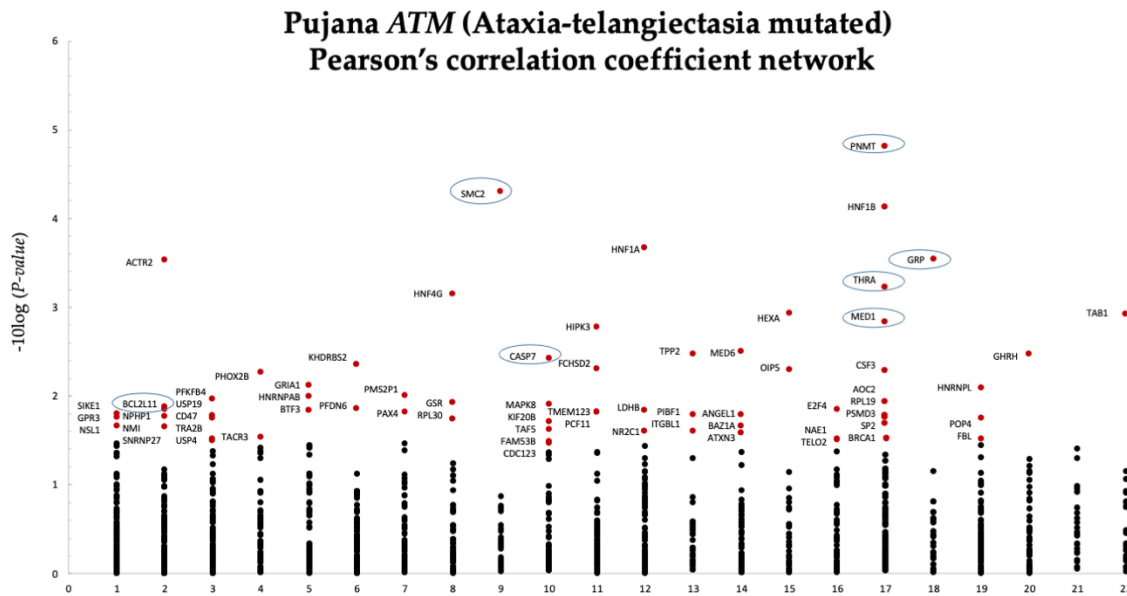


Figure 5.3: Genes associated with PDAC in the Pujana ATM Pearson correlation coefficient (PCC) network gene set ($P = 1.25E-05$). A total of 67 genes (red circles, $P < 0.02$) were identified from the gene set as contributing the most to PDAC association in the GWAS pathway analysis. Genes of interest highlighted using circle. Modified without permission from Walsh et al. [122].

5.2.1.1. SNPs identified in DDR genes associated with PDAC

GWAS SNPs, GWAS pathway analysis SNPs and SNPs from literature searches resulted in the identification of 12 SNPs in genes associated with the DDR pathway. The genes of interest identified, and their role in the normal pancreas is outlined in (Table 5.1). The SNPs identified are listed in Table 5.2. SNPs in LD with these were identified using LD link, and LD heatmaps were generated. To refine the SNPs identified for post-GWAS functional validation, the best candidate from the SNP blocks was identified using the criteria shown in Figure 5.4 and brought forward for functional analysis.

Table 5.1: DNA damage repair (DDR) pathway genes associated with SNPs identified in literature searches.

Gene Name	Symbol	Function
B-cell lymphoma 2 like 11	<i>BCL2L11</i>	Apoptotic regulator
Caspase 7	<i>CASP7</i>	Executioner protein of apoptosis
Gastrin releasing peptide	<i>GRP</i>	Stimulates the release of gastrin
Mediator complex subunit 1	<i>MED1</i>	Coactivator of RNA polymerase II-dependent genes
Phenylethanolamine N-methyltransferase	<i>PNMT</i>	Methylation of norepinephrine
Structural maintenance of chromosomes 2	<i>SMC2</i>	Conversion of interphase chromatin
Thyroid hormone receptor alpha	<i>THRA</i>	Nuclear hormone receptor

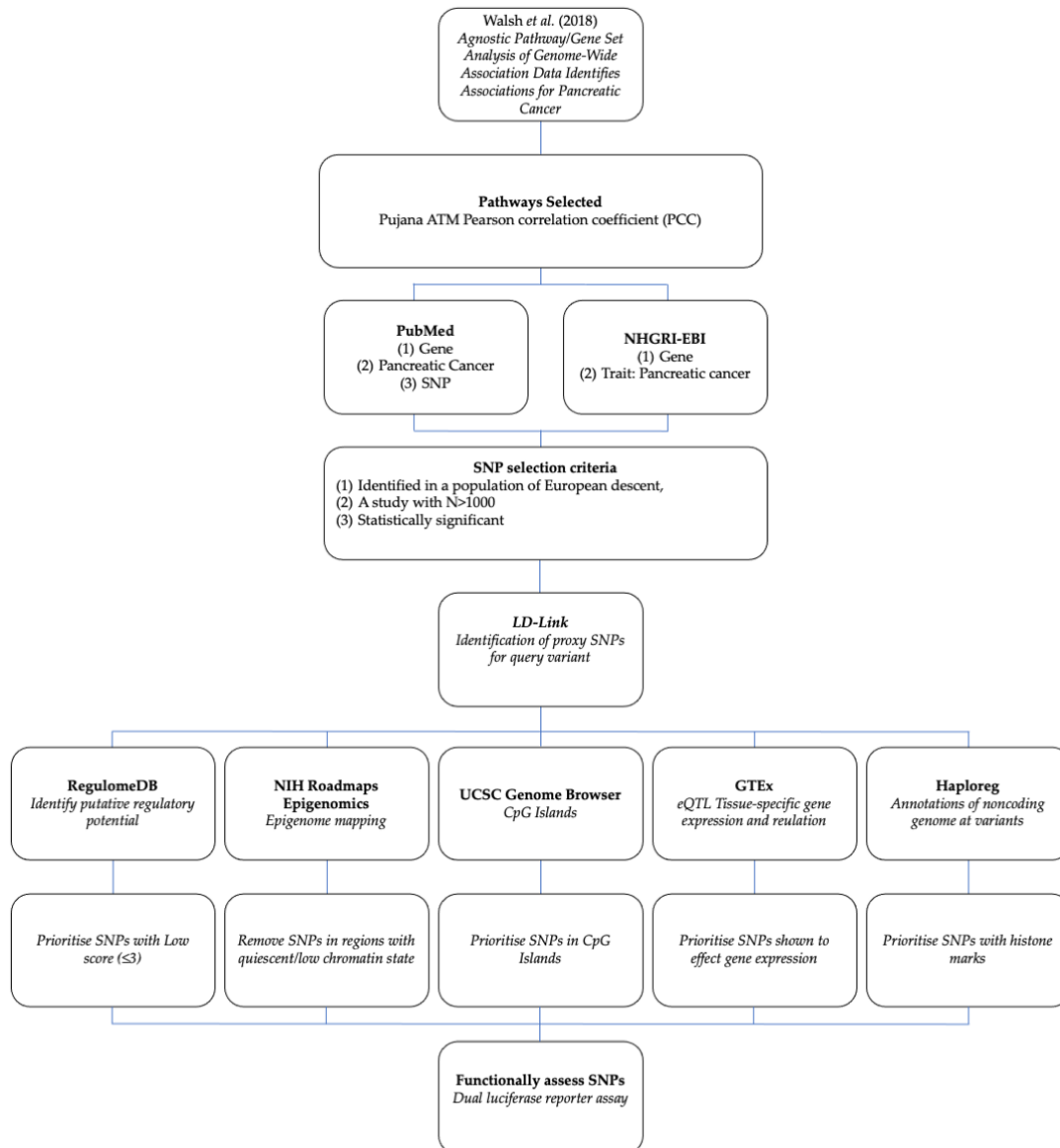


Figure 5.4: The criteria used for SNP identification, in silico assessment and prioritisation of SNPs from the Pujana ATM PCC gene set from the study by Walsh et al. [122].

All candidate SNPs identified in PDAC GWAS studies, and SNPs in LD with a regulomeDB (regDB) score ≤ 3 and an $R^2 > 0.8$ which met the criteria (Figure 5.4) were assessed for functional analysis. In total, 12 SNPs were identified (Table 5.2) and a dual luciferase reporter assay was performed to assess if prioritised SNPs had a functional effect on gene expression.

Table 5.2: SNPs located in DDR pathway genes predicted to have a role in the development of PDAC.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	P-value	Cases Controls	Allelic OR (95% CI)	Location/ Function	Ref
BCL2L1	rs13396983	2	111900598	A G	6.80E-04	3,851 3,934	1.12 (1.05–1.20)	Intronic	[121]
BCL2L1	rs2015454	2	111872148	A G	7.00E-04	3,851 3,934	1.12 (1.05–1.20)	Intronic	[121]
CASP7	rs3124737#	10	115497095	A G	3.29E-05	9,040 12,496	0.91 (0.88-0.95)	6.4kb 3' of CASP7	[122]
GRP	rs1517037*	18	56878274	C T	7.56E10-06	9,040 12,496	0.82 (0.75–0.89)	9.1kb 5' of GRP	[116]
GRP	rs57791062#	18	56880211	C T	1.79E-06	9,040 12,496	0.88 (0.83-0.93)	7.2kb 5' of GRP	[122]
MEDI	rs12947620	17	37605364	T C	5.46E-05	8,477 6,946	1.11 (1.06-1.17)	Intronic	[375]
PNMT	rs876493#	7	37824545	G A	1.27E-06	9,040 12,496	0.9 (0.86-0.94)	5'UTR	[122]
SMC2	rs3818626	9	106856633	C T	2.20E-06	8,477 6,946	1.12 (1.07–1.17)	5'UTR	[376]
SMC2	rs2417487*	9	106887581	A G	1.49E-07	9,040 12,496	1.11 (1.07-1.15)	Intronic	[116]
SMC2	rs10991043*	9	106797388	T C	3.53E-07	9,040 12,496	1.12 (1.07-1.16)	59kb 5' of SMC2	[116]
SMC2	rs7859034#	9	106856590	G T	3.07E-07	9,040 12,496	0.9 (0.86-0.93)	Intronic	[122]
THRA	rs8078692#	17	38215117	A G	6.71E-06	9,040 12,496	0.91 (0.87-0.95)	—	[122]

This table includes name of nearest functional gene (20kb upstream/downstream), SNP marker, HG19 chromosome position, reference and alternative alleles, allelic odds ratio, and reference paper in which SNP was identified. *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

5.2.1.2. Analysis of *BCL2L1* associated SNPs

B-cell lymphoma 2 like 11 (*BCL2L1*) is a gene located at 2q13 and contains 14 exons. *BCL2L1* has high levels of expression in lymph nodes, thyroid and bone marrow, and low levels of expression in the pancreas, where it acts as an apoptotic regulator. Two SNPs identified from GWAS literature searches, and the five SNPs identified through proxy analysis are outline in an LD block ($R^2 > 0.8$; $n=7$) and highlighted in a heatmap (Figure 5.5). The full *in silico* analysis of the 7 SNPs associated with *BCL2L1* is outlined in Table 5.3. SNPs were then prioritised for functional analysis based on the criteria outlined in Figure 5.4. SNPs rs11378324 and rs2241845 was selected for functional analysis as both are in CpG islands, and rs11378324 is within a predicted enhancer site, and rs2241845 is in an active transcription start site in the pancreas. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.6.

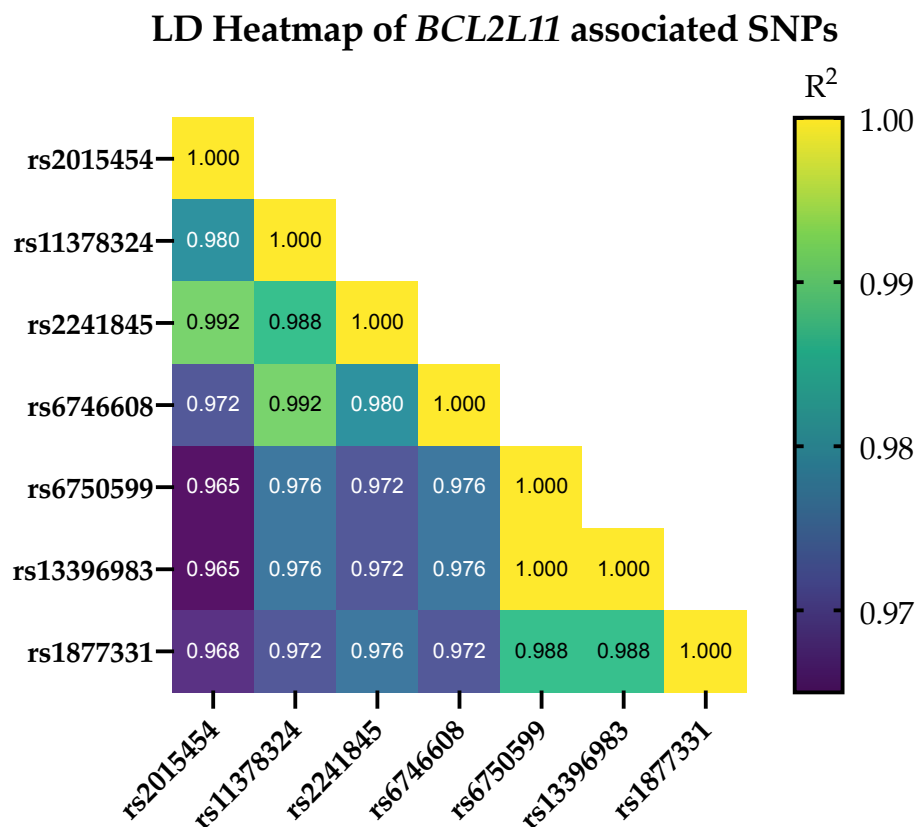


Figure 5.5. LD heatmap of *BCL2L1* SNPs of interest based on R^2 values in European populations (LDLink).

Table 5.3: Full in silico analysis of SNPs associated with the gene *BCL2L1I*. Underlined SNPs were selected for functional analysis.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function	H3K27ac	CpG Island	eQTL (NES)
<i>BCL2L1I</i>	rs13396983	2	111900598	A G	3a	Strong transcription	Intronic	—	—	↓ -0.57, Brain
<i>BCL2L1I</i>	rs2015454	2	111872148	A G	7	Quiescent/Low	Intronic	—	—	↓ -0.61, Brain
<u><i>BCL2L1I</i></u>	<u>rs2241845</u>	<u>2</u>	<u>111879100</u>	<u>A G</u>	<u>2b</u>	<u>Active TSS</u>	<u>Intronic</u>	—	<u>Yes</u>	↓ -0.62, Brain
<u><i>BCL2L1I</i></u>	<u>rs11378324</u>	<u>2</u>	<u>111877346</u>	<u>T GGGGGG</u>	<u>3a</u>	<u>Enhancers</u>		—	<u>Yes</u>	↓ -0.57, Brain
<i>BCL2L1I</i>	rs6746608	2	111892984	A G	1f	Strong transcription	Intronic	—	—	↓ -0.57, Brain
<i>BCL2L1I</i>	rs1877331	2	111906510	G A	3a	Strong transcription	Intronic	Yes	—	↓ -0.62, Brain
<i>BCL2L1I</i>	rs6750599	2	111893869	A T	1f	Strong transcription	Intronic	—	—	↓ -0.55, Brain

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG island. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

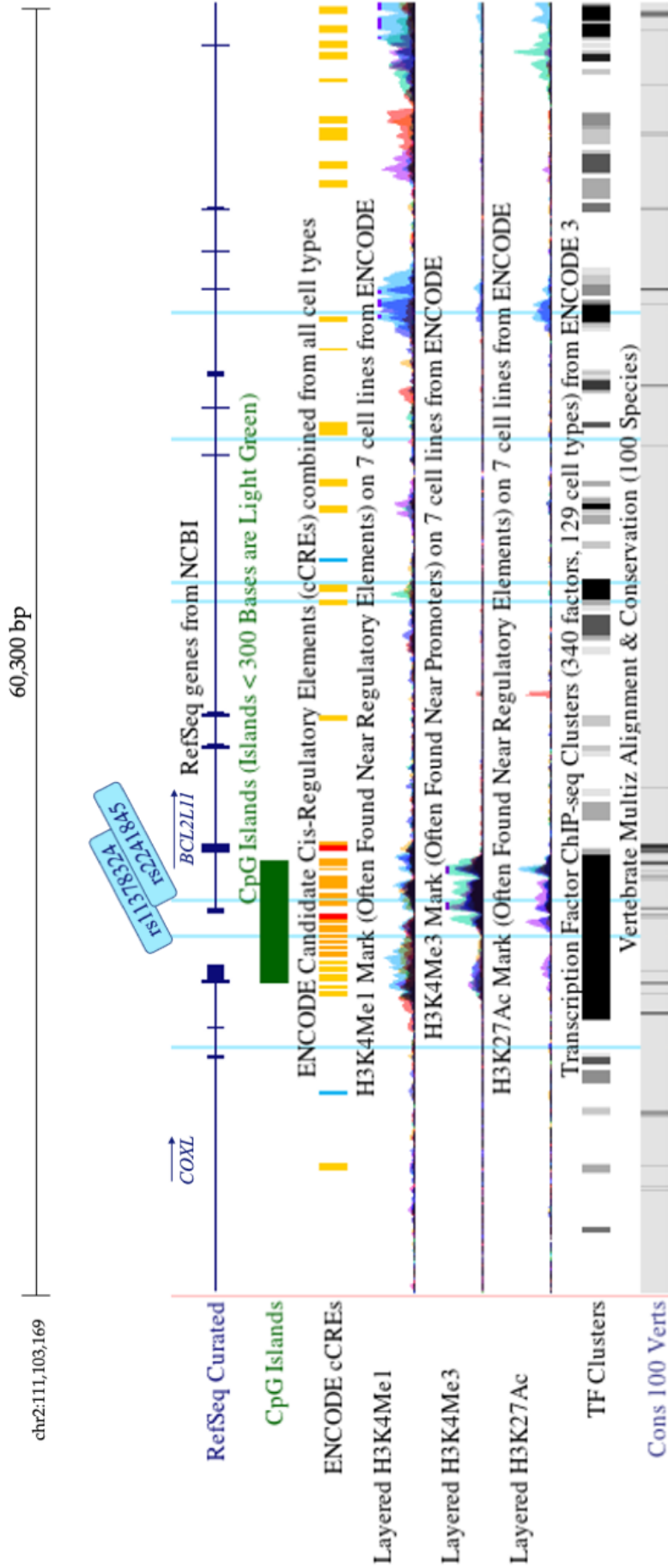


Figure 5.6: UCSC Genome Browser view of BCL2L1 gene regions surrounding SNPs of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.3. Analysis of CASP7 associated SNPs

Caspase7 (*CASP7*) is a gene located at 10q25.3 and contains 13 exons. *CASP7* has high levels of expression in colon, small intestine and duodenum, and low levels of expression in the pancreas, where it acts as the executioner of apoptosis. One SNP was identified from GWAS pathway analysis, and no SNPs within the selection criteria (RegulomeDB <4; $R^2 > 0.8$) were identified through proxy analysis. The full *in silico* analysis of rs3124737 is outlined in Table 5.4. rs3124737 was selected for functional analysis as it was identified as having a positive influence on *CASP7* expression in normal pancreas using GTEx. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.7.

Table 5.4: Full in silico analysis of SNPs associated with the gene *CASP7*. Underlined SNPs were selected for functional analysis. # Highlights GWAS pathway analysis SNPs.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function	eQTL (NES)
------	--------	-----	--------------------	--------------------	-------	-----------------------------	-----------------------	------------

<u>CASP7</u>	<u>rs3124737#</u>	<u>10</u>	<u>115497095</u>	<u>A</u> <u>G</u>	<u>6</u>	<u>Strong transcription</u>	<u>6.4kb 3' of CASP7</u>	<u>↑ 0.48, Pancreas</u>
--------------	-------------------	-----------	------------------	-------------------	----------	-----------------------------	--------------------------	-------------------------

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas. eQTL includes NES and tissue. #Highlights GWAS pathway analysis SNPs.

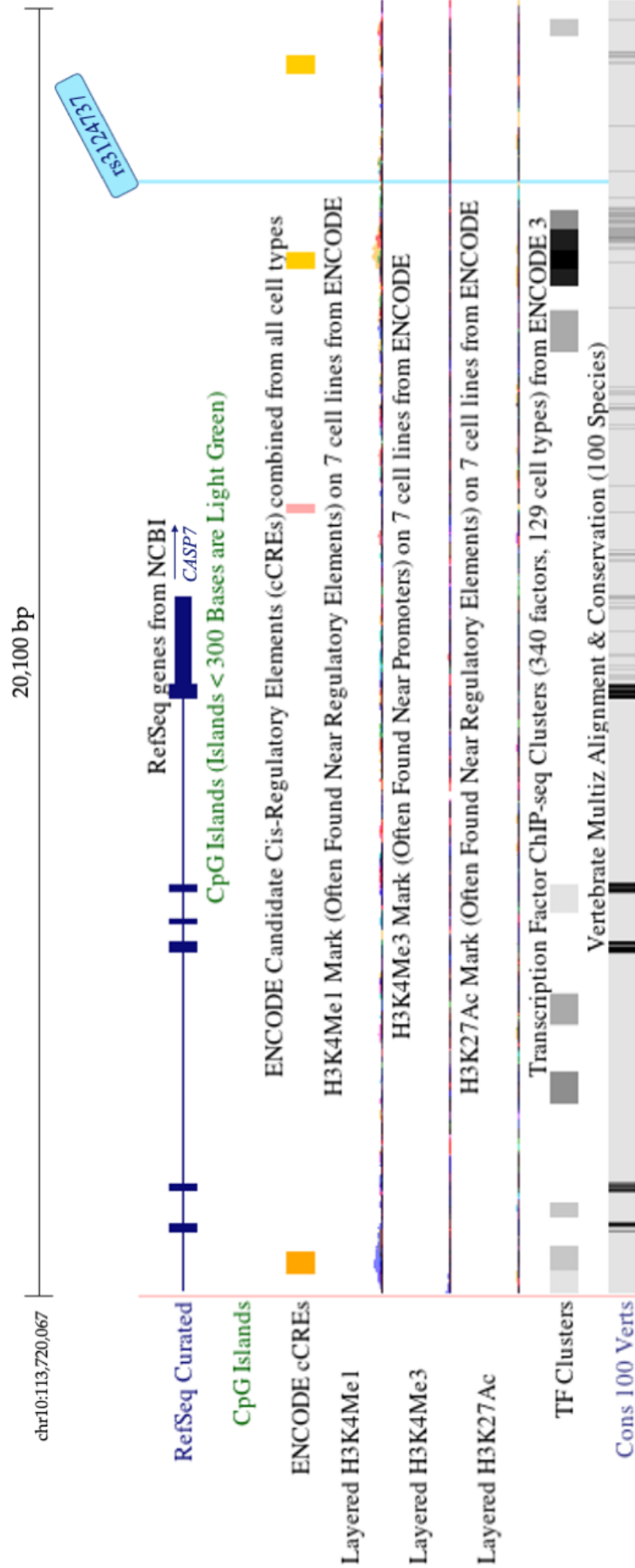


Figure 5.7: UCSC Genome Browser view of *CASP7* gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.4. Analysis of *PNMT* associated SNP

Phenylethanolamine N-methyltransferase (*PNMT*) is a gene located at 17q12 and contains 5 exons. *PNMT* has high levels of expression in adrenal glands and low expression in fat, where it methylates norepinephrine to form epinephrine. One SNP was identified from GWAS pathway analysis, and no SNPs were identified within the selection criteria ($\text{RegulomeDB} \leq 3$; $R^2 > 0.8$) through proxy analysis. The full *in silico* analysis of rs876493 is outlined in Table 5.5. Rs876493 was selected for functional analysis as it was present in a CpG island and with an H3K27ac histone mark in the pancreas. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.8.

Table 5.5: Full in silico analysis of SNPs associated with the gene PNMT. Underlined SNPs were selected for functional analysis.

Gene	Marker	Chr	Position(HG19)	Alleles Ref Alt	RegDB	Chromatin State	Pancreas	Location/ Function	H3K27ac	CpG Island	eQTL (NES)
PNMT	rs876493#	7	37824545	G A	4	Enhancers		5'UTR	Yes	Yes	↓ -0.24, Heart

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG Island. eQTL includes NES and tissue. #Highlights GWAS pathway analysis SNPs.

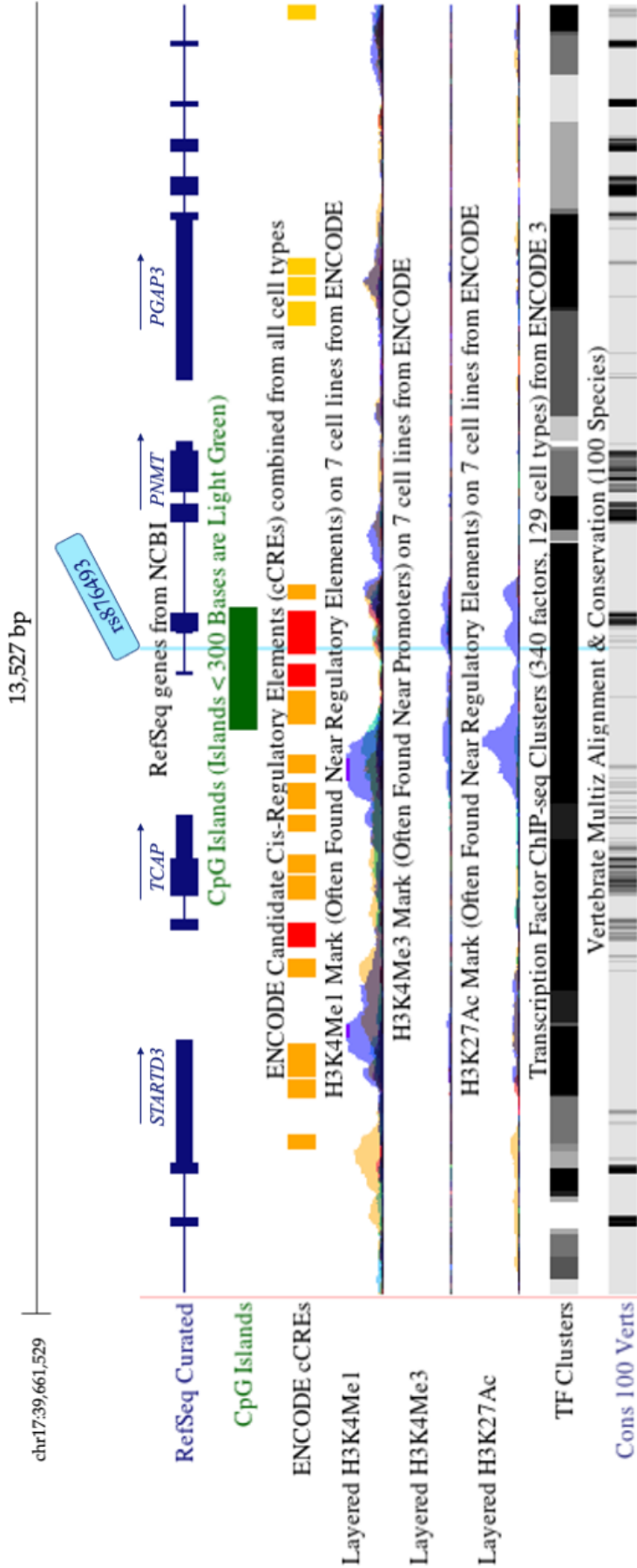


Figure 5.8: UCSC Genome Browser view of PNMT gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.5. Analysis of GRP associated SNP

Gastrin releasing peptide (*GRP*) is a gene located at 18q21.32 and contains 5 exons. *GRP* has high levels of expression in testis, lung and prostate, and low expression in the pancreas, where it controls the release of gastrointestinal hormones, smooth muscle cell contraction, and epithelial cell proliferation. SNP rs57791062 identified in GWAS pathway analysis, one SNP identified in literature as a GWAS SNP, and an additional SNP was identified through proxy analysis are presented in an LD block ($R^2 > 0.8$; $n=3$) and highlighted in a heatmap (Figure 5.9). The full *in silico* analysis of SNPs is outlined in Table 5.6. No SNPs were selected for functional analysis, as no SNPs had a low regulomeDB score, were present in a CpG island or had histone marks in the pancreas. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.10.

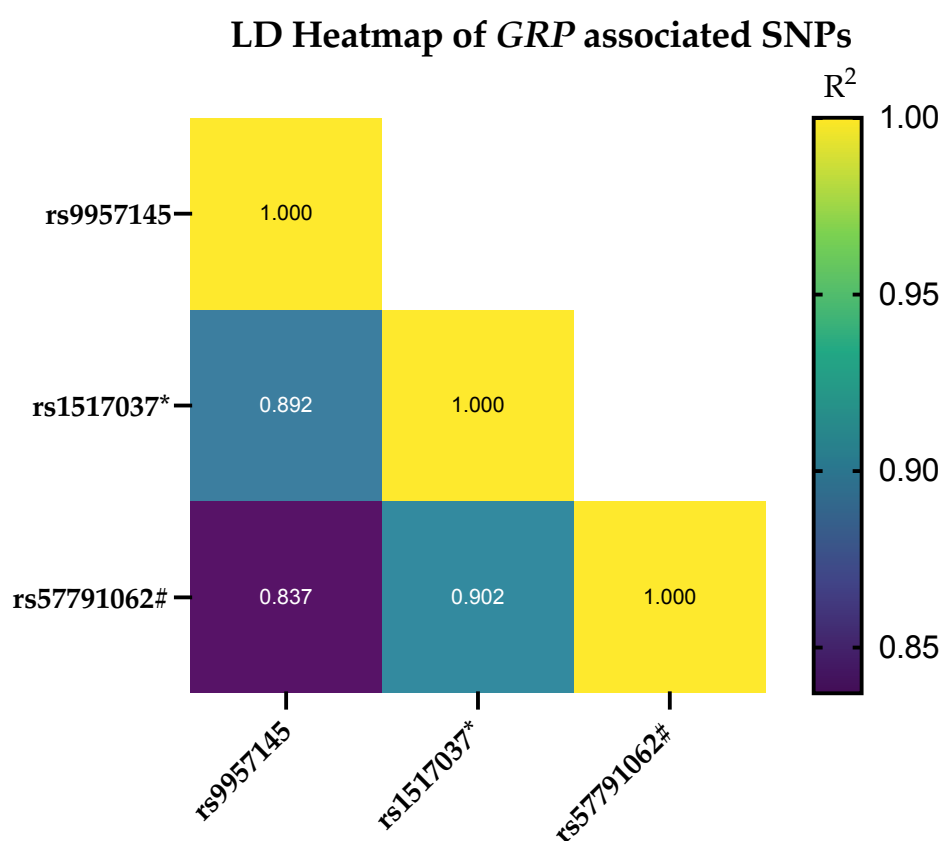


Figure 5.9: LD heatmap of *GRP* SNPs of interest based on R^2 values in European populations (LDLink). *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

Table 5.6: Full in silico analysis of SNPs associated with the gene GRP. *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function
GRP	rs1517037*	18	56878274	C T	5	Strong transcription	9.1 kb 5' of GRP
GRP	rs57791062#	18	56880211	C T	5	Strong transcription	7.2 kb 5' of GRP
GRP	rs9957145	18	56876228	G A	5	Strong transcription	11 kb 5' of GRP

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG Island. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

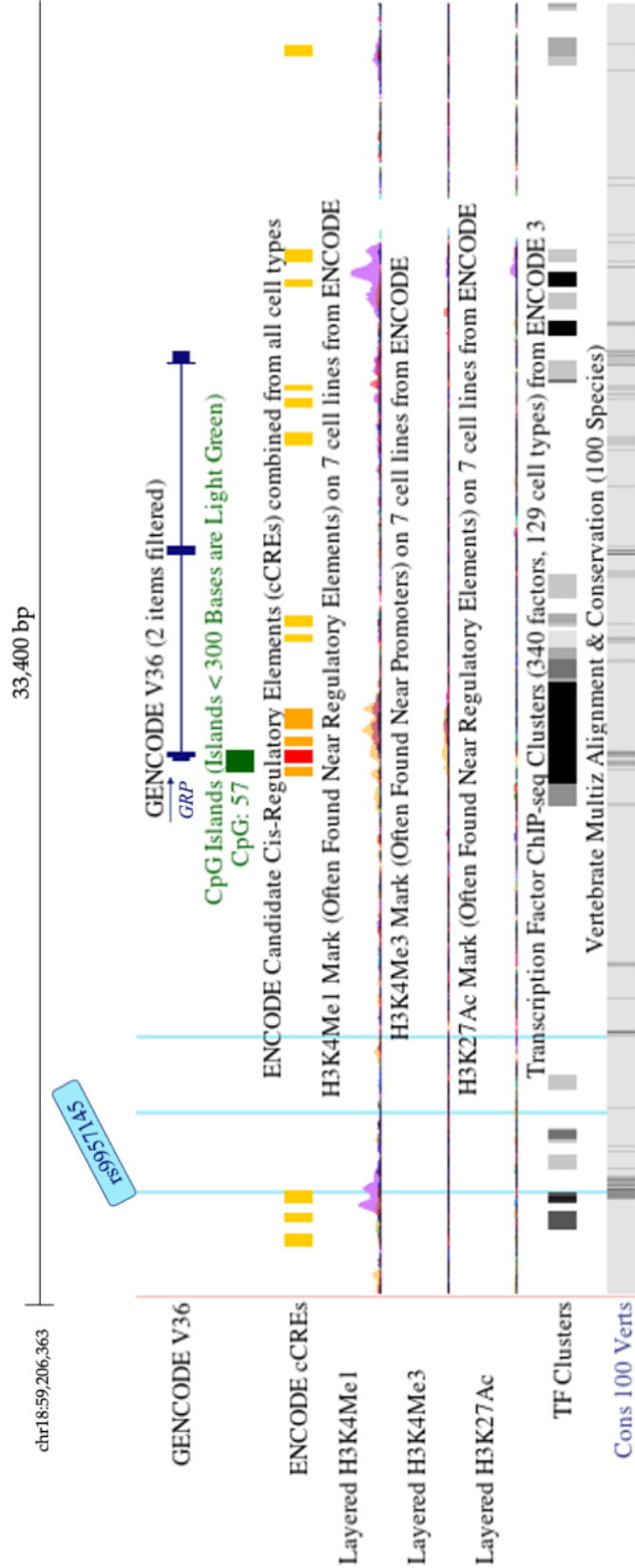


Figure 5.10: UCSC Genome Browser view of GRP gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.6. Analysis of *MED1* associated SNP

Mediator complex subunit 1 (*MED1*) is a gene located at 17q12 and it contains 17 exons. *MED1* has high levels of expression in a number of tissues, including, lymph nodes, appendix and endometrium, and low expression in the pancreas, where it acts with RNA polymerase II to control transcriptional regulation. One SNP was identified in literature searches, and 12 SNPs were identified through proxy analysis in are presented an LD block ($R^2 > 0.8$; $n=13$) and highlighted in a heatmap (Figure 5.11). The full *in silico* analysis of the SNPs is outlined in Table 5.7. No SNPs were selected for functional analysis. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.12.

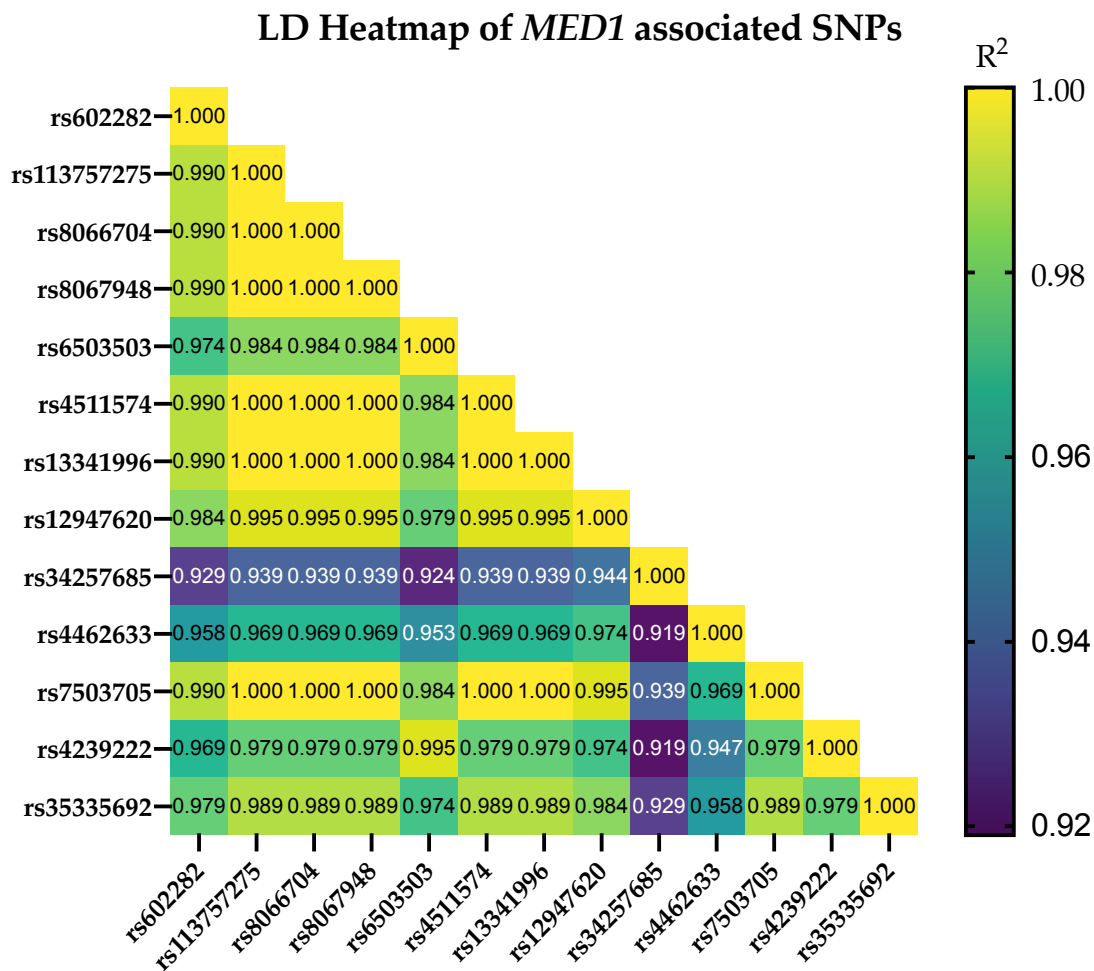


Figure 5.11: LD heatmap of *MED1* SNPs of interest based on R^2 values in European populations (LDLink).

Table 5.7: Full in silico analysis of SNPs associated with the gene *MEDI*.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function	Histone Marks	CpG Island	eQTL (NES)
<i>MEDI</i>	rs12947620	17	37605364	T C	4	Strong transcription	Intronic	H3K4me1 H3K27ac	—	↓ -0.068, Fibroblasts
<i>MEDI</i>	rs602282	17	37433021	C T	2c	Strong transcription	Intronic	—	—	↓ -0.066, Fibroblasts
<i>MEDI</i>	rs13757275	17	37462407	TTTT T(N=14)	3a	Strong transcription	Intronic	—	—	—
<i>MEDI</i>	rs8066704	17	37477218	C T	3a	Quiescent/Low	Intronic	—	—	↑ 0.067, Fibroblasts
<i>MEDI</i>	rs8067948	17	37477644	A G	3a	Quiescent/Low	Intronic	—	Yes	↑ 0.067, Fibroblasts
<i>MEDI</i>	rs6503503	17	37514373	G A	2a	Quiescent/Low	Intronic	—	—	↑ 0.067, Fibroblasts
<i>MEDI</i>	rs4511574	17	37532916	C T	2c	Quiescent/Low	Intronic	—	—	↑ 0.067, Fibroblasts
<i>MEDI</i>	rs13341996	17	37536862	C T	2b	Strong transcription	Intronic	—	—	↑ 0.067, Fibroblasts
<i>MEDI</i>	rs34257685	17	37621053	A C	3a	Enhancers	Intronic	—	—	—
<i>MEDI</i>	rs4462633	17	37641867	A C	3a	Strong transcription	Intronic	—	—	↓ -0.065, Fibroblasts
<i>MEDI</i>	rs7503705	17	37669704	A G	3a	Strong transcription	Intronic	—	—	↓ -0.065, Fibroblasts
<i>MEDI</i>	rs4239222	17	37696235	G T	3a	Strong transcription	—	—	—	↓ -0.063, Fibroblasts
<i>MEDI</i>	rs35335692	17	37710006	A G	2b	Strong transcription	—	—	—	↓ -0.066, Fibroblasts

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG Island. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

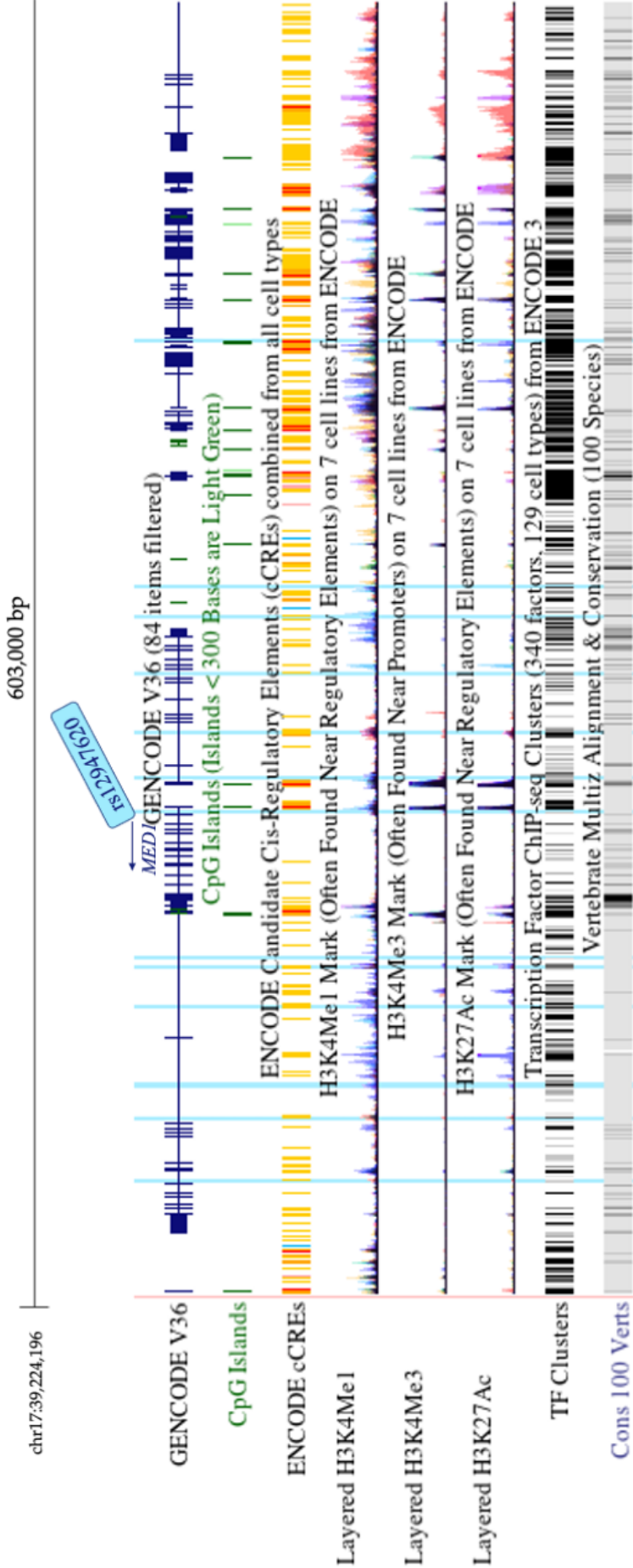


Figure 5.12: UCSC Genome Browser view of MED1 gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in seven cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.7. Analysis of SMC2 associated SNPs

Structural maintenance of chromosomes 2 (*SMC2*) is a gene located at 9q31.1 and it contains 29 exons. *SMC2* has high levels of expression in a number of tissues, including, lymph nodes, testis and appendix, and low expression in the pancreas, where it acts as central component of the condensin complex. One SNP was identified in GWAS pathway analysis, three SNPs were identified in literature searches, and 10 SNPs were identified through proxy analysis. These are presented in an LD block ($R^2 > 0.8$; $n=14$) and highlighted in a heatmap (Figure 5.13). The full *in silico* analysis of the SNPs is outlined in Table 5.8. SNPs rs4742901 and rs4742902 were selected for functional analysis, as both were in transcription start sites, CpG islands and had a H3K27ac histone mark in the pancreas. UCSC genome browser view of the region surrounding the SNPs is presented in Figure 5.14.

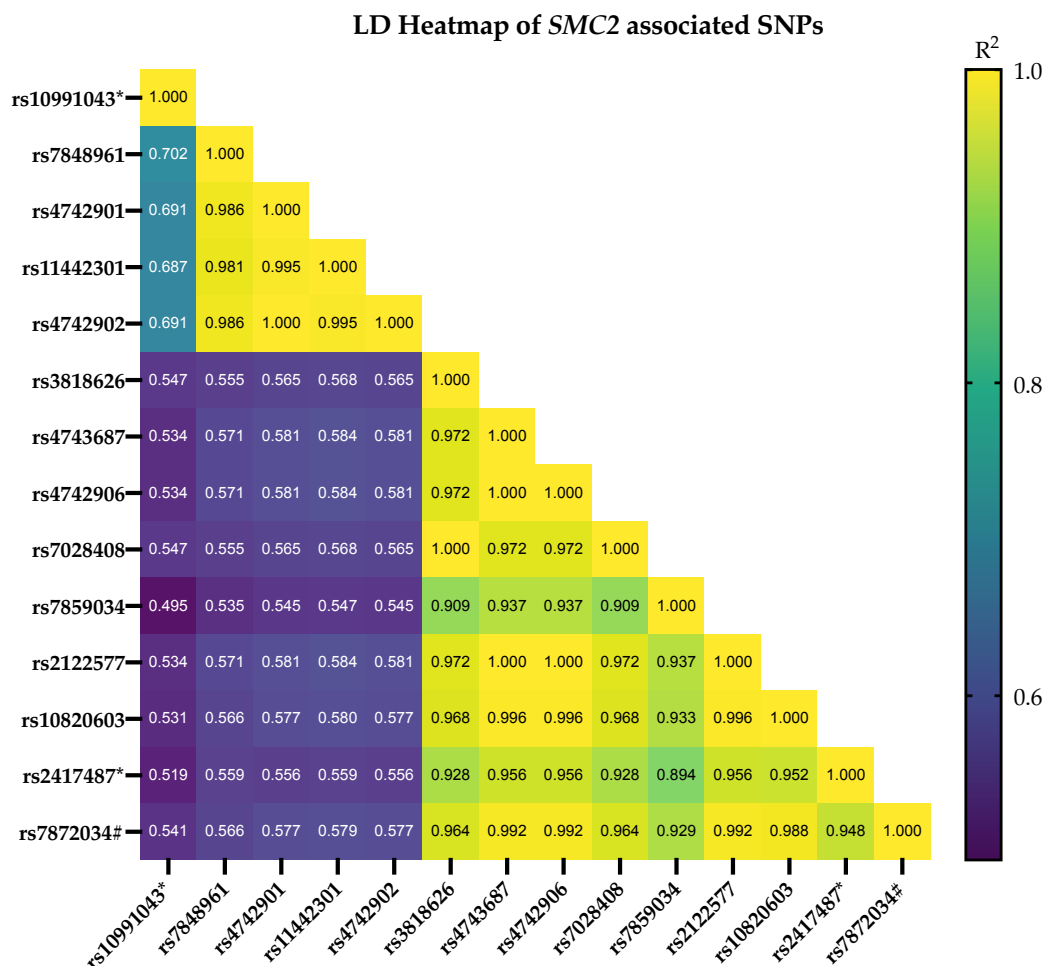


Figure 5.13: LD heatmap of SMC2 SNPs of interest based on R^2 values in European populations (LDLink). *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

Table 5.8: Full in silico analysis of SNPs associated with the gene *SMC2*. Underlined SNPs were selected for functional analysis

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt	RegDB	Chromatin State Pancreas	Location/ Function	H3K27ac	CpG Island	eQTL (NES)
<u>SMC2</u>	<u>rs4742901</u>	<u>9</u>	<u>106856243</u>	T C	<u>2b</u>	<u>Active TSS</u>	<u>497bp 5' of SMC2</u>	<u>H3K4me1</u> , <u>H3K4me3</u> <u>H3K27ac</u>	<u>Yes</u>	↑ 0.46, Thyroid
SMC2	rs1442301	9	106856145	A AG	2b	Active TSS	397bp 5' of SMC2	H3K27ac	Yes	↑ 0.47, Thyroid
SMC2	rs7848961	9	106852661	G A	3a	Enhancers	3.9kb 5' of SMC2	—	—	↑ 0.47, Thyroid
SMC2	rs2122577	9	106866703	C T	2b	Strong transcription	Intronic	—	—	↓ -0.55, Heart
SMC2	rs10820603	9	106877939	G A	6	Strong transcription	Intronic	—	—	↓ -0.55, Heart
SMC2	rs7872034	9	106896809	G A	1f	Strong transcription	Synonymous	H3K27ac	—	↓ -0.54, Heart
SMC2	rs4743687	9	106856910	C T	4	Active TSS	Intronic	H3K27ac	Yes	↓ -0.54, Heart
SMC2	rs4742906	9	106857078	A G	4	Active TSS	5'UTR	H3K27ac	—	↓ -0.55, Heart
SMC2	rs7028408	9	106859811	G A	7	Strong transcription	Intronic	—	—	↓ -0.55, Heart
<u>SMC2</u>	<u>rs4742902</u>	<u>9</u>	<u>106856043</u>	A G	<u>2b</u>	<u>Active TSS</u>	<u>297bp 5' of SMC2</u>	<u>H3K4me3</u> , <u>H3K27ac</u>	<u>Yes</u>	↑ 0.47, Thyroid
SMC2	rs3818626	9	106856633	C T	2b	Active TSS	5'UTR	H3K27ac	Yes	↓ -0.53, Heart
SMC2	rs2417487*	9	106887581	A G	6	Strong transcription	Intronic	—	—	↓ -0.54, Heart
SMC2	rs10991043*	9	106797388	T C	6	Quiescent/Low	59kb 5' of SMC2	—	—	↑ 0.43, Thyroid
SMC2	rs7859034#	9	106865692	G T	7	Strong transcription	Intronic	—	—	↓ -0.55, Heart

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers and if SNP is present in a CpG Island. eQTL includes NES and tissue *Highlights GWAS SNPs. #Highlights GWAS pathway analysis SNPs.

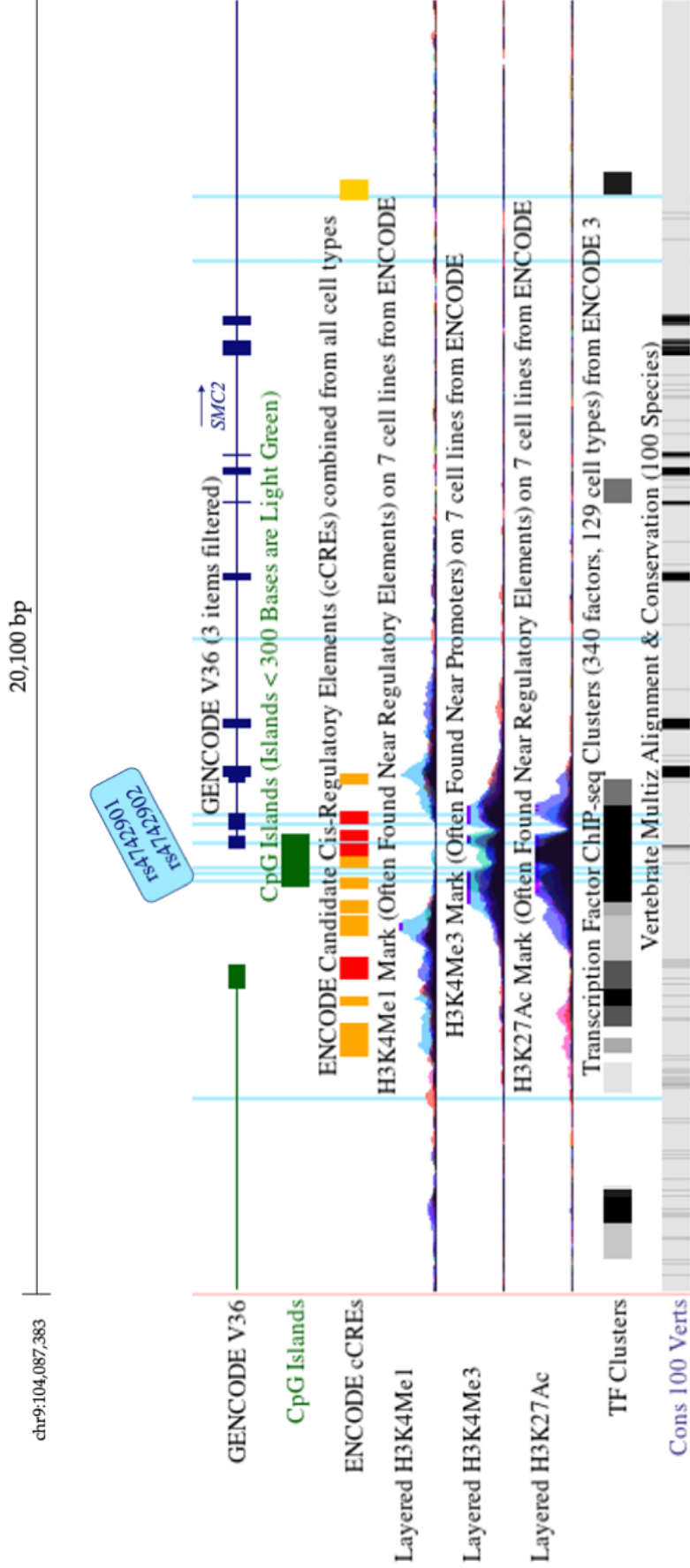


Figure 5.14: UCSC Genome Browser view of SMC2 gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in 7 cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.8. Analysis of *THRA* associated SNP

Thyroid hormone receptor alpha (*THRA*) is a gene located at 17q21.1 and contains 11 exons. *THRA* has high levels of expression in the brain and ovary, and low expression in the pancreas, nuclear hormone receptor for triiodothyronine. One SNP was identified in GWAS pathway analysis and no SNPs were within the selection criteria ($\text{RegulomeDB} \leq 3$; $R^2 > 0.8$) were identified through proxy analysis. The full *in silico* analysis of the SNP is outlined in Table 5.9. The SNP was not selected for functional analysis. Overview of the potential regulatory tracks of this SNPs are shown in the UCSC genome browser view in Figure 5.15.

Table 5.9: Full in silico analysis of SNPs associated with the gene THRA.

Gene	Marker	Chr	Position (HG19)	Alleles Ref Alt		RegDB	Chromatin State Pancreas	Location/ Function	eQTL (NES)
THRA	rs8078692	17	38215117	A	G	6	Strong transcription	—	↓ -0.063, Oesophagus

This table includes gene name, SNP marker, HG19 chromosome position, reference and alternative alleles, regulomeDB score (RegDB), chromatin state in the pancreas, the function of the SNP, histone markers. eQTL includes NES and tissue.

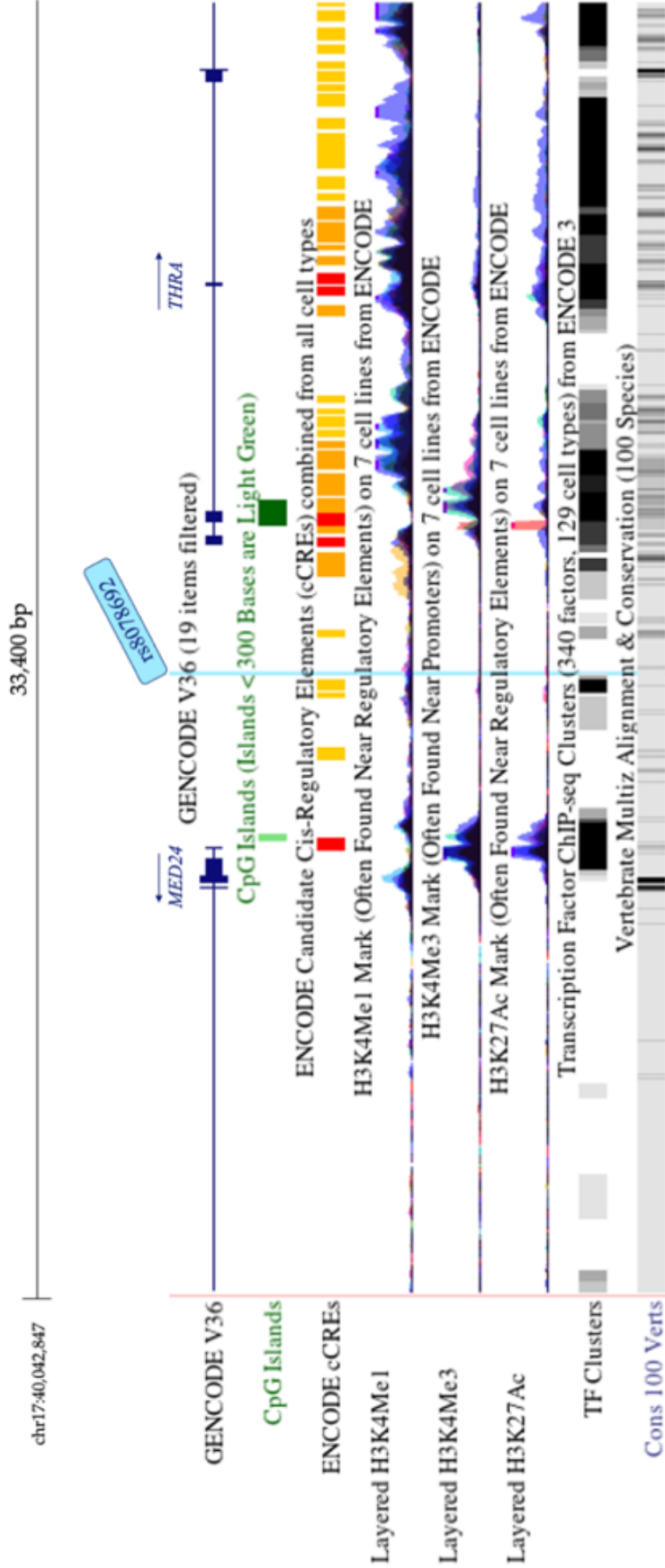


Figure 5.15: UCSC Genome Browser view of THRA gene regions surrounding SNP of interest (highlighted in light blue) with selected ENCODE regulation tracks and conservation among 100 vertebrates. H3K27ac (enhancer), H3K4Me1 (enhancer) and H3K4Me3 (promoters) in 7 cell lines (non-PDAC/pancreas). CpG islands are indicative of a promoter region and ENCODE candidate Cis-Regulatory Elements from all cell types. ChIP-seq signals of transcription factors from ENCODE.

5.2.1.9. SNPs selected for functional validation

In total, six SNPs from four genes were selected for functional validation. Two SNPs from gene *SMC2* (rs4742901 and rs4742902) and *CASP7* SNP (rs3124737) were successfully cloned and edited using site directed mutagenesis. *PNMT* SNP (rs876493) was not brought forward for dual luciferase reporter assays as attempts to optimise PCR cloning of the region of interest were unsuccessful. Optimisation included the use of three different primer pairs, and extension times of 30 seconds and one minute. No colonies were observed following three attempts to perform site directed mutagenesis of *BCL2L1* SNPs (rs11378324 and rs2241845) therefore these SNPs not brought forward for dual luciferase reporter assays (Table 5.10).

Table 5.10: Final list of SNPs selected for functional validation, including nearest gene, marker, HG19 position, reference and alternative SNP and reasons for exclusion.

Gene	Marker	Chr	Position (HG19)	Ref	Alt	Reasons for exclusion
<i>BCL2L1I</i>	rs2241845	2	111879100	A	G	Site directed mutagenesis unsuccessful
<i>BCL2L1I</i>	rs11378324	2	111877346	G	GGGGGG	Site directed mutagenesis unsuccessful
<i>PNMT</i>	rs876493	7	37824545	G	A	Unable to clone region of interest
<i>CASP7</i>	rs3124737	10	115497095	A	G	Included
<i>SMC2</i>	rs4742901	9	106856243	T	C	Included
<i>SMC2</i>	rs4742902	9	106856243	A	G	Included

5.2.2. Role of CASP7 in PDAC

5.2.2.1. Functional validation of SNPs associated with CASP7

In a PDAC GWAS pathway analysis previously conducted by our group a candidate risk allele rs3124737-G in close proximity to *CASP7* in the Pujana ATM PCC Network was gene set was identified. An eQTL was identified between rs3124737-G and *CASP7* where the risk allele resulted in an increased expression in two normal pancreas tissue databases (GTEx $P=2.5E-08$; $\beta=0.48$) and (LTG $P=0.02$; $\beta=0.28$) [122]. The eQTL effect of the alternative allele of SNP rs3124737-G was associated with higher expression of the *CASP7* gene in normal pancreatic tissue using GTEx (Figure 5.16. A).

A genomic region containing the selected SNP rs3124737 (*CASP7*) was cloned into a luciferase reporter vector to test the effect of each allelic variant on transcription in a dual luciferase reporter assay, as per methodology described in Section 2.13. The SNP rs3124737-G influenced the reporter expression level in both established pancreatic cell lines PANC1 and ASPC1, but not in the cell line HEK293T. The dual luciferase reporter assay showed an increased activity of the G allele at rs3124737 relative to the A allele in PANC1 and ASPC1 (Figure 5.16. B).

HOCOMOCO-v11 was used to generate PWM for changes in transcription factor binding due to the presence of the SNP. The presence of the alternative allele (G) is predicted to result in a loss of binding of MYBB (41.36-fold, $p=3.23E-03$) and ONEC2 (9.43-fold, $2.66E-03$). Whereas rs3124737-G is predicted to result in increased binding of transcription factors E2F8, KLF1 and BRCA, of 23.87-fold ($P=4.59E-04$), 20.6-fold, ($P=3.60E-04$) and 6.74-fold ($P=3.95E-04$) respectively (Figure 5.16 C).

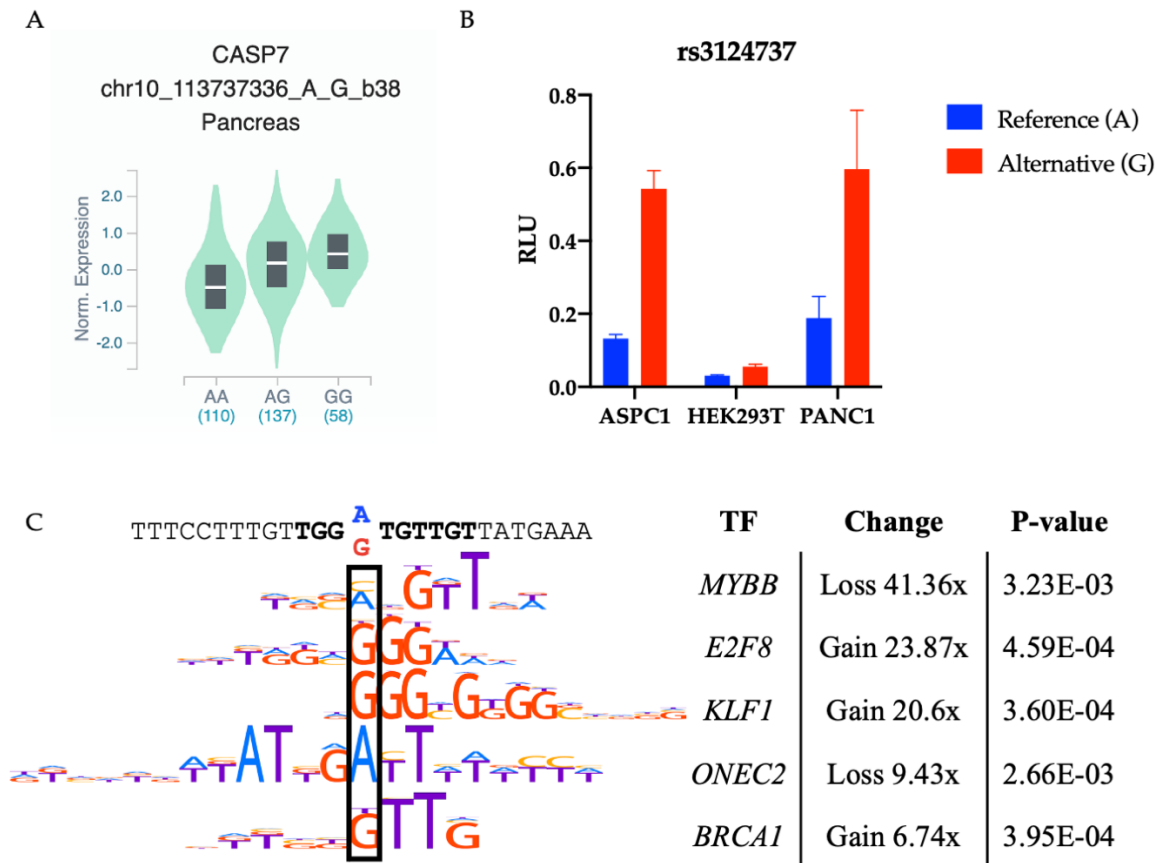


Figure 5.16: (A) GTEX – violin plot of expression of different alleles in normal pancreas. (B) Effects of the candidate function SNPs on gene expression in a dual luciferase reporter assay. Two allelic variants of putative enhancers containing each SNP were cloned upstream of the Firefly luciferase gene. A panel of cell lines – HEK293T, ASPC1 and PANC1 – were transfected with reporter constructs together with Renilla. Luminescence was measured 24 hours after transfection using Dual-Glo Luciferase Assay System (Promega). Luciferase expression was normalised to Renilla and made relative to p2Luc plasmid with no insertion. Blue boxes – reference allele, red boxes – alternative allele. (C). DNA sequence surrounding rs3124737 and ChIP-seq based motif logos for MYBB, E2F8, KLF1, ONEC2 and BRCA1 according to HOCOMOCO-vII [285]. A to G substitution alters motif P-value (fold-change is given according to PERFECTOS-APE [286]).

5.2.2.2. Role of CASP7 in PDAC

CASP7 is critical in apoptosis induction, acting as a candidate for susceptibility to insulin-dependent diabetes. Inactivating mutations in CASP7 have been reported to contribute to the pathogenesis of some human solid cancers; therefore, CASP7 was identified as a candidate for further exploration and validation. IHC was performed to assess the expression of CASP7 in a cohort of pancreatic cancer patients. The protein and mRNA expression of CASP7 was also assessed in established pancreatic cancer cell lines through Western blotting and RT-qPCR, and in publicly available data.

5.2.2.3. Immunohistochemistry of CASP7

FFPE tumour specimens from 49 patients and 3 normal pancreas tissue specimens were analysed by means of IHC for expression of CASP7. Samples were stained with a CASP7 antibody recognising both the precursor and active forms of the protein, were scored and presented as positive or negative staining. Negative staining was observed in 33 (67%), and positive staining was observed in 16 (33%) of the samples, with representative images in Figure 5.17. Information on patient age at diagnosis was available for all 49 cases; ages ranged from 38 to 81 years and the median age was 64 years. Average survival was 25.9 months, with 71% of the patients deceased at the time of latest available follow up (31/01/2017). Tumour type for all cases was ductal, one case with tumour invading the colon, and another invading the intestine. 13 tumours were poorly differentiated, 31 were moderately differentiation and 5 were well differentiated. Relevant histopathological information was available for all cases (Table 5.11).

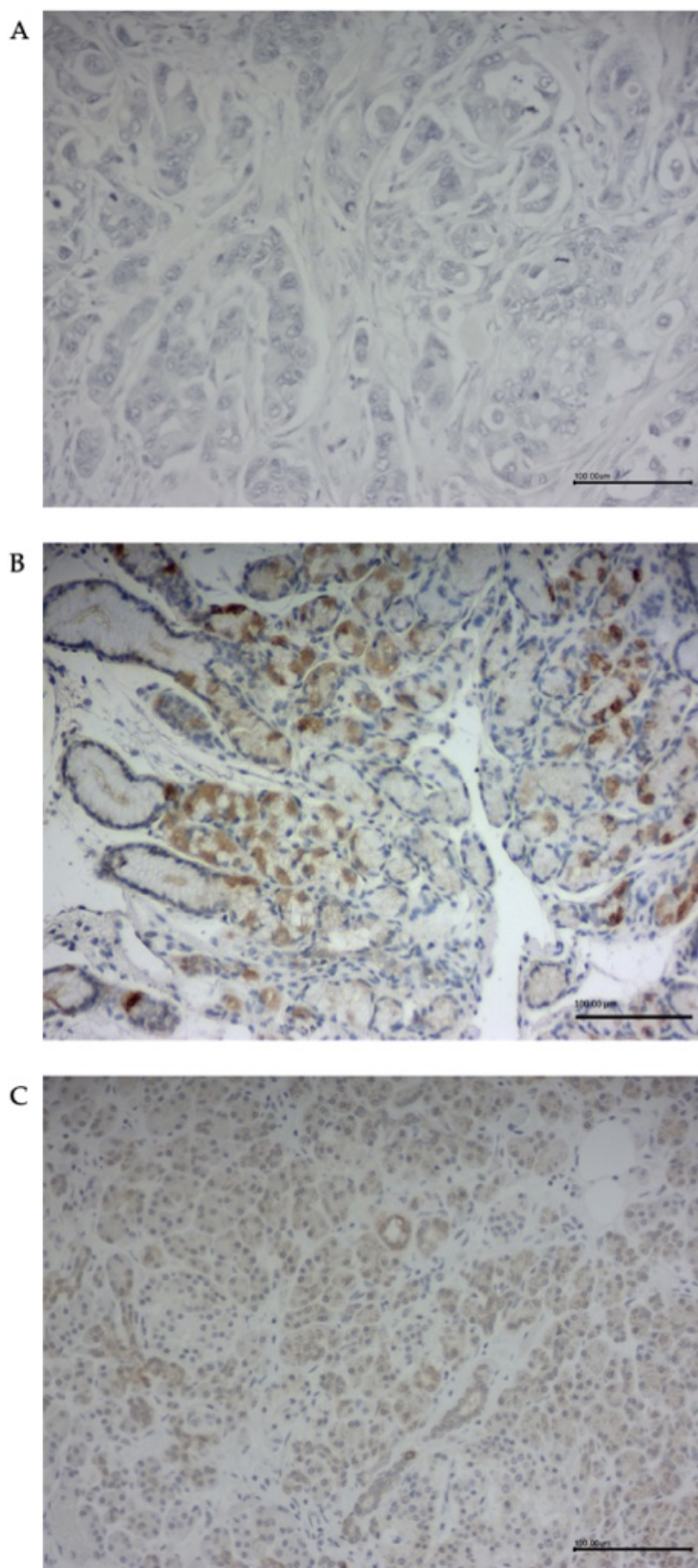


Figure 5.17: Representative images of CASP7 staining (A) PDAC tumour with negative expression (B) PDAC tumour with positive expression and (C) positive CASP7 staining in normal pancreas. Images taken at 200X, scale bar 100 μm.

Table 5.11: Patient information and statistics divided by patients with negative and positive CASP7 staining.

	Negative Staining n=33 (67%)	Positive Staining n=16 (33%)	χ^2 P-value
Gender*			0.7541
Male	20 (60.61 %)	11 (68.75 %)	
Female	13 (39.39 %)	5 (31.25 %)	
Age at surgery (median)	60	69	—
T stage**			0.2578
1	3 (9.09 %)	0 (0.00 %)	
2	2 (6.06 %)	1 (6.25 %)	
3	28 (84.85 %)	15 (93.75 %)	
N stage**			0.3354
0	8 (24.24 %)	6 (37.50 %)	
1	25 (75.76 %)	10 (62.50 %)	
R0 (>1mm)*			>0.9999
Yes	32 (96.97 %)	16 (100.00 %)	
No	1 (3.03 %)	0 (0.00 %)	
Differentiation**			0.0003
Poor	8 (24.24 %)	5 (31.25 %)	
Moderate	22 (66.67 %)	2 (12.50 %)	
Well	3 (9.09 %)	9 (56.25 %)	
Max pathological axis	2.98 cm	2.91 cm	—
Lymphatic invasion*			0.0032
Absent	2 (6.06 %)	7 (43.75 %)	
Present	31 (93.94 %)	9 (56.25 %)	
Perineural invasion*			>0.9999
Absent	2 (6.06 %)	0 (0.00 %)	
Present	31 (93.94 %)	16 (100.00 %)	
Portal vein involvement*			>0.9999
No	32 (96.97 %)	15 (93.75 %)	
Yes	1 (3.03 %)	1 (6.25 %)	
Survival Status*			0.5010
Dead	25 (75.76 %)	10 (62.50 %)	
Alive	8 (24.24 %)	6 (37.50 %)	
Survival (Months)	24.2	29.4	—

χ^2 – *Fisher's exact test; ** Chi-squared.

Kaplan-Meier survival (Figure 5.18) analysis revealed that patients with CASP7 expression have significantly longer (P-value=0.015, log rank) survival times than patients with no CASP7 expression. Patients with no expression have a median survival of 24.15 months, while patients with positive CASP7 expression have a median survival of 29.43 months. Positive (medium to high) protein expression of CASP7 was observed in the three normal pancreas samples included for staining.

CASP7 Expression Survival Analysis

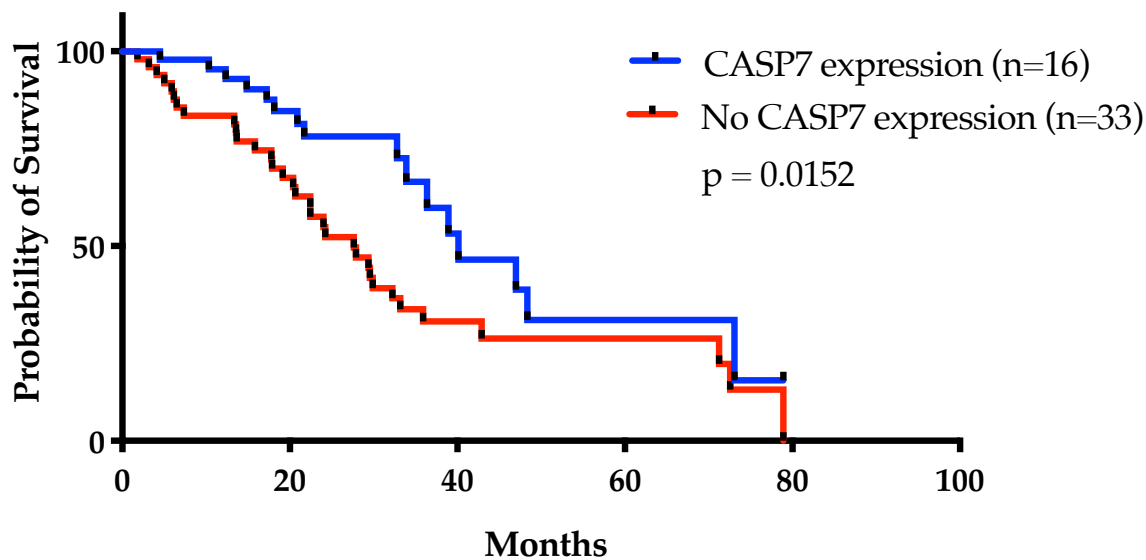


Figure 5.18: Kaplan-Meier survival analysis of PDAC patients with CASP7 expression (blue) compared to PDAC patients without CASP7 expression (red).

5.2.2.4. CASP7 expression in PDAC established cell lines

RT-qPCR was performed to assess the mRNA levels of CASP7 in a panel of PDAC cell lines (Figure 5.19) and in a subset of tumour versus adjacent normal PDAC samples (Figure 5.20). In the cell line panel, all cell lines developed from the primary cancer in the pancreas (blue, Figure 5.19) have increased expression of CASP7 compared to the transformed normal epithelial cell line H6c7 (red, Figure 5.19). Metastatic cell lines (green, Figure 5.19) have lower expression than the cell lines established from the primary tumour.

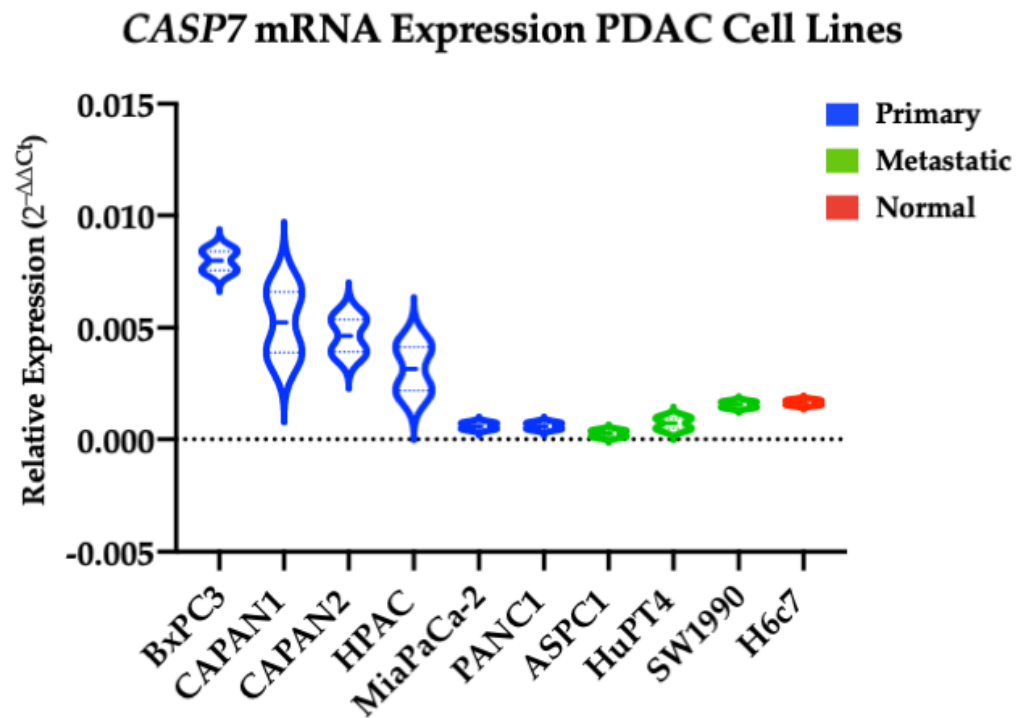


Figure 5.19: Relative quantification of ($2^{-\Delta\Delta C_t}$) of CASP7 mRNA expression in transformed normal pancreas (red) (H6c7), established cell lines derived from primary cancer site (blue) (BxPC3, CAPAN1, CAPAN2, HPAC, MiaPaCa-2, PANC1) and metastatic cancer sites (green) (SW1990, HuPT4, ASPC1). GAPDH used as endogenous control.

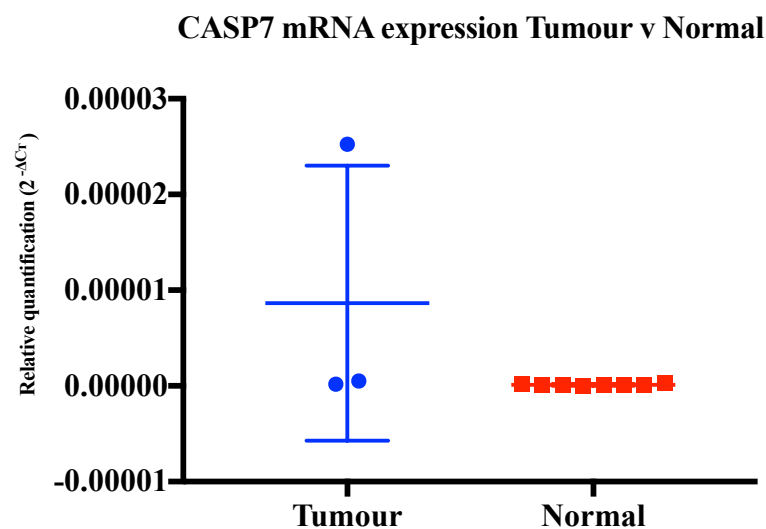


Figure 5.20: Relative quantification of ($2^{-\Delta\Delta C_t}$) of CASP7 mRNA expression in PDAC tumour (blue) (n=3) vs. adjacent normal (red) (n=8) panel. 18S used as endogenous control.

Protein levels of CASP7 in established PDAC cell lines were also assessed. Western blot analysis of PDAC cell line panel indicated that there is varied expression of full length CASP7 in cell lines. High levels of CASP7 expression were observed in PDAC established cell lines BxPC3, Capan1, Capan2, and in H6c7, a transformed immortalised normal pancreas cell line. Low levels of expression were observed in AsPC1, HuPT4, MiaPaCa-2, PANC1, PT213, and no detectable expression was observed in HPAC, SW1990 and Fib102, a fibroblast cell line derived from tumour-adjacent tissue. Non-specific binding was observed at a number of molecular weights (Figure 5.21).

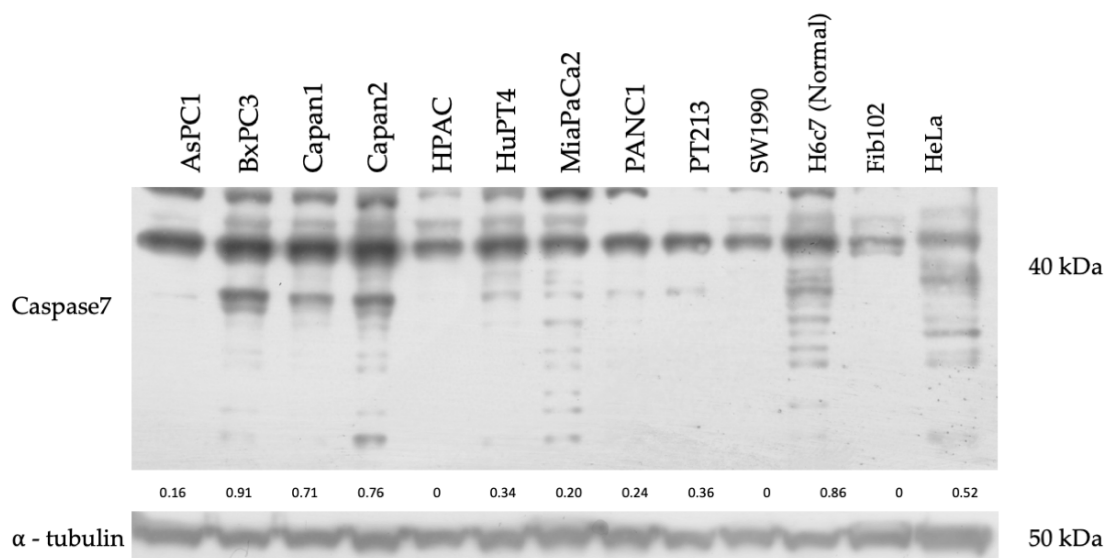


Figure 5.21: Western blot analysis for the expression of full length 7 CASP7 in 9 PDAC established cell lines (AsPC1, BxPC3, Capan1, Capan2, HPAC, HuPT4, MiaPaCa-2, PANC1, and SW1990), primary PDAC line (PT213), fibroblast line derived from tumour-adjacent tissue (Fib102), transformed normal cell line (H6c7) and positive control (HeLa) ImageJ software was used to perform densitometry.

5.2.2.5. In silico analysis of CASP7 expression in publicly available datasets

The expression of CASP7 in PDAC and normal pancreas tissue was assessed using publicly available datasets. GEPIA2 was used to assess mRNA expression in different tumour and normal tissues across all cancer types (Figure 5.22 A). The effect of the expression of CASP7 on the survival of pancreatic cancer patients was also plotted using GEPIA2 (Figure 5.22 B). GEPIA2 was also used to plot RNA

expression levels during pancreatic cancer stages I-IV (Figure 5.22 C). Data from www.proteinatlas.org indicates that staining of CASP7 is medium in normal samples, and no staining was detected in 58% of PDAC tumours tested.

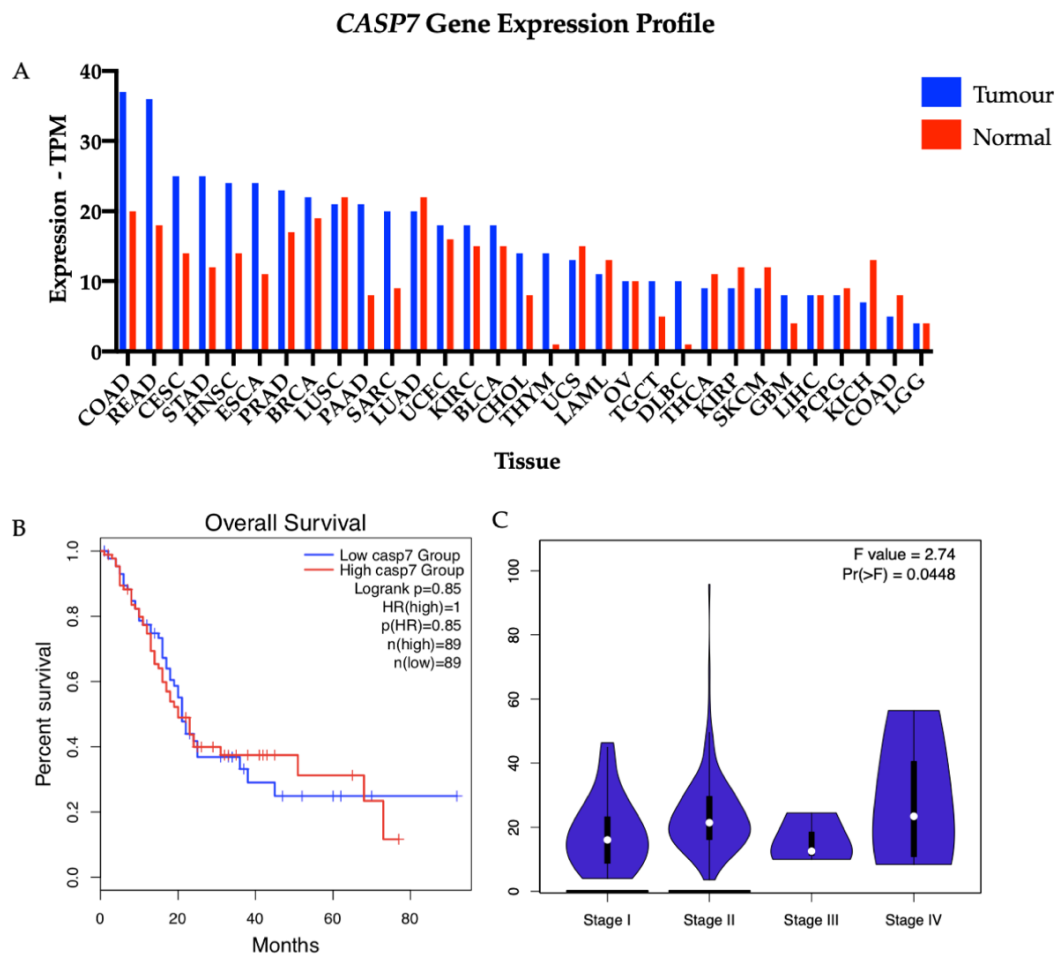


Figure 5.22: The GEPIA2 database was used to examine the expression of CASP7 in pancreatic cancer and the normal pancreas. Plots were generated using data from TCGA for pancreatic cancer and GTEx for normal pancreas. (A) mRNA expression in different tumour and normal tissues, TCGA tumour abbreviations used. (B) The effect of CASP7 expression on the overall survival of pancreatic cancer. (C) RNA expression levels during pancreatic cancer stages I-IV.

5.2.2.6. Functional assessment of the effect of SNPs rs4742902 and rs4742901 on gene transcription

SNPs rs4742901 and rs4742902 act as eQTLs in multiple non-pancreas tissues in GTEx whereby the alternative reference alleles result in increased expression of *SMC2* gene. Figure 5.23 and Figure 5.24 presents the effect size in various tissues of rs4742901 and rs4742902 allelic configuration on *SMC2* gene expression. The SNPs are 200 bp apart, upstream of *SMC2* and occur in a CpG island, and are in high LD ($R^2 \geq 0.8$, LD Link) with rs7859034, the top gene signal in the Pujana ATM PCC Network in the PDAC GWAS pathway analysis previously conducted by our group [122].

The dual luciferase reporter assay was performed using established PDAC cell lines ASPC1 and PANC1, and HEK293T cells for *SMC2* SNPs rs4742901, and rs4742902. No effect was observed for rs4742901 (Figure 5.25 A). In rs4742901, HOCOMOCO-vII PWM analysis shows that the presence of the alternative allele (C) is predicted to result in a gain of binding of MXI1-long, MXI1, BHE, MYC and HES5 by 49.56-fold ($P=1.29E-04$), 38.84-fold ($P=4.83E-04$), 27.62-fold ($P=3.27E-04$), 22.71-fold ($P=3.60E-04$) and 14.64-fold ($P=2.44E-04$) respectively (Figure 5.25 B).

rs4742902 results in increased luciferase expression compared to reference allele in HEK293T cells, but no difference observed in PANC1 and ASPC1 cells (Figure 5.26 A). PWM analysis for rs4742902 showed that the presence of the alternative allele (G) is predicted to result in a loss of binding of RUNX2, RUNX1, RUN3, and ZN563 by 35.25-fold ($P=3.30E-05$), 26.34-fold ($P=6.80E-05$), 17.03-fold ($P=1.06E-04$) and 12.65-fold ($P=4.60E-05$) respectively. The alternative SNP (G) is also is predicted to result in a gain of 6.39-fold ($P=8.80E-05$) of transcription factor ARNT2 (Figure 5.26 B).

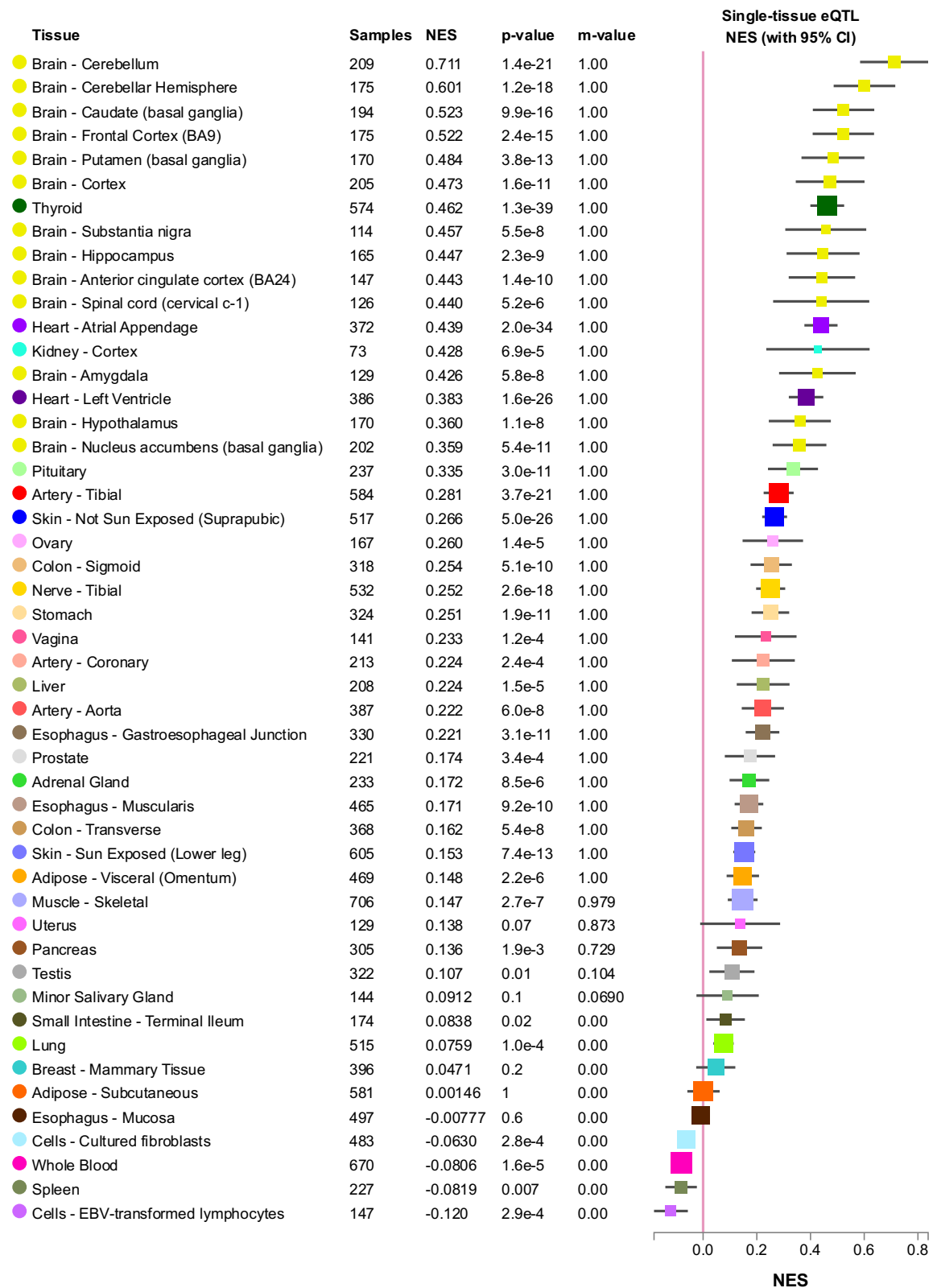


Figure 5.23: eQTL prediction of top signals in normal tissues using GTEx for *rs4742901*. Normalised effect size (NES) is the effect of the alternative allele relative to the reference allele. M-value posterior probability that an effect exists in a tissue. eQTL expression quantitative trait loci- identification of genetic associations with gene expression level.

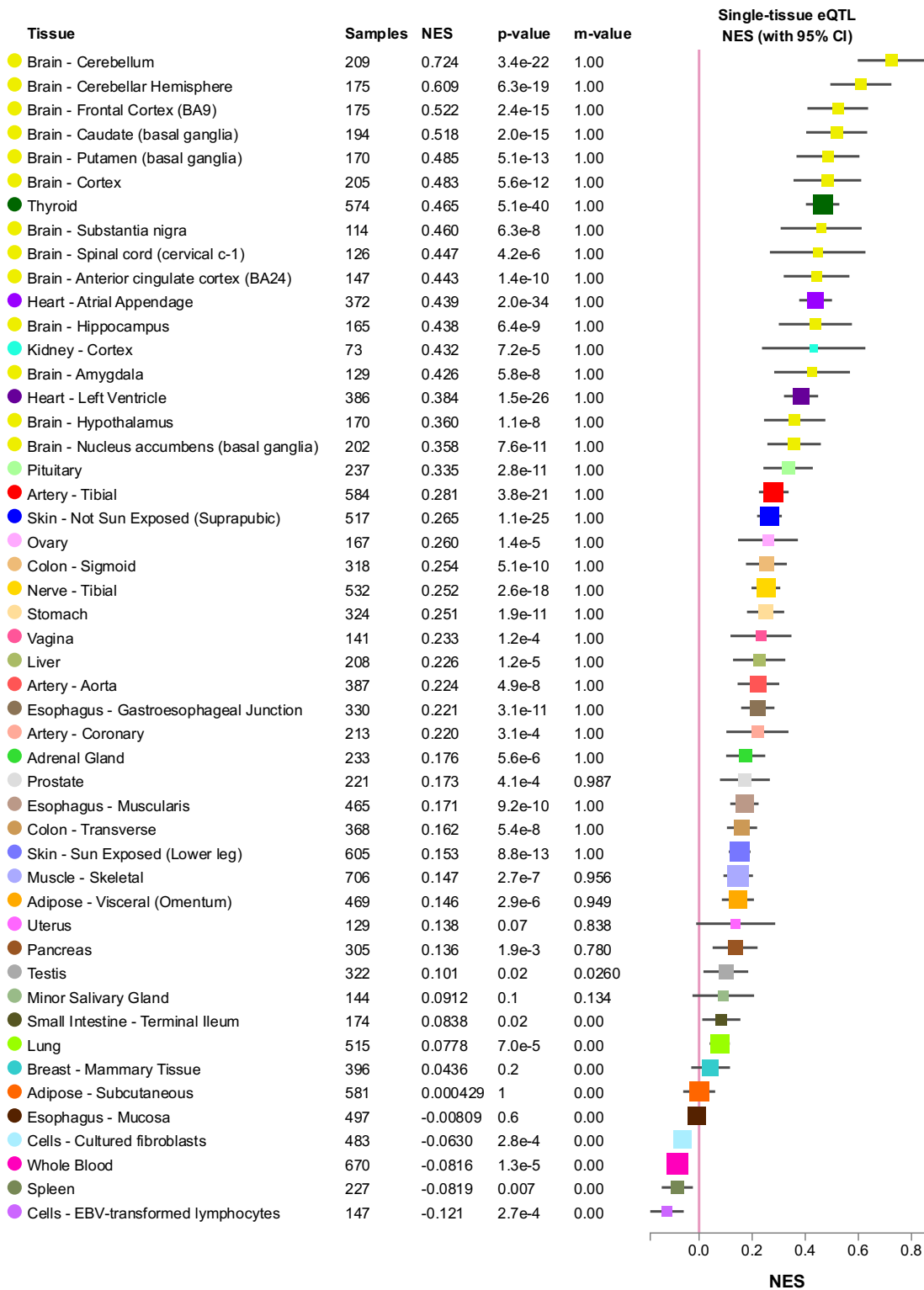


Figure 5.24: eQTL prediction of top signals in normal tissues using GTEx for rs4742902. Normalised effect size (NES) is the effect of the alternative allele relative to the reference allele. M-value posterior probability that an effect exists in a tissue. eQTL expression quantitative trait loci- identification of genetic associations with gene expression level.

A

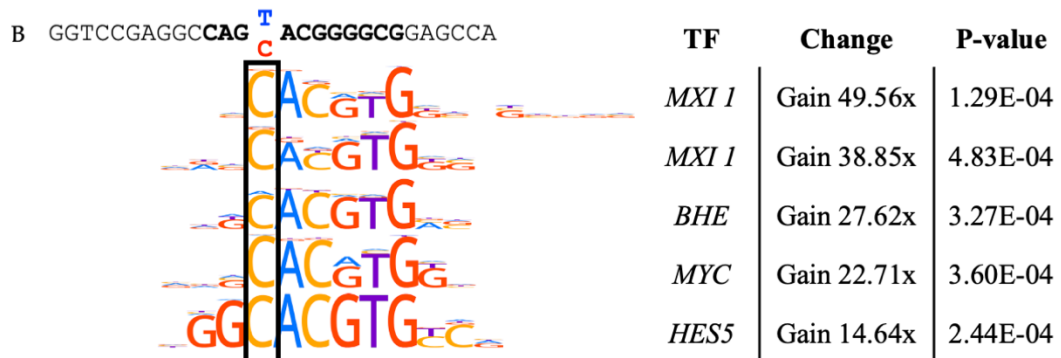
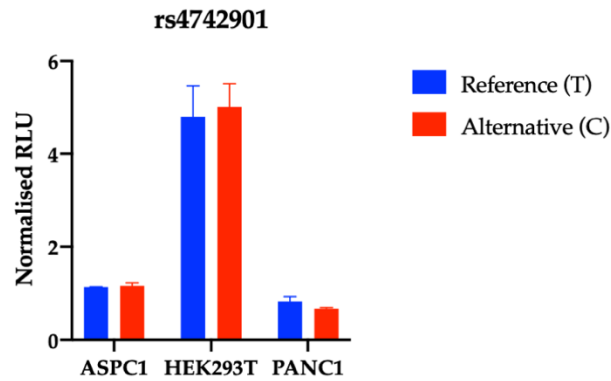


Figure 5.25: (A) Effects of the candidate SNP rs4742901 on gene expression in a dual luciferase reporter assay. Two allelic variants of 5'UTR regions containing each allele variant were cloned upstream of the Firefly luciferase gene. A panel of cell lines – HEK293T, ASPC1 and PANC1 – were transfected with reporter constructs together with Renilla. Luminescence was measured 24 hours after transfection using Dual-Glo Luciferase Assay System (Promega). Luciferase expression was normalised to Renilla and made relative to p2Luc plasmid with no insertion. Blue boxes – reference allele, red boxes – alternative allele. (B) DNA sequence surrounding rs4742901 and ChIP-seq based motif logos for MXI1, BHE, MYC and HE5 according to HOCOMOCO-v11 [285]. T to C substitution alters motif P-value (fold-change is given according to PERFECTOS-APE [286]).

A

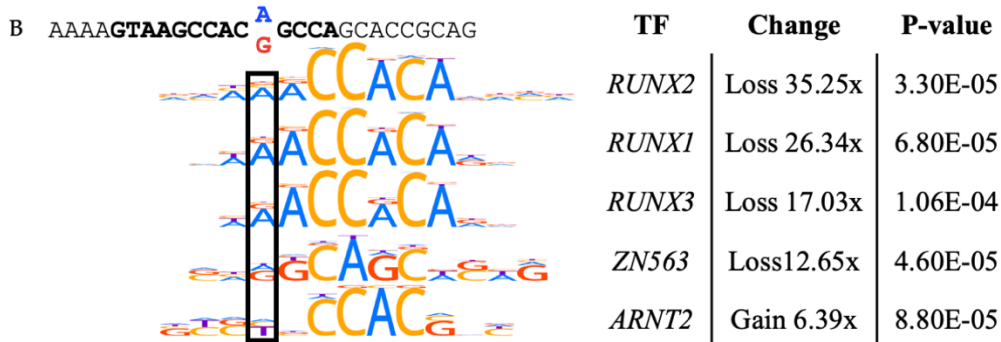
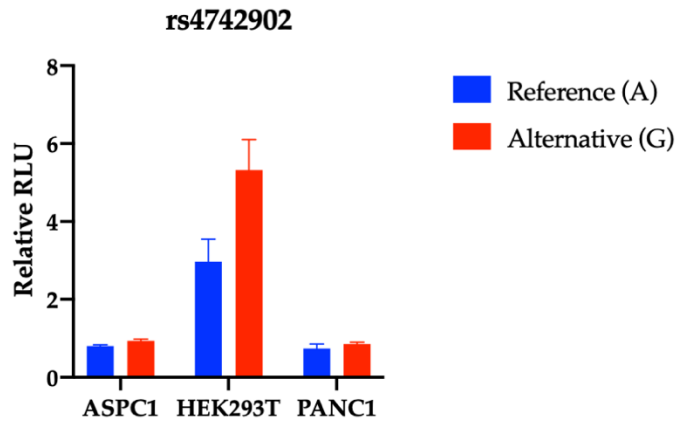


Figure 5.26: (A) Effects of the candidate SNP rs4742902 on gene expression in a dual luciferase reporter assay. Two allelic variants of 5'UTR regions containing each SNP were cloned upstream of the Firefly luciferase gene. A panel of cell lines – HEK293T, ASPC1 and PANC1 – were transfected with reporter constructs together with Renilla. Luminescence was measured 24 hours after transfection using Dual-Glo Luciferase Assay System (Promega). Luciferase expression was normalised to Renilla and made relative to p2Luc plasmid with no insertion Blue boxes – reference allele, red box – alternative allele. (B) DNA sequence surrounding rs4742901 and ChIP-seq based motif logos for RUNX1, RUNX2, RUNX3, ZN563 and ARNT2 according to HOCOMOCO-v11 [285]. A to G substitution alters motif P-value (fold-change is given according to PERFECTOS-APE [286]).

5.2.2.7. Role of SMC2 in PDAC

The expression of *SMC2* in PDAC and normal pancreas tissue was assessed using publicly available datasets. GEPIA2 was used to assess mRNA expression in different tumour and normal tissues across multiple cell types (Figure 5.27 A). The effect of the expression of *SMC2* on the survival of pancreatic cancer patients was plotted using GEPIA2 (Figure 5.27 B). GEPIA2 was also used to plot RNA expression levels during pancreatic cancer stages I-IV (Figure 5.27 C).

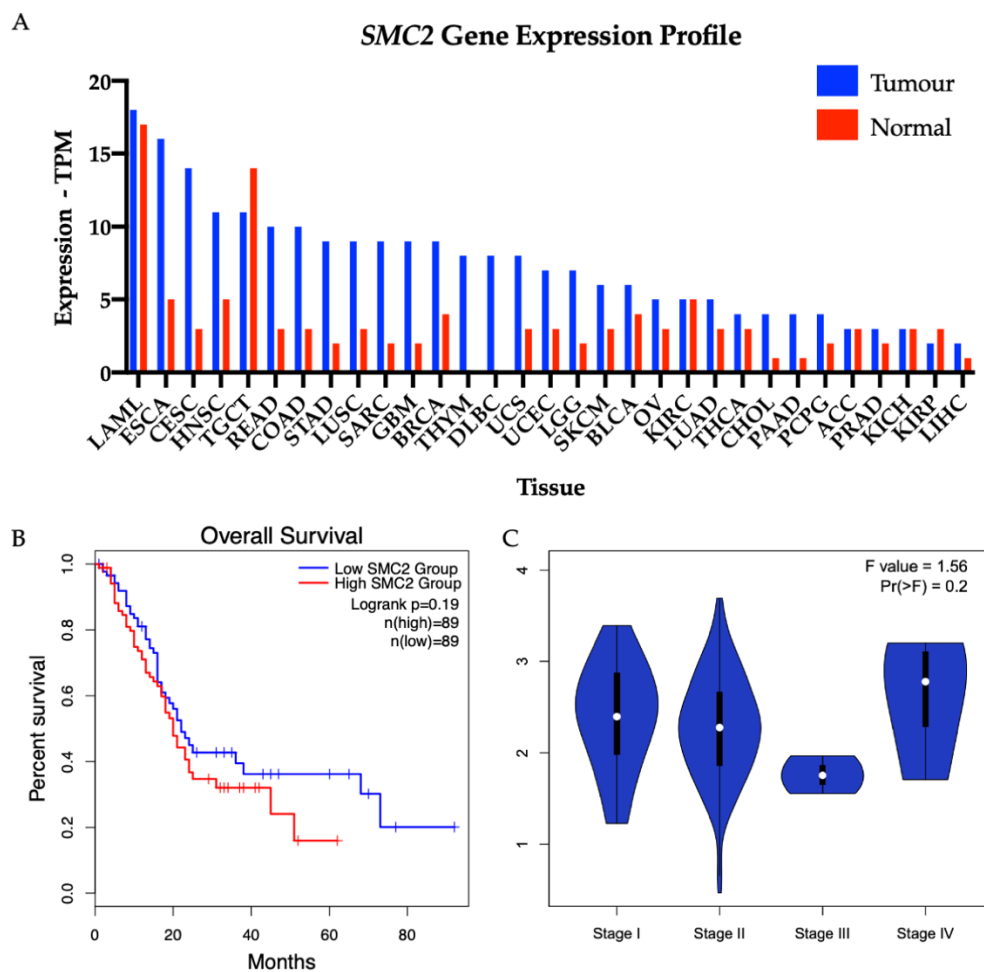


Figure 5.27: The GEPIA2 database was used to examine the expression of *SMC2* in pancreatic cancer and the normal pancreas. Plots were generated using data from TCGA for pancreatic cancer and GTEx for normal pancreas. (A) *SMC2* mRNA expression in different tumour and normal tissues, TCGA tumour abbreviations used. (B) Effect of *SMC2* expression on the overall survival of pancreatic cancer. (C) *SMC2* RNA expression levels during pancreatic cancer stages I-IV.

5.2.3. Targeting the DDR pathway in non-*BRCA* mutant PDAC cells using a novel compound capable of mimicking the effect of *BRCA2* mutation

The function of a novel compound, a dihydroquinoline pyrazoline-based molecule (ARN24089) that disrupts the RAD51-*BRCA2* protein-protein interaction was investigated. The hypothesis is that the drug mimics the effect of *BRCA2* mutation and sensitises cells to DDR targeting drugs, such as cisplatin and olaparib, a PARP inhibitor, as per methodology described in Section 2.6.

5.2.3.1. Single drug response assays

Drug response assays were performed to assess the response to olaparib treatment in PDM37, PDM41, PDM106 and PT291 organoids, and PT127 CLOs. An IC_{50} value for olaparib was not obtained for any line at maximum the concentration set out in this experiment (10 μ M) (Figure 5.28). This was selected as the maximum concentration as plasma concentrations achieved in adults receiving olaparib (100 mg BID) are (C_{min} =1 μ M and C_{max} = 8.5 μ M). For patients receiving 400 mg BID olaparib as a single agent, the plasma concentration (C_{max}) is 18 μ M.

PDM37, PDM41 and PT291 were responsive to treatment with ARN24089, with PDM37 achieving an IC_{50} of 13.95 μ M (95% CI, 9.51 – 20.49). At the maximum soluble concentration of ARN24089 (40 μ M), PDM41 had a mean viability of 46.60% and PT291 had a mean viability of 34.70% (Figure 5.29).

In treatment with cisplatin, PDM41 and PT291 had IC_{50} values of 9.51 μ M (95% CI, 6.2 – 14.4) and 9.74 μ M (95% CI, 6.2 – 15) respectively. PDM106 was not responsive to treatment with cisplatin.

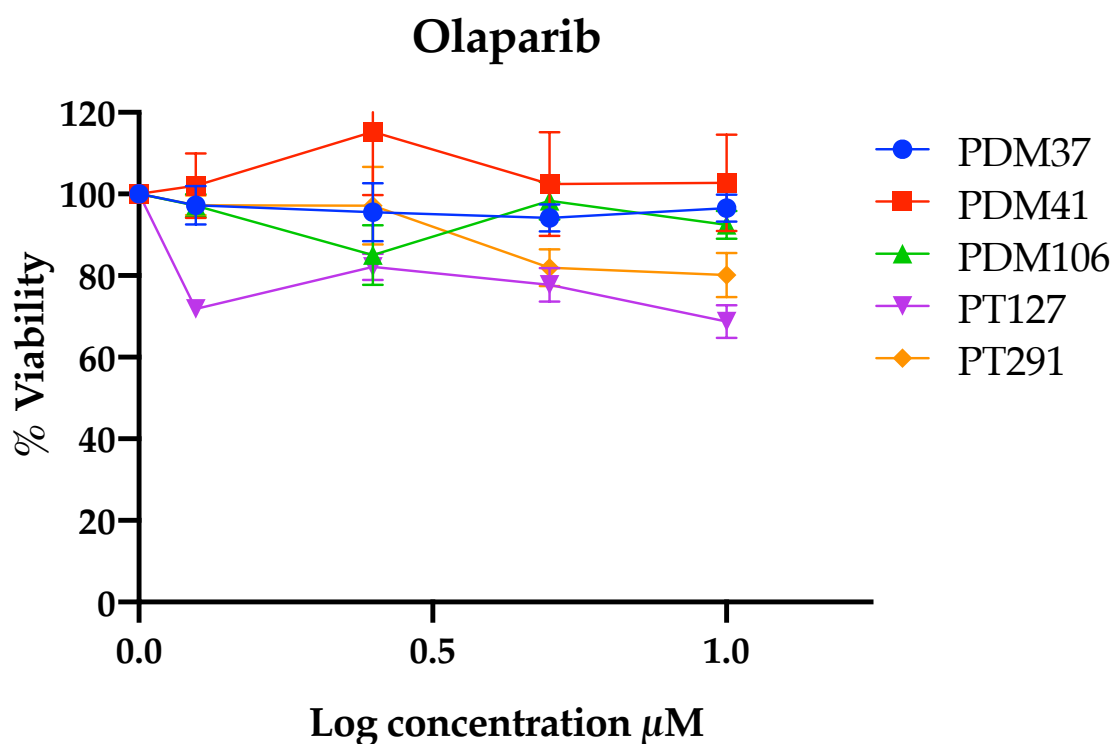


Figure 5.28: Growth inhibition of PDM37 organoids (blue), PDM41 organoids (red), PDM106 organoids (green), PT127 CLOs (purple) and PT291 organoids (orange) over 5 days of treatment with olaparib. Viability of organoids and CLOs was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error of the mean of biological triplicate experiments.

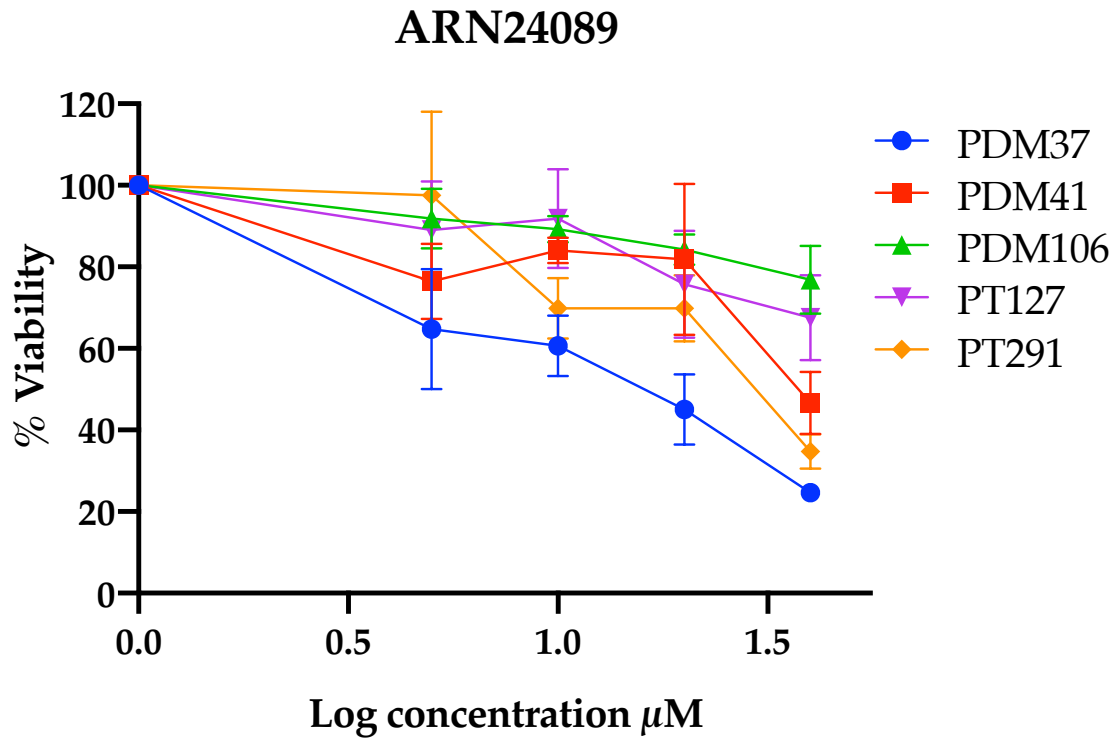


Figure 5.29: Growth inhibition of PDM37 organoids (blue), PDM41 organoids (red), PDM106 organoids (green), PT127 CLOs (purple) and PT291 organoids (orange) over 5 days of treatment ARN24089. Viability of organoids and CLOs was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard deviation of biological triplicate experiments.

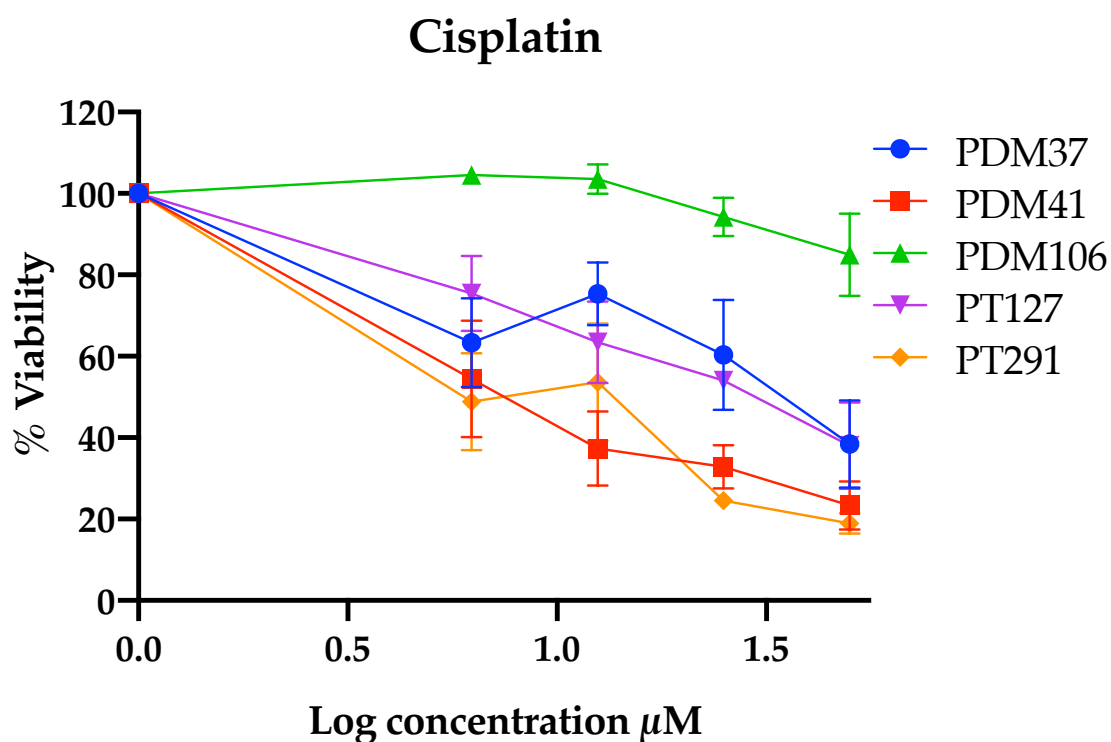


Figure 5.30: Growth inhibition of PDM37 organoids (blue), PDM41 organoids (red), PDM106 organoids (green), PT127 CLOs (purple) and PT291 organoids (orange) over 5 days of treatment with cisplatin. Viability of organoids and CLOs was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error the of mean the of mean of biological triplicate experiments.

5.2.3.2. Combination drug response assays

The novel compound, ARN24089, was tested in combination with olaparib and cisplatin. All organoid lines were tested with the agents in single, double and triple combinations at the concentrations outlined in Table 5.12. In PDM37, no significant difference was found between the treatment with ARN24089 as a single agent or in combination with the other drugs. This is due to the use of a single concentration higher than the IC₅₀ (13.95 μ M) (Figure 5.31). PDM106 showed no significant response to any treatment combination (Figure 5.32). In PDM41 and PT291, no significant difference was observed between single and combination treatments with 5 μ M cisplatin (Figure 5.34 and Figure 5.36). In both PDM41 and PT291, a significant difference was observed between treatment with olaparib alone, and in combination with ARN24089, (PDM41 P=0.021; PT291 P<0.0001). In both PDM41 and PT291, this double combination treatment was more effective than the triple combination with 1 μ M cisplatin, however this difference was not significant (Figure 5.33 and Figure 5.35).

Table 5.12: Combinations used for treatment with organoids with cisplatin, olaparib and ARN24089, with low cisplatin concentration (1 μ M) and high cisplatin concentration (5 μ M).

Combinations – 1 μM cisplatin	Combinations – 5 μM cisplatin
No drug control	No drug control
Cisplatin (1 μ M)	Cisplatin (5 μ M)
Olaparib (10 μ M)	Olaparib (10 μ M)
ARN24089 (20 μ M)	ARN24089 (20 μ M)
Cisplatin (1 μ M) & olaparib (10 μ M)	Cisplatin (5 μ M) & olaparib 10 μ M
Cisplatin (1 μ M) & ARN24089 (20 μ M)	Cisplatin (5 μ M) & ARN24089 (20 μ M)
Olaparib (10 μ M) & ARN24089 (20 μ M)	Olaparib (10 μ M) & ARN24089 (20 μ M)
Cisplatin (1 μ M), olaparib (10 μ M) & ARN24089 (20 μ M)	Cisplatin (5 μ M), olaparib (10 μ M) & ARN24089 (20 μ M)

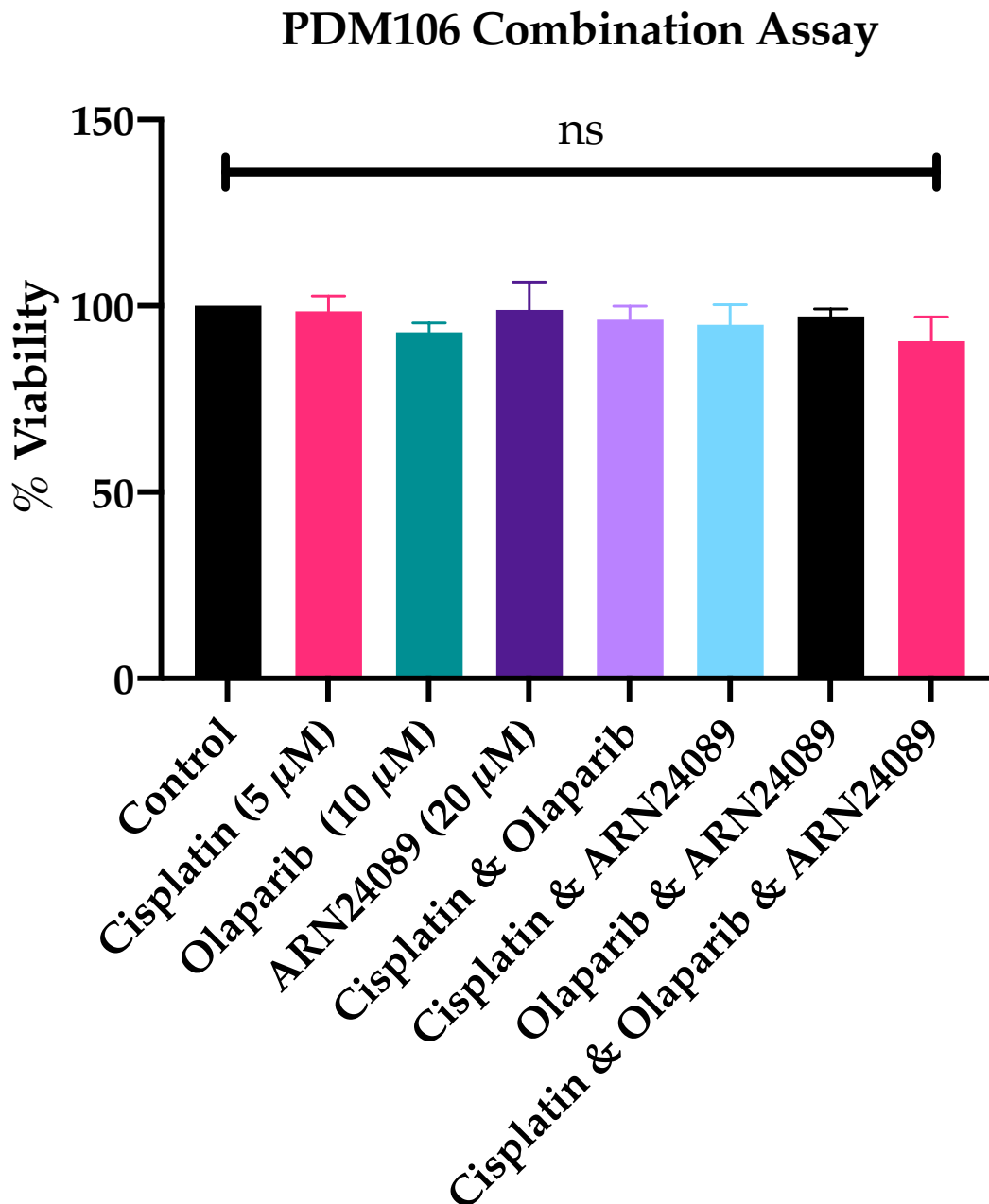


Figure 5.32: Growth inhibition of PDM106 organoids treated with cisplatin (5 μ M), olaparib (10 μ M) and ARN24089 (20 μ M), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. The viability of organoids was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error of the mean of triplicate experiments. A one-way ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

PDM41 Combination Assay

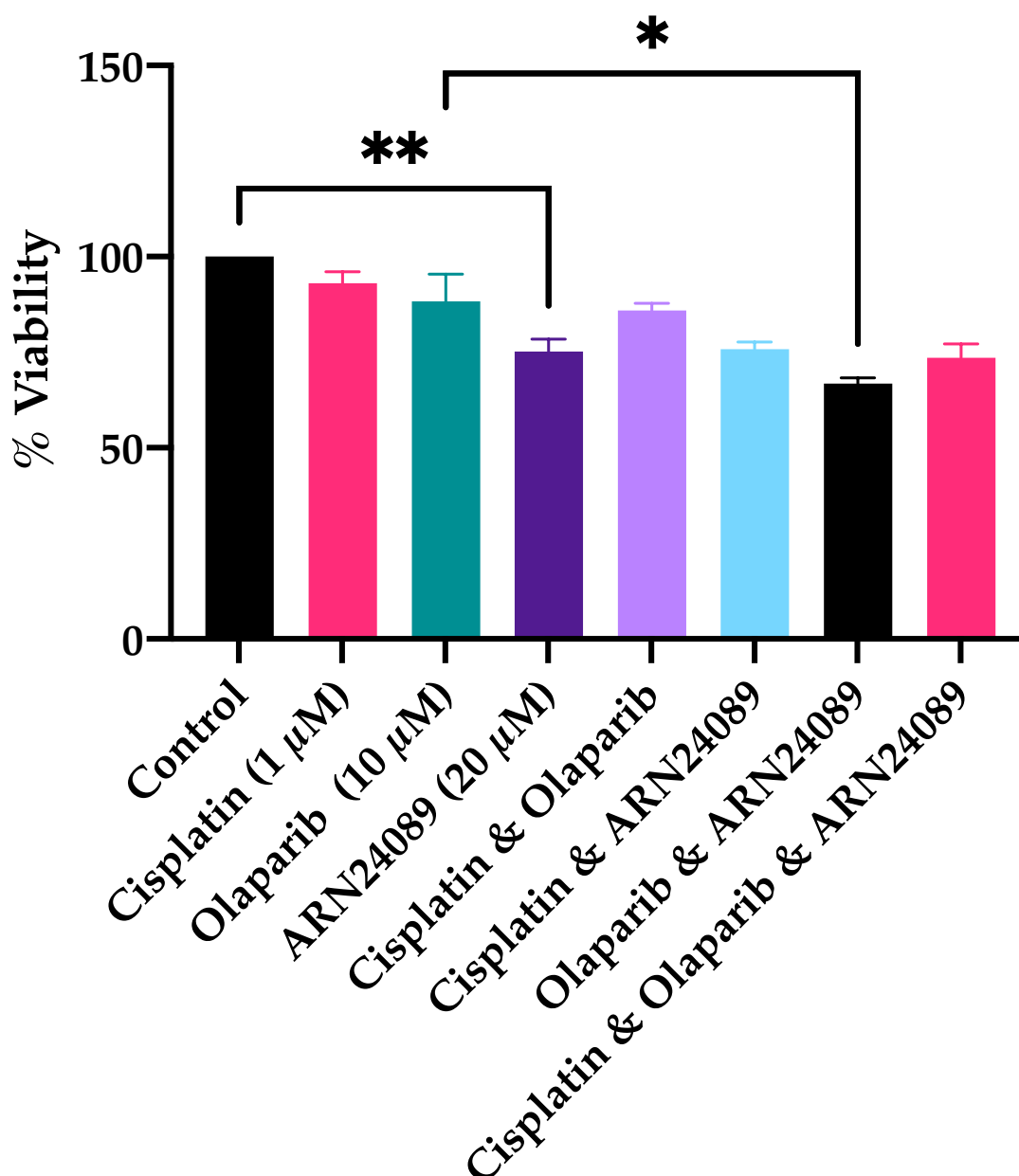


Figure 5.33: Growth inhibition of PDM41 organoid cells treated with cisplatin (1 μ M), olaparib (10 μ M) and ARN24089 (20 μ M), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. The viability of organoids was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error the of mean of triplicate experiments. A one-way ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

PDM41 Combination Assay

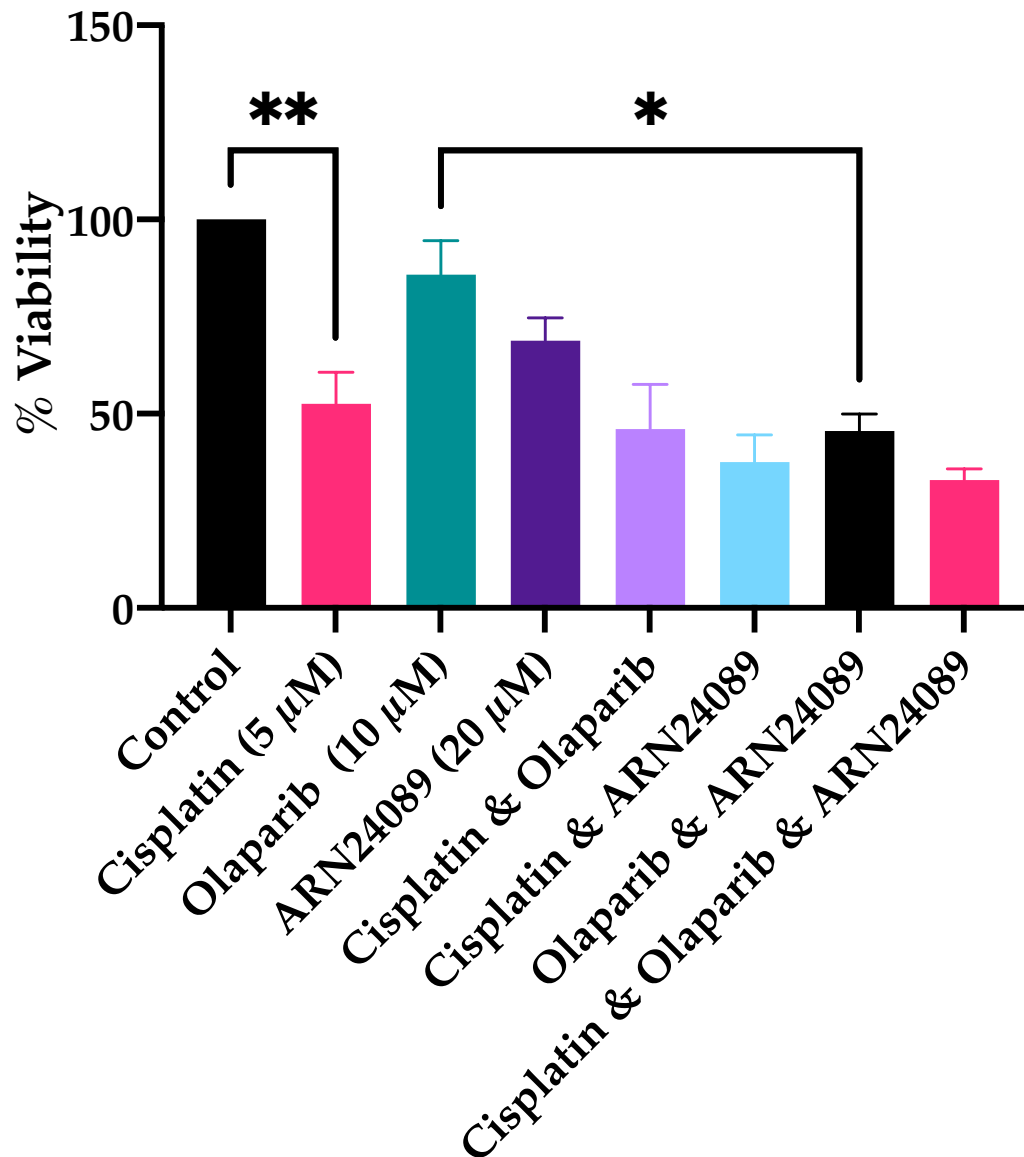


Figure 5.34: Growth inhibition of PDM41 organoid cells treated cisplatin (5 µM), olaparib (10 µM) and ARN24089 (20 µM), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. The viability of organoids was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error the of mean of triplicate experiments. A one-way ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

PT291 Combination Assay

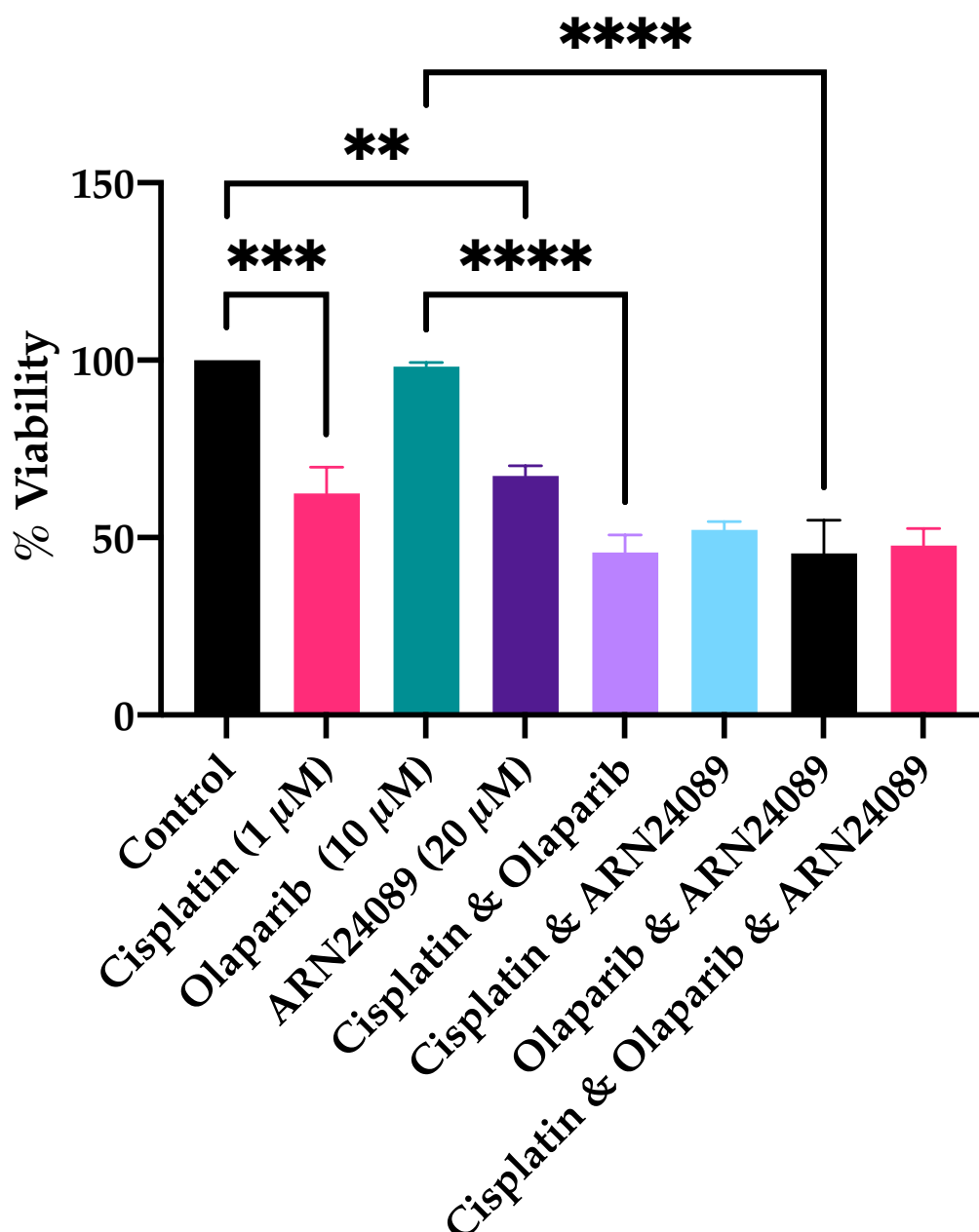


Figure 5.35: Growth inhibition of PT291 organoid cells treated with cisplatin (1 µM), olaparib (10 µM) and ARN24089 (20 µM), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. The viability of organoids was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error of the mean of triplicate experiments. A one-way ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

PT291 Combination Assay

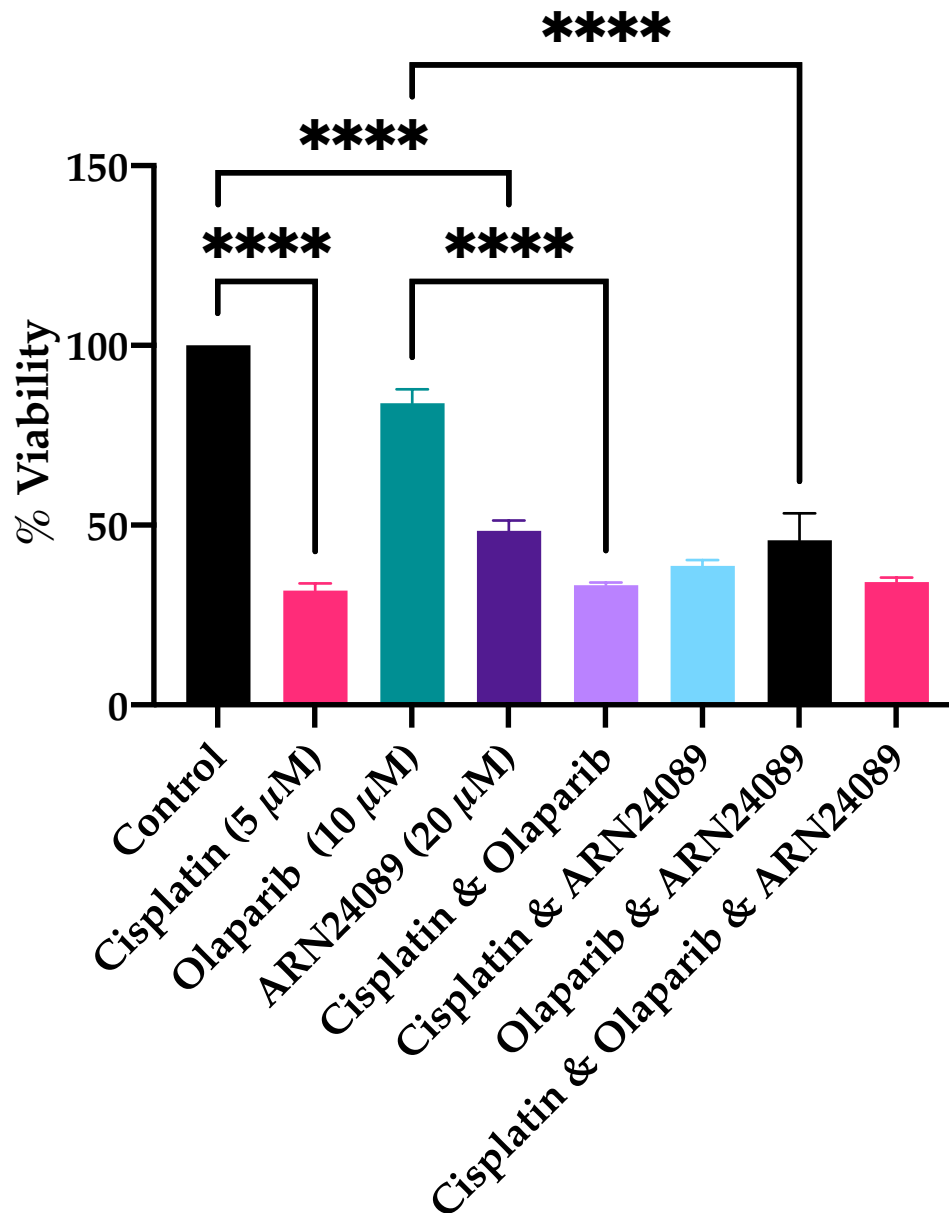


Figure 5.36: Growth inhibition of PT291 organoid cells treated with cisplatin (5 µM), olaparib (10 µM) and ARN24089 (20 µM), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. The viability of organoids was measured by Cell Titre Glo. Percentage viability was calculated relative to untreated control cells. Error bars represent standard error the of mean of triplicate experiments. A one-way ANOVA was used to determine statistical significance. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

5.2.3.3. Translocation of RAD51

RAD51 is a protein which acts with BRCA2 and PALB2 to repair double-stranded breaks in DNA. The expression of RAD51 was assessed in PDM41 and PT291 organoids which had been treated for 24 hours with cisplatin (1 μ M), olaparib (10 μ M) and ARN24089 (20 μ M), alone and in combination. Separated colour channel images show a clear nuclear localisation of RAD51 in cisplatin treated samples compared to the controls (Figure 5.37 and Figure 5.38). In both PDM41 and PT291, significantly increased expression of RAD51 was observed in the samples treated with cisplatin relative to the control (Figure 5.39). There was also a significant reduction in RAD51 expression in olaparib and ARN24089 combination compared to ARN24089 alone.

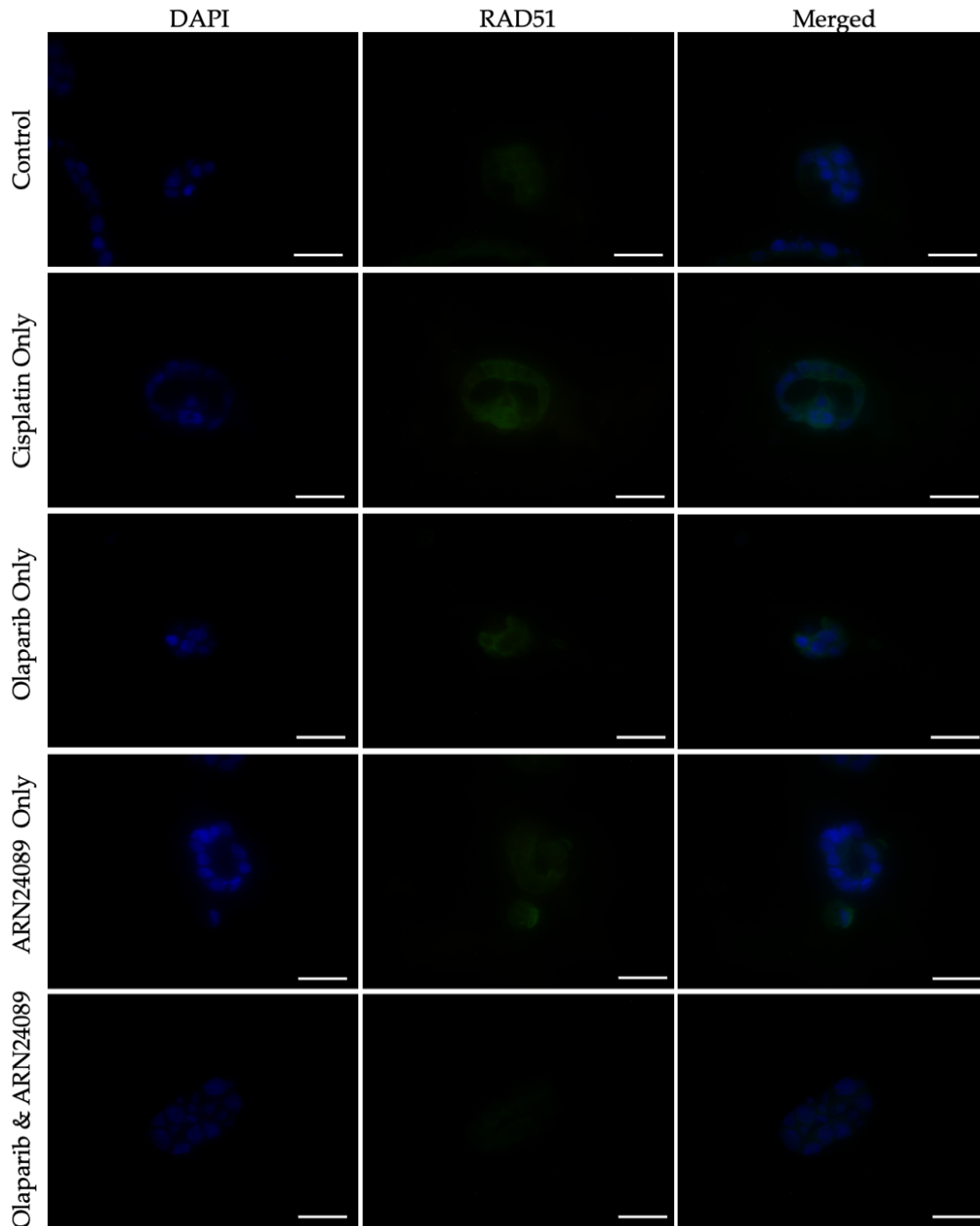


Figure 5.37: Images of immunofluorescence staining of DAPI (blue), RAD51 (green) and merged in PDM41 organoids 24 hours after treatment with cisplatin (1 μ M), olaparib (10 μ M), ARN24089 (20 μ M), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. Organoids were imaged at 630X. Scale bars 40 μ m.

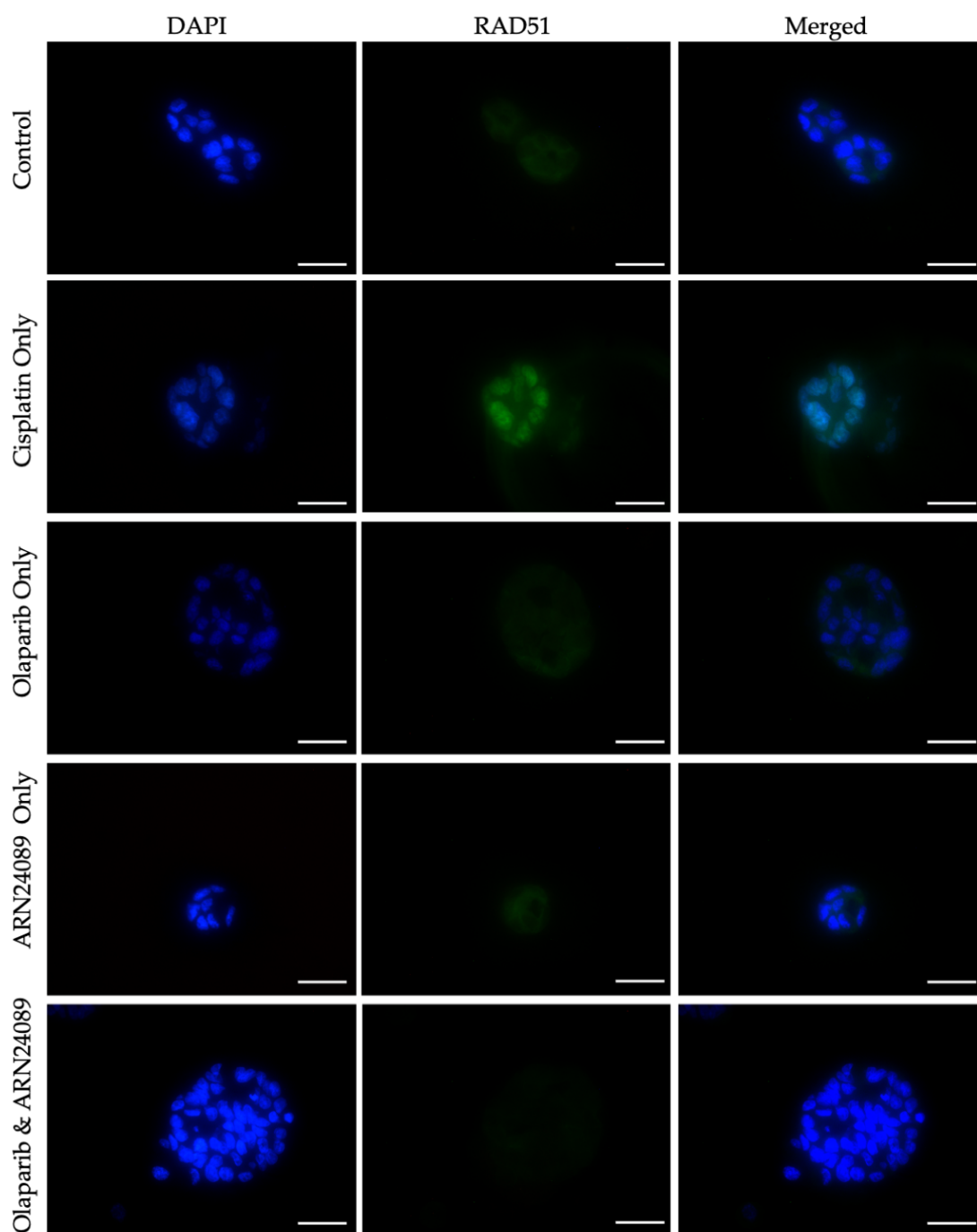


Figure 5.38: Images of immunofluorescence staining of DAPI (blue), RAD51 (green) and merged in PT291 organoids 24 hours after treatment cisplatin (1 μ M), olaparib (10 μ M), ARN24089 (20 μ M), cisplatin plus olaparib, cisplatin plus ARN24089, olaparib plus ARN24089 and cisplatin plus olaparib and ARN24089. Organoids were imaged at 630X. Scale bars 40 μ m.

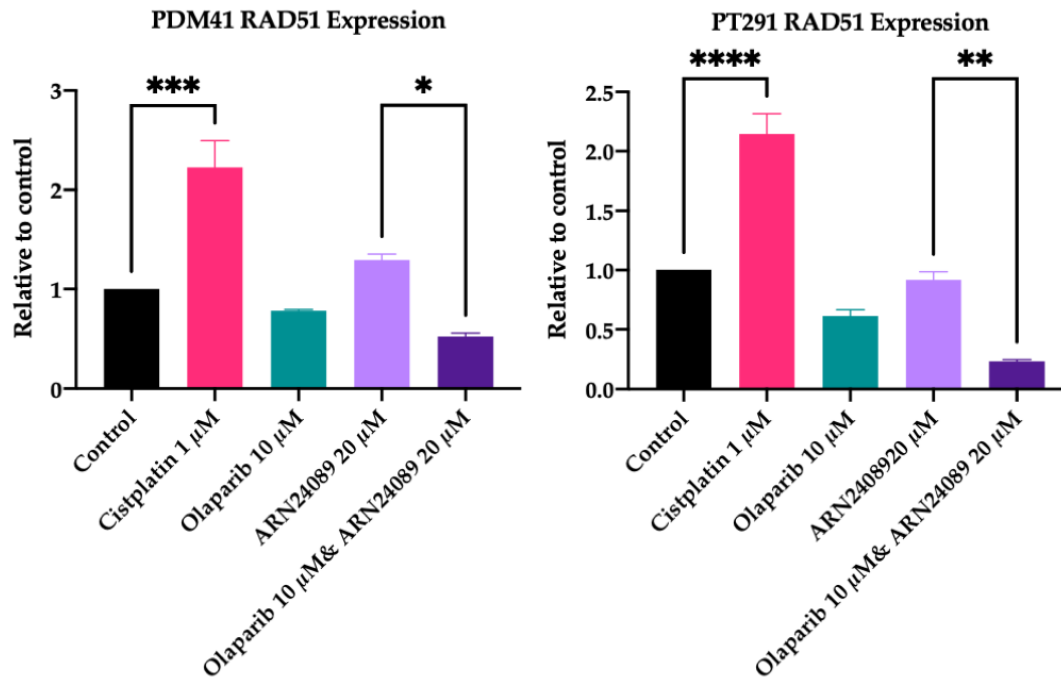


Figure 5.39: A bar graph showing the expression of RAD51 relative to the untreated control 24 hours after treatment with cisplatin (1 μ M), olaparib (10 μ M), ARN24089 (20 μ M), and olaparib plus ARN24089 after in PDM41 and PT291. A one-way ANOVA was performed. * denotes $p < 0.05$, ** denotes $p < 0.01$, *** denotes $p < 0.001$.

5.2.3.4. Caspase-3/7 mediated apoptosis

The Incucyte® Caspase-3/7 dye binds to cleaved caspase-3/7 and allows for the identification of cells undergoing caspase-3/7 mediated apoptosis. This assay was performed in PDM41 and PT291 organoids which had been treated with cisplatin (1 μ M), olaparib (10 μ M) and ARN24089 (20 μ M), alone and in combination. Images were taken every two hours over the course of 5 days (Figure 5.40 and Figure 5.42). Treatment of PDM41 (Figure 5.41) shows that organoids treated with the combination of olaparib and ARN24089 had the highest levels of caspase-3/7 cleavage, with levels increasing until approximately 40 hours. ARN24089 alone had the second highest levels of cleaved caspase-3/7, and olaparib, cisplatin and the control all had similar levels of cleaved caspase-3/7. In PT291, cisplatin alone had the highest levels of cleaved caspase-3/7, followed by olaparib alone. Treatment with the combination of olaparib and ARN24089 showed similar levels to the control (Figure 5.43).

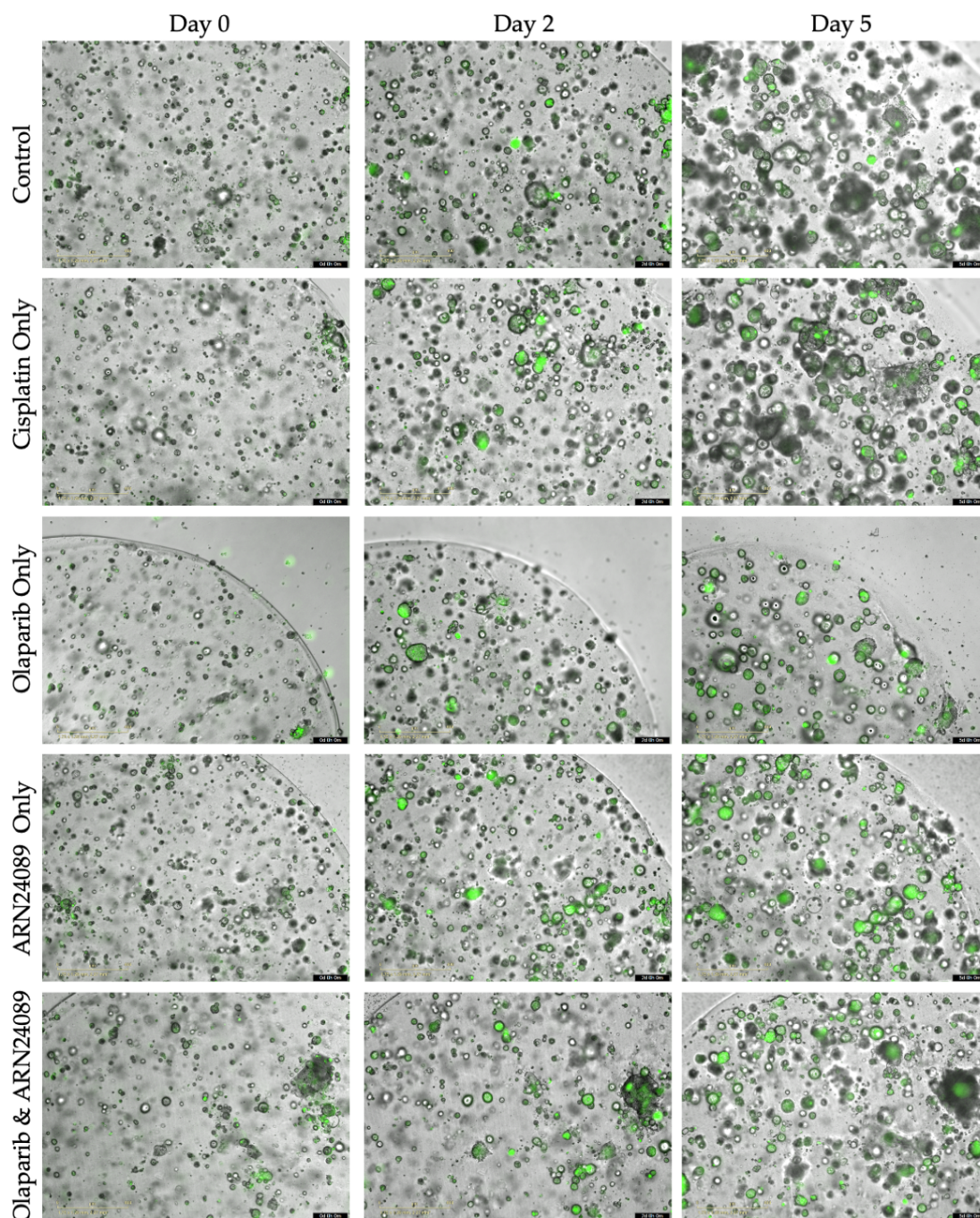


Figure 5.40: Representative images from Day 0, Day 2 and Day 5, taken on Incucyte Live Cell Imaging System of PDM41 organoids treated with cisplatin ($1\ \mu\text{M}$), olaparib ($10\ \mu\text{M}$), ARN24089 ($20\ \mu\text{M}$), and olaparib plus ARN24089 after in the presence of Caspase-3/7 Green Dye. Magnification 100x, scale bars $400\ \mu\text{m}$.

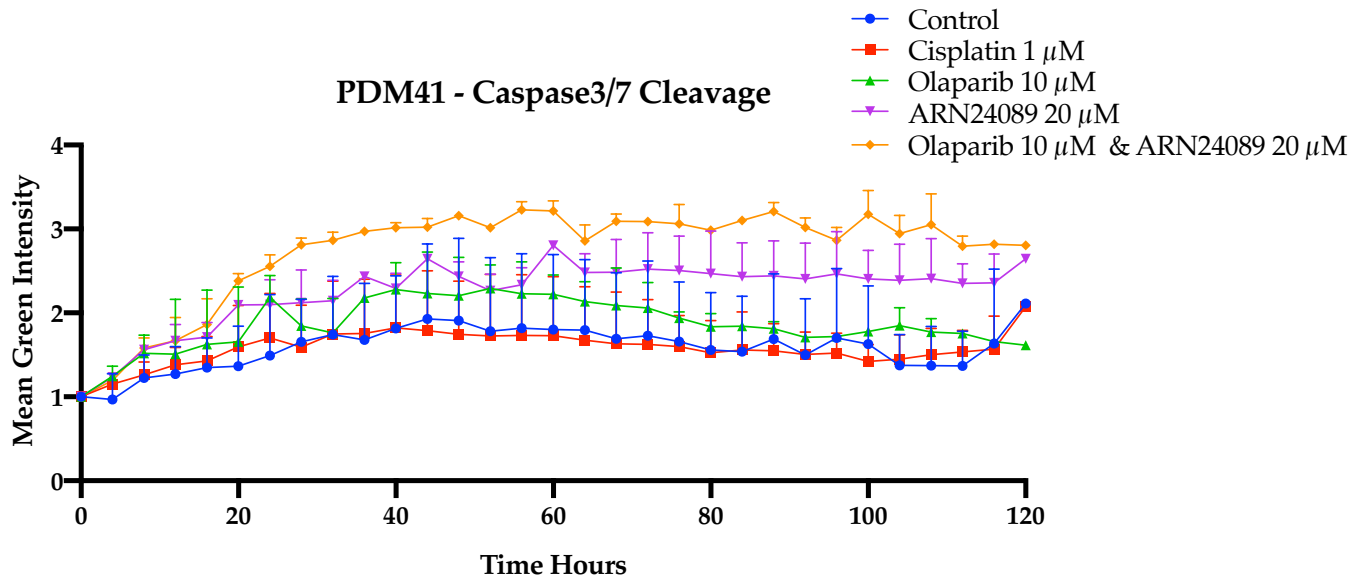


Figure 5.41: Time course of apoptosis of PDM41 organoid in the presence of caspase-3/7 green dye and with cisplatin (1 μ M), olaparib (10 μ M), ARN24089 (20 μ M), and olaparib plus ARN24089. Images taken every 2 hours over the course of five days. Error bars represent standard error the of mean of triplicate experiments.

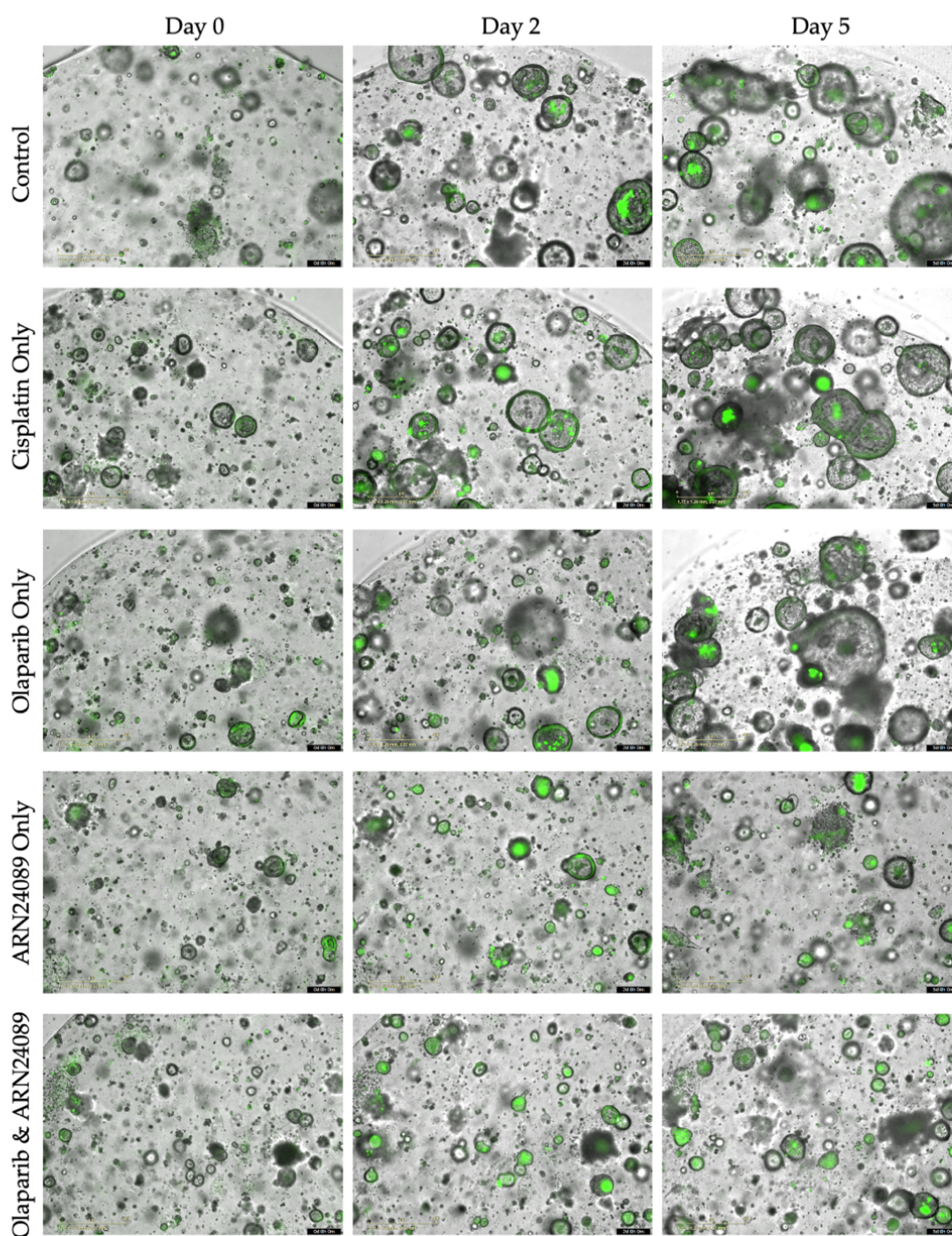


Figure 5.42: Representative images from Day 0, Day 2 and Day 5, taken on Incucyte Live Cell Imaging System of PDM291 organoids treated with cisplatin ($1\ \mu\text{M}$), olaparib ($10\ \mu\text{M}$), ARN24089 ($20\ \mu\text{M}$), and olaparib plus ARN24089 after in the presence of Caspase-3/7 Green Dye. Magnification 100x, scale bars $400\ \mu\text{m}$.

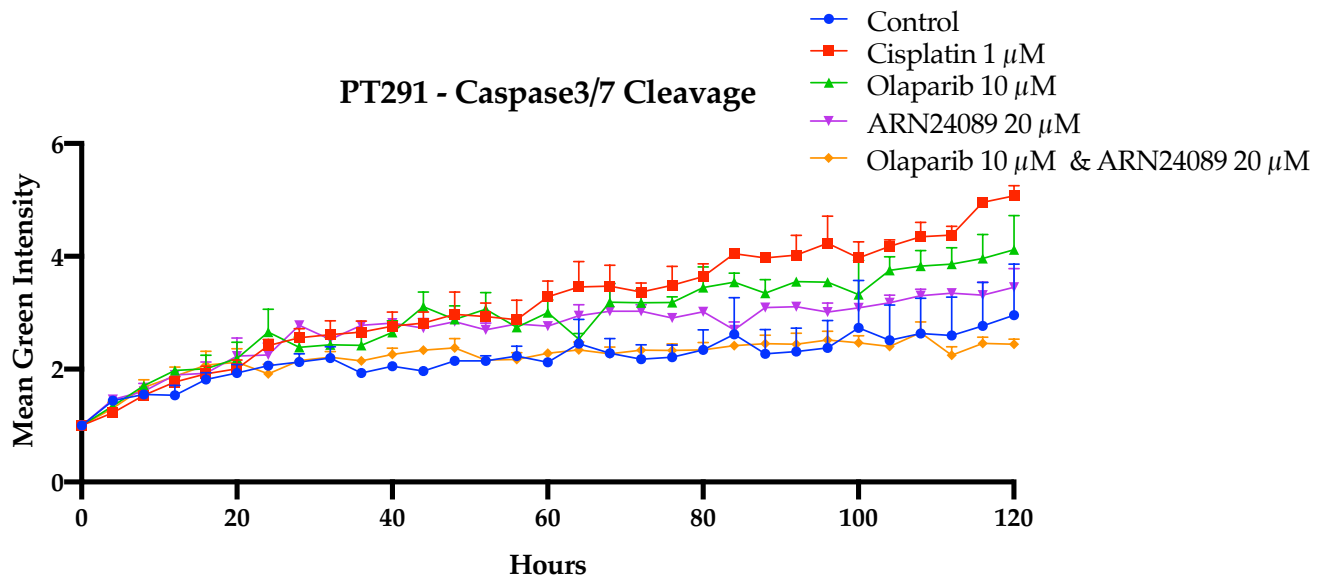


Figure 5.43: Time course of apoptosis of PT291 organoid in the presence of caspase-3/7 green dye and with cisplatin (1 μM), olaparib (10 μM), ARN24089 (20 μM), and olaparib plus ARN24089. Images taken every 2 hours over the course of five days.

Error bars represent standard error the of mean of triplicate experiments.

5.3. Discussion

The GWAS pathway analysis study performed by Walsh *et al.* [122] identified five pathways and gene sets associated with the development of PDAC. An important step in improving outcomes for patients with pancreatic cancer includes validation of genomic variants of PDAC identified in GWAS and pathway analysis studies. The gene set is based on is a network of transcripts whose expression has a positive correlation with ATM in normal tissues [186]. In this chapter, a functional analysis of candidate SNPs and genes identified from the Pujana ATM PCC Network gene set and all SNPs in LD with a high $R^2 \geq 0.8$, was performed.

The pathway analysis highlighted *CASP7* as a gene of interest within the Pujana ATM PCC, as it found that that patients with the SNP rs3124737, had reduced probability of developing PDAC (OR=0.9; CI 0.81-0.95, P-value=2.5E-8). This study also identified that rs3124737 acts as an eQTL as the alternative allele has positive influence on *CASP7* expression in normal pancreas.

Potential functional variants for *CASP7* and *SMC2* were selected for *in vitro* functional characterisation using dual luciferase reporter assays in two established pancreatic cancer cell lines, ASPC1 and PANC1, and the HEK293T cell line to confirm the alternative alleles of these SNPs affected the transcriptional promoter activity.

The role of *CASP7* polymorphisms have been studied in a number of different disease settings and have been found to have an effect on disease risk. A study looking at the role of SNPs in the 3'UTR of *CASP7* found rs4353229 TT was associated with a decreased risk of developing gastric cancer in an Eastern Chinese population [377]. Research into ischaemic stroke (IS) indicates that the rs12415607-A allele induces lower levels of transcriptional activity and *CASP7* mRNA, and thus is associated with a reduced risk of IS [378]. *CASP7* polymorphisms have also been associated with the development of rheumatoid arthritis and diabetes in a Bedouin Arab family [379,380].

Caspases play an important role in the execution of apoptosis; however, the potential role of *CASP7* in pancreatic cancer is unclear. These results found that

CASP7 SNP rs3124737-G increased luciferase activity compared to the reference allele, in ASPC1 and PANC1, indicating that *CASP7* may play a role in pancreatic cancer development. HOCOMOCO analysis predicts that the presence of rs3124737-G alters the binding site of a number of proteins. This alteration results in increased binding potential of BRCA1, a known tumour suppressor by 6.74-fold [381]. It also results in decreased binding potential of transcription factor MYBB (MYBL2). MYBB has previously been shown to be overexpressed in PDAC [382]. In lung cancer cells, knockdown of MYBB resulted in the induction of apoptosis, highlighting the potential link between MYBB and *CASP7* expression [383].

In order to establish the protein expression of *CASP7* in PDAC, IHC was performed on a cohort of 49 PDAC cancer patients, and three normal pancreas samples. Our results found that 67% of the samples were *CASP7* negative. Survival analysis on the patient cohort was performed and patient tumours expressing *CASP7* had a significantly increased survival time compared to those who did not express *CASP7* (24.15 vs. 29.43 months). However, it is important to note that scoring by a pathologist and multivariate analysis is required before any definitive conclusion on survival can be determined from this data. Similar levels of expression were obtained from proteintlas.org, which found 58% of samples stained had no *CASP7* expression, indicating differential expression among PDAC cases. Survival analysis on the expression of *CASP7* expression from GEPIA2 shows different overall survival levels to the one performed from the patient cohort. There are a number of potential reasons for this discrepancy – GEPIA2 analysis used mRNA data, whereas analysis on the cohort used protein expression data. Additionally, the GEPIA2 dataset was divided into above and before the 50th percentile, however, the patient cohort was divided into expression and no expression. Furthermore, analysis of the expression of *CASP7* in established cell lines was performed. Differential expression of *CASP7* in PDAC cell lines was observed, with higher levels of *CASP7* in established PDAC cell lines derived from primary tumours compared to those derived from metastatic site of PDAC. Parallel analysis of protein expression of total *CASP7* revealed that the metastatic derived cells lines, ASPC1, HuPT4 and SW1990 had the lowest *CASP7* protein expression, correlating with the mRNA data.

Inactivating mutations of the *CASP7* gene are rare; although functional analysis shows that inactivating *CASP7* mutations *in vitro* may promote loss of its apoptotic function, thereby contributing to the pathogenesis of some solid cancers [384]. *CASP7* is found to be downregulated in colon and gastric [385,386]. In ER positive breast cancer, high expression of *CASP7* showed better relapse-free survival [387]. The presence of cleaved *CASP7* in breast cancer results in a better prognosis [388].

SNPs in close proximity (< 20 kb upstream/downstream) to the *SMC2* gene were selected for functional analysis as these were in CpG islands and had H3K27ac histone marks in the pancreas. In addition, these SNPs were in LD with rs7859034, the top gene signal in the Pujana ATM PCC Network in the PDAC GWAS pathway analysis previously conducted by our group [122]. The SNPs selected for functional analysis by dual luciferase reporter assay, rs4742901 and rs4742902 are 200 bps apart, and are located upstream of the *SMC2* gene. The luciferase assay data showed an increase in the functional activity of the construct harbouring the alternative rs4742901-C allele in the HEK293T cell line only. Interestingly, HOCOMOCO-v11 motif analysis of the DNA regions surrounding the SNPs showed that rs4742901-C resulted in a gain of binding potential of MXI1-long, MXI1, BHE, MYC and HES5. A strong enhancer activity was observed for rs4742902-G (alt) compared to the A (ref) allele in all three cell lines tested. Analysis of motifs suggest that the rs4742902-G (alt) results in a loss of binding of RUNX2, RUNX1, RUN3, and ZNF563, but preferential binding to TF ARNT2.

The structural maintenance of chromosome 2 (*SMC2*) plays an important role in packaging of chromatin before cell division and DDR, which is required for proper chromosome segregation and maintenance of chromosomal stability [389]. *SMC2* is suggested to have a pro-oncogenic role, and is direct transcriptional target of the oncogenic Wnt signalling pathway [390]. *SMC2* recruitment to mitotic chromosomes is controlled by DDR checkpoints, such as Chk2, which facilitates effective repair of genomic damage, therefore further highlighting the role of the DDR pathway in PDAC development [391].

The DDR pathway is a complex network of cellular pathways capable of repairing abnormal DNA lesions. The PARP pathway is used as a salvage mechanism to detect and repair DNA damage. The use of PARP inhibitors is well established in *BRCA*-associated cancers, in which *BRCA* mutant cancers show better responses to platinum-based agents [392]. FOLFIRINOX is the only platinum-based treatment regime that has been approved for PDAC, retrospective studies have shown that treatment with FOLFIRINOX shows a significant survival in patients with germline mutations in *BRCA1/BRCA2* and *PALB2*.

The phase III POLO study evaluated PARPi olaparib, as maintenance therapy in metastatic PDAC and germline *BRCA1/2* mutations. Patients that had stable disease following 16 weeks of platinum-based therapy were then randomised to receive olaparib or placebo [180]. Patients in the olaparib-treated arm displayed significantly longer progression free survival (median PFS: 7.4 versus 3.8 months) and an improvement in objective response rate (23.1% vs. 11.5%) and median duration of response (24.9 vs. 3.7 months), but no OS difference was observed. This practice changing trial was the first to stratify PDAC based on genetic mutations and led to the FDA approval of olaparib for the maintenance treatment of adult patients with deleterious or suspected deleterious germline *BRCA* mutated (gBRCAm) in metastatic PDAC. However, interesting findings in this trial and others have found that between 22-40% of gBRCAm PDAC patients progressed or did not respond to platinum-based/PARPi therapies [181,393]. Furthermore, recent studies found that the addition of the PARPi veliparib to gemcitabine plus cisplatin did not improve response rate in stage III-IV PDAC with gBRCAm or *PALB2* mutations, indicating the presence of patients with alternative non-HR deficient mechanisms [394]. Additionally, retrospective studies have shown that treatment with FOLFIRINOX results in a significant increase in survival in patients with mutations in *BRCA1/BRCA2* and *PALB2* [395]. However, the use of high dose PARP inhibitors and multi-chemotherapy regime FOLFIRINOX for the treatment of PDAC can be extremely toxic, and only 25% of patients are eligible due to the high levels of adverse effects [38,396,397].

This suggests that there are sub-populations in both sporadic and hereditary PDAC where other distinct modulators could be exploited to provide significant benefit to PDAC patients. This could also open up new avenues for novel anti-cancer concepts. Therefore, to further elucidate if the DDR pathway can be treated in non-germline *BRCA1/2* mutated PDAC, a novel dihydroquinoline pyrazoline-based molecule (ARN24089) was tested in organoid models [372]. During the DDR process, RAD51 is recruited by BRCA2 to the double-stranded break sites to by PALB2. Activated RAD51 allows the incursion of sister chromatids to allow for homologous repair. ATM loss also results in a failure to make this complex, leading to error prone NHEJ [392]. ARN24089 disrupts the RAD51-BRCA2 protein-protein interaction, mimicking a *BRCA2* mutation thus sensitising the cells to treatment with PARP inhibitor [398]. The novel compound ARN24089 disrupts the RAD51-BRCA2 protein-protein interaction, mimicking a *BRCA2* mutation thus sensitising the cells to treatment with a PARP inhibitor [372]. This study aimed to determine, for the first time, the efficacy of ARN24089 in organoids models. *In vitro* growth inhibitory assays were performed with ARN24089 (20 μ M), olaparib (10 μ M) and cisplatin (1 μ M & 5 μ M) as single, double and triple combinations. These assays found that after 5 days of treatment, the double combination of ARN24089 (20 μ M) and olaparib (10 μ M) was more effective at growth inhibition than the triple combination of ARN24089 (20 μ M), olaparib (10 μ M) and cisplatin (1 μ M).

Immunofluorescence imaging of RAD51 organoids treated with single, double and triple combinations of cisplatin (1 μ M), olaparib (10 μ M) and ARN24089 (20 μ M) show clear differences in the localisation of RAD51 after treatment for 24 hours. Cisplatin only treated cells show a clear localisation of RAD51 to the nucleus, whereas treatment with the combination of ARN24089 and olaparib show significantly reduced levels of RAD51 in the organoids compared to the control, and treatment with ARN24089 only. A study by Cruz *et al.* [399] found that the detection of RAD51 foci in the nucleus correlates with PARPi resistance. Nuclear localisation of RAD51 occurs in response to DNA damage, however this process requires BRCA2 to occur [400,401]. Together, this suggests that RAD51 foci induction to the nucleus was blocked by ARN24089 and olaparib, indicating that in PDAC organoid models, when used in combination with olaparib, ARN24089

disrupts the RAD51-BRCA2 protein-protein interaction. Further confirmation would require analysis of the pH2AX foci number, indicating the accumulation of DNA double-strand breaks within the combination.

The results show that when used in the treatment of PDAC organoid lines PDM41 and PT291 show the same level, or greater response to treatment with a dual combination of ARN24089 (20 μ M) and olaparib (10 μ M) compared to treatment with a triple combination of cisplatin (1 μ M), ARN24089 (20 μ M) and olaparib (10 μ M) over 5 days. Using the Incucyte Imaging Software, it was observed that in PDM41 the highest levels of casapase-3/7 cleavage was in the double (olaparib/ARN24089) combination, with levels of cleaved caspase-3/7 increasing until the 40-hour timepoint, and then remaining stable. Treatment with cisplatin (1 μ M) resulted in similar levels of caspase cleavage as the control. In PT291, the lowest levels of caspase3/7 cleavage are observed in organoids treated with a combination of ARN24089 and olaparib, and the highest levels were observed in cisplatin treated organoids. However, differences in size of the organoids, resulting from decreased proliferation due to drug treatment may account for these differences. The variable results of apoptosis induction by the ARN24089/olaparib combination in both organoids may require further analysis such as cell cycle analysis, assessing levels of cleaved-PARP and pH2AX foci number to determine the mechanisms of cell death. Furthermore, as we observed that the DNA damaging agent cisplatin induces translocation of RAD51 to the nucleus, *in vitro* pre-treatment with cisplatin prior to ARN24089/olaparib may improve the effectiveness of these drug combinations.

PARPs are activated by DNA breaks and use of PARP inhibitors are well established in BRCA-associated cancers [392]. However, use of high dose PARP inhibitors for the treatment of PDAC can be extremely toxic, with side effects including sporadic acute myeloid leukaemia [396]. High levels of >3 grade 3 toxicities are associated in patients treated with combinations of PARP and cisplatin/gemcitabine [396]. In order to improve results in *in vitro* assays using organoids to assess the effectiveness of these drugs, pre-treatment with a DNA damaging agent, such as cisplatin, may be required.

While the work in this chapter has shed light on the importance role of the DDR pathway in PDAC research, it has a number of limitations. Due to the limited solubility of ARN24089 (maximum soluble concentration of 40 μ M), it was not possible to study its potential use as a single agent in PDAC organoids, as a reliable IC_{50} value could not be obtained. This drug will also require improvements in its solubility before it can potentially be used *in vivo* studies. Finally, there are a number of limitations of the dual luciferase reporter assay, as the plasmid had limited available restriction enzyme sites, therefore sequences which contained test RE sites were excluded from further analysis. We were also unable to obtain the plasmid without a SV40 promoter, so the negative control used in the experiment was cells with transfection reagent only.

In conclusion, here we propose that SNPs in the Pujana ATM PCC Network gene set may be involved in regulating key genes *CASP7* and *SMC2*, further underpinning the potential role of DDR in PDAC development. Direct targeting of the RAD51-BRCA2 protein-protein complex in non-*BRCA2* mutated PDAC organoids highlights a potential new therapeutic strategies for the treatment of pancreatic cancer.

[This page is intentionally left blank]

Chapter 6. Summary and future work

6.1. Summary

The overall aim of this project was to develop a novel *in vitro* model for the study of PDAC and to validate genomic variants of the disease. By establishing novel methods for the development of organoids from PDAC tumour samples, the work in this thesis has built on the advancements in the field of *in vitro* cancer modelling to bridge the *in vivo*/ *in vitro* gap in PDAC research. In addition, methods for the generation of isogenic PDAC cell lines, and a novel model for PDAC, CLOs have been outlined in this thesis. RNA-sequencing has validated that the organoid and CLO models are transcriptomically similar and are more representative of the PDX tumour than the isogenic cell line. Organoid and CLO models have also been shown to be morphologically and phenotypically similar. Finally, an *in vivo* study has shown that when implanted subcutaneously, the CLO model was capable of recapitulating the morphology of adenocarcinoma. These CLO models makes the use of more representative *in vivo* organotypic models more accessible for all PDAC researchers.

In addition, work in this thesis has prioritised, and functionally validated GWAS pathway analysis associated SNPs through dual luciferase reporter assay and proposes that SNPs in the DDR pathway (rs4742902, rs3124737) and MODY pathway (rs488087) have a functional effect on gene expression. Analysis of the role of the associated genes (*SMC2*, *CASP7* and *CEL*) suggests that these genes may play a role in the development and progression of the disease. Further analysis of MODY transcription factors has highlighted differences in expression of *HNFLA* and *HNF4G* effects the of overall survival of patients, and differential expression of both genes has downstream effects on a number of molecular pathways. Further underpinning the potential role of the MODY pathway in PDAC.

Additionally, this thesis explored the role of DDR in drug response, and organoid models were used to assess the effectiveness of the novel compound ARN24089. This compound has shown effectiveness alone, and in combination with olaparib, in a number of different organoid models

6.2. Conclusions

In conclusion, this thesis presents a novel *in vitro* model of PDAC: CLOs, which are more representative of the tumour growth and development than established cell lines. This model overcomes a number of the issues associated with the use of organotypic models, including sourcing of tumour samples, slow growth rates and use of expensive reagents, making organoid culture assessable to more researchers. Additionally, the use of CLOs is not limited to PDAC, and therefore has the potential to improve *in vitro* modelling for all cancer types.

The functional validation of GWAS pathway analysis associated SNPs through dual luciferase reporter assay has highlighted the potential mechanisms underlying PDAC carcinogenesis. These SNPs could also be used for the development of multigenic predictors for identifying individuals at high risk of developing pancreatic cancer. Finally, the use of a novel compound, ARN24089, in the direct targeting of the RAD51-BRCA2 protein-protein complex in non-*BRCA2* mutated PDAC organoids highlights new potential therapeutic strategies for the treatment of PDAC.

6.3. Future work

1. Further development of organoid and CLO models to make them more representative of the PDAC tumours in patients. This includes developing methods to use cells present in the PDAC tumour microenvironment, stromal cells, cancer-associated fibroblasts and immune cells.
2. Optimisation of the method for establishing organoids and isogenic primary cell lines from PDX tumours for other cancer types, as well as validating the use of CLOs in different cancer types.

3. The generation of CRISPR-knockout cell lines and organoids would allow for further assessment of the function of MODY and DDR pathway genes, such as *HNFI1A*, *HNF4G*, *SMC2* and *CASP7* in PDAC, in the development and progression of the disease.
4. To further assess the effect of SNPs on the development and progression of PDAC, functional validation through dual luciferase reporter assay of all SNPs identified in a GWAS and pathway analysis studies should be performed using a large panel of PDAC cell lines. This would allow for identification of subtypes of the disease in which the SNPs may have an effect.
5. Generation of CRISPR edited organoids and cell lines, would allow for the full functional validation of the GWAS SNPs in PDAC, and results from this could be used in the development of a multigenic predictors for identifying people at risk of developing the disease.
6. Further investigations into the use of novel compound ARN24089 in the *in vivo* treatment of PDAC. Including optimisation of treatment regimens in combinations of ARN24089, cisplatin and olaparib, including pre-treatment priming with cisplatin, followed by treatment with a combination of ARN24089 and olaparib. Other treatment combinations using drugs, such as immunotherapy pembrolizumab could be assessed. Establishing the mechanism of action of the drug through assessment of cytostatic/cytotoxic effects including cell cycle progression, type of cell death mechanism (apoptosis, necrosis, necroptosis). Additionally, the effect of treatment on DDR pathway could be assessed through evaluating the expression of γ -H2AX and cleaved-PARP in treated cells and performing p21/cell cycle analysis.

[This page is intentionally left blank]

Chapter 7. References

1. Bray, F.; Laversanne, M.; Weiderpass, E.; Soerjomataram, I. The ever-increasing importance of cancer as a leading cause of premature death worldwide. *Cancer* **2021**.
2. Omran, A.R. The epidemiologic transition. A theory of the epidemiology of population change. *Milbank Mem. Fund Q.* **1971**, *79*, 161–170.
3. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **2021**, *71*, 209–249.
4. Siegel, R.L.; Miller, K.D.; Fuchs, H.E.; Jemal, A. Cancer Statistics, 2021. *CA. Cancer J. Clin.* **2021**, *71*, 7–33.
5. Chiorean, E.G.; Covelev, A.L. Pancreatic cancer: optimizing treatment options, new, and emerging targeted therapies. *Drug Des Devel Ther* **2015**, *9*, 3529–3545.
6. Rawla, P.; Sunkara, T.; Gaduputi, V. Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. *World J. Oncol.* **2019**, *10*, 10.
7. Becker, A.E.; Hernandez, Y.G.; Frucht, H.; Lucas, A.L. Pancreatic ductal adenocarcinoma: Risk factors, screening, and early detection. *World J. Gastroenterol.* **2014**, *20*, 11182–11198.
8. Lynch, S.M.; Vrieling, A.; Lubin, J.H.; Kraft, P.; Mendelsohn, J.B.; Hartge, P.; Canzian, F.; Steplowski, E.; Arslan, A.A.; Gross, M.; et al. Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *Am. J. Epidemiol.* **2009**, *170*, 403–413.
9. Stolzenberg-Solomon, R.Z.; Amundadottir, L.T. Epidemiology and inherited predisposition for sporadic pancreatic adenocarcinoma. *Hematol. Oncol. Clin. North Am.* **2015**, *29*, 619–640.
10. Arslan, A.A.; Helzlsouer, K.J.; Kooperberg, C.; Shu, X.-O.; Steplowski, E.; Bueno-de-Mesquita, H.B.; Fuchs, C.S.; Gross, M.D.; Jacobs, E.J.; LaCroix, A.Z. Anthropometric measures, body mass index, and pancreatic cancer: a pooled analysis from the Pancreatic Cancer Cohort Consortium (PanScan). *Arch.*

- Intern. Med.* **2010**, *170*, 791–802.
11. Elena, J.W.; Steplowski, E.; Yu, K.; Hartge, P.; Tobias, G.S.; Brotzman, M.J.; Chanock, S.J.; Stolzenberg-Solomon, R.Z.; Arslan, A.A.; Bueno-de-Mesquita, H.B. Diabetes and risk of pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. *Cancer Causes Control* **2013**, *24*, 13–25.
 12. Johansson, K.A.; Grapin-Botton, A. Development and diseases of the pancreas. *Clin. Genet.* **2002**, *62*, 14–23.
 13. Whitcomb, D.C.; Lowe, M.E. Human pancreatic digestive enzymes. *Dig. Dis. Sci.* **2007**, *52*, 1–17.
 14. Oregon State University The Pancreas Available online: https://open.oregonstate.edu/aandp/chapter/17-9-the-pancreas/#fig-ch18_09_01.
 15. Friedlander, S.Y.G.; Chu, G.C.; Snyder, E.L.; Girnius, N.; Dibelius, G.; Crowley, D.; Vasile, E.; DePinho, R.A.; Jacks, T. Context-Dependent Transformation of Adult Pancreatic Cells by Oncogenic K-Ras. *Cancer Cell* **2009**, *16*, 379–389.
 16. Kanda, M.; Matthaei, H.; Wu, J.; Hong, S.M.; Yu, J.; Borges, M.; Hruban, R.H.; Maitra, A.; Kinzler, K.; Vogelstein, B.; et al. Presence of somatic mutations in most early-stage pancreatic intraepithelial neoplasia. *Gastroenterology* **2012**, *142*, 730–733.
 17. Hezel, A.F.; Kimmelman, A.C.; Stanger, B.Z.; Bardeesy, N.; DePinho, R.A. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev.* **2006**, *20*, 1218–1249.
 18. Schafer, K.A. The Cell Cycle: A Review. *Vet. Pathol.* **1998**, *35*, 461–478.
 19. Bertoli, C.; de Bruin, R.A.M. Cell Division: Turning cell cycle entry on its head. *Elife* **2014**, *3*, e03475.
 20. Ballehaninna, U.K.; Chamberlain, R.S. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J. Gastrointest. Oncol.* **2012**, *3*, 105–119.
 21. Buxbaum, J.L.; Eloubeidi, M.A. Molecular and clinical markers of pancreas cancer. *JOP. Journal of the Pancreas* **2010**, *11*, 536–544.
 22. Ni, X.G.; Bai, X.F.; Mao, Y.L.; Shao, Y.F.; Wu, J.X.; Shan, Y.; Wang, C.F.; Wang,

- J.; Tian, Y.T.; Liu, Q. The clinical value of serum CEA, CA19-9, and CA242 in the diagnosis and prognosis of pancreatic cancer. *Eur. J. Surg. Oncol.* **2005**, *31*, 164–169.
23. Kim, J.; Bamlet, W.R.; Oberg, A.L.; Chaffee, K.G.; Donahue, G.; Cao, X.J.; Chari, S.; Garcia, B.A.; Petersen, G.M.; Zaret, K.S. Detection of early pancreatic ductal adenocarcinoma with thrombospondin-2 & CA19-9 blood markers. *Sci. Transl. Med.* **2017**, *9*.
 24. Jenkinson, C.; Elliott, V.L.; Evans, A.; Oldfield, L.; Jenkins, R.E.; O'Brien, D.P.; Apostolidou, S.; Gentry-Maharaj, A.; Fourkala, E.O.; Jacobs, I.J.; et al. Decreased serum thrombospondin-1 levels in pancreatic cancer patients up to 24 months prior to clinical diagnosis: Association with diabetes mellitus. *Clin. Cancer Res.* **2016**, *22*, 1734–1743.
 25. Ryan, D.P.; Hong, T.S.; Bardeesy, N. Pancreatic Adenocarcinoma. *N. Engl. J. Med.* **2014**, *371*, 1039–1049.
 26. Neoptolemos, J.P.; Palmer, D.H.; Ghaneh, P.; Psarelli, E.E.; Valle, J.W.; Halloran, C.M.; Faluyi, O.; O'Reilly, D.A.; Cunningham, D.; Wadsley, J.; et al. Comparison of adjuvant gemcitabine and capecitabine with gemcitabine monotherapy in patients with resected pancreatic cancer (ESPAC-4): a multicentre, open-label, randomised, phase 3 trial. *Lancet* **2017**, *389*, 1011–1024.
 27. Nishio, K.; Kimura, K.; Amano, R.; Yamazoe, S.; Ohnira, G.; Nakata, B.; Hirakawa, K.; Ohira, M. Preoperative predictors for early recurrence of resectable pancreatic cancer. *World J. Surg. Oncol.* **2017**, *15*, 16.
 28. Egawa, S.; Takeda, K.; Fukuyama, S.; Motoi, F.; Sunamura, M.; Matsuno, S. Clinicopathological aspects of small pancreatic cancer. *Pancreas* **2004**, *28*, 235–240.
 29. Ariyama, J.; Suyama, M.; Satoh, K.; Sai, J. Imaging of small pancreatic ductal adenocarcinoma. *Pancreas* **1998**, *16*, 396–401.
 30. Chand, S.; O'Hayer, K.; Blanco, F.F.; Winter, J.M.; Brody, J.R. The landscape of pancreatic cancer therapeutic resistance mechanisms. *Int. J. Biol. Sci.* **2016**, *12*, 273.
 31. Yao, S.Y.; Ng, A.M.; Cass, C.E.; Baldwin, S.A.; Young, J.D. Nucleobase transport by human equilibrative nucleoside transporter 1 (hENT1). *J Biol*

- Chem* **2011**, *286*, 32552–32562.
32. Spratlin, J.; Sangha, R.; Glubrecht, D.; Dabbagh, L.; Young, J.D.; Dumontet, C.; Cass, C.; Lai, R.; Mackey, J.R. The absence of human equilibrative nucleoside transporter 1 is associated with reduced survival in patients with gemcitabine-treated pancreas adenocarcinoma. *Clin. Cancer Res.* **2004**, *10*, 6956–6961.
 33. Greenhalf, W.; Ghaneh, P.; Neoptolemos, J.P.; Palmer, D.H.; Cox, T.F.; Lamb, R.F.; Garner, E.; Campbell, F.; Mackey, J.R.; Costello, E.; et al. Pancreatic cancer hENT1 expression and survival from gemcitabine in patients from the ESPAC-3 trial. *J Natl Cancer Inst* **2014**, *106*, djt347.
 34. Fitzgerald, T.L.; Lertpiriyapong, K.; Cocco, L.; Martelli, A.M.; Libra, M.; Candido, S.; Montalto, G.; Cervello, M.; Steelman, L.; Abrams, S.L.; et al. Roles of EGFR and KRAS and their downstream signaling pathways in pancreatic cancer and pancreatic cancer stem cells. *Adv. Biol. Regul.* **2015**, *59*, 65–81.
 35. Moore, M.J.; Goldstein, D.; Hamm, J.; Figer, A.; Hecht, J.R.; Gallinger, S.; Au, H.J.; Murawa, P.; Walde, D.; Wolff, R.A.; et al. Erlotinib plus gemcitabine compared with gemcitabine alone in patients with advanced pancreatic cancer: a phase III trial of the National Cancer Institute of Canada Clinical Trials Group. *J. Clin. Oncol.* **2007**, *25*, 1960–1966.
 36. Von Hoff, D.D.; Ervin, T.; Arena, F.P.; Chiorean, E.G.; Infante, J.; Moore, M.; Seay, T.; Tjulandin, S.A.; Ma, W.W.; Saleh, M.N.; et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. *N. Engl. J. Med.* **2013**, *369*, 1691–1703.
 37. Thota, R.; Pauff, J.M.; Berlin, J.D. Treatment of metastatic pancreatic adenocarcinoma: a review. *Oncology (Williston Park)*. **2014**, *28*, 70–74.
 38. Conroy, T.; Desseigne, F.; Ychou, M.; Bouché, O.; Guimbaud, R.; Bécouarn, Y.; Adenis, A.; Raoul, J.-L.; Gourgou-Bourgade, S.; de la Fouchardière, C.; et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. *N. Engl. J. Med.* **2011**, *364*, 1817–1825.
 39. Katayama, E.S.; Hue, J.J.; Bajor, D.L.; Ocuin, L.M.; Ammori, J.B.; Hardacre, J.M.; Winter, J.M. A comprehensive analysis of clinical trials in pancreatic cancer: What is coming down the pike? *Oncotarget* **2020**, *11*, 3489.

40. Jones, S.; Zhang, X.; Parsons, D.W.; Lin, J.C.-H.C.-H.; Leary, R.J.; Angenendt, P.; Mankoo, P.; Carter, H.; Kamiyama, H.; Jimeno, A.; et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* (80-.). **2008**, *321*, 1801–1806.
41. Cicens, J.; Kvederaviciute, K.; Meskinyte, I.; Meskinyte-Kausiliene, E.; Skeberdyte, A. KRAS, TP53, CDKN2A, SMAD4, BRCA1, and BRCA2 mutations in pancreatic cancer. *Cancers (Basel)*. **2017**, *9*, 42.
42. Waddell, N.; Pajic, M.; Patch, A.-M.; Chang, D.K.; Kassahn, K.S.; Bailey, P.; Johns, A.L.; Miller, D.; Nones, K.; Quek, K. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature* **2015**, *518*, 495–501.
43. Biankin, A. V.; Waddell, N.; Kassahn, K.S.; Gingras, M.-C.; Muthuswamy, L.B.; Johns, A.L.; Miller, D.K.; Wilson, P.J.; Patch, A.-M.; Wu, J.; et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* **2012**, *491*, 399–405.
44. Witkiewicz, A.K.; McMillan, E.A.; Balaji, U.; Baek, G.; Lin, W.C.; Mansour, J.; Mollaei, M.; Wagner, K.U.; Koduru, P.; Yopp, A.; et al. Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nat Commun* **2015**, *6*, 6744.
45. Liu, P.; Wang, Y.; Li, X. Targeting the untarable KRAS in cancer therapy. *Acta Pharm. Sin. B* **2019**, *9*, 871–879.
46. Rachakonda, P.S.; Bauer, A.S.; Xie, H.; Campa, D.; Rizzato, C.; Canzian, F.; Beghelli, S.; Greenhalf, W.; Costello, E.; Schanne, M.; et al. Somatic mutations in exocrine pancreatic tumors: association with patient survival. *PLoS One* **2013**, *8*, e60870.
47. di Magliano, M.P.; Logsdon, C.D. Roles for KRAS in pancreatic tumor development and progression. *Gastroenterology* **2013**, *144*, 1220–1229.
48. Rowinsky, E.K.; Windle, J.J.; Von Hoff, D.D. Ras protein farnesyltransferase: A strategic target for anticancer therapeutic development. *J Clin Oncol* **1999**, *17*, 3631–3652.
49. Van Cutsem, E.; van de Velde, H.; Karasek, P.; Oettle, H.; Vervenne, W.L.; Szawlowski, A.; Schoffski, P.; Post, S.; Verslype, C.; Neumann, H.; et al. Phase III trial of gemcitabine plus tipifarnib compared with gemcitabine plus placebo in advanced pancreatic cancer. *J Clin Oncol* **2004**, *22*, 1430–1438.

50. Calhoun, E.S.; Kern, S.E. Molecular Genetics of Pancreatic Cancer. *Pancreat. Cancer* **2008**, 27–39.
51. Hong, S.-M.; Park, J.Y.; Hruban, R.H.; Goggins, M. Molecular signatures of pancreatic cancer. *Arch. Pathol. Lab. Med.* **2011**, 135, 716–727.
52. Hruban, R.H.; Iacobuzio-Donahue, C.; Wilentz, R.E.; Goggins, M.; Kern, S.E. Molecular pathology of pancreatic cancer. *Cancer J* **2001**, 7, 251–258.
53. Tatarian, T.; Winter, J.M. Genetics of Pancreatic Cancer and Its Implications on Therapy. *Surg. Clin. North Am.* **2016**, 96, 1207–1221.
54. Sirivatanauksorn, V.; Sirivatanauksorn, Y.; Lemoine, N.R. Molecular pattern of ductal pancreatic cancer. *Langenbecks Arch Surg* **1998**, 383, 105–115.
55. Al-Kaabi, A.; Van Bockel, L.W.; Pothén, A.J.; Willems, S.M. p16INK4A and p14ARF gene promoter hypermethylation as prognostic biomarker in oral and oropharyngeal squamous cell carcinoma: A review. *Dis. Markers* **2014**.
56. Bailey, J.M.; Hendley, A.M.; Lafaro, K.J.; Pruski, M.A.; Jones, N.C.; Alsina, J.; Younes, M.; Maitra, A.; McAllister, F.; Iacobuzio-Donahue, C.A. p53 mutations cooperate with oncogenic Kras to promote adenocarcinoma from pancreatic ductal cells. *Oncogene* **2016**, 35, 4282–4288.
57. Tokino, T.; Nakamura, Y. The role of p53-target genes in human cancer. *Crit Rev Oncol Hematol* **2000**, 33, 1–6.
58. Li, D.; Xie, K.; Wolff, R.; Abbruzzese, J.L. Pancreatic cancer. *Lancet* **2004**, 363, 1049–1057.
59. Rockland Immunochemicals, I. p53 Pathway Antibodies Available online: <https://rockland-inc.com/p53-pathway.aspx>.
60. Weissmueller, S.; Manchado, E.; Saborowski, M.; Morris IV, J.P.; Wagenblast, E.; Davis, C.A.; Moon, S.H.; Pfister, N.T.; Tschaharganeh, D.F.; Kitzing, T.; et al. Mutant p53 drives pancreatic cancer metastasis through cell-autonomous PDGF receptor β signaling. *Cell* **2014**, 157, 382–394.
61. Principe, D.R.; Doll, J.A.; Bauer, J.; Jung, B.; Munshi, H.G.; Bartholin, L.; Pasche, B.; Lee, C.; Grippo, P.J. TGF- β : Duality of function between tumor prevention and carcinogenesis. *J. Natl. Cancer Inst.* **2014**, 106.
62. Xia, X.; Wu, W.; Huang, C.; Cen, G.; Jiang, T.; Cao, J.; Huang, K.; Qiu, Z. SMAD4 and its role in pancreatic cancer. *Tumour Biol* **2015**, 36, 111–119.
63. Ahmed, S.; Bradshaw, A.D.; Gera, S.; Zahidunnabi Dewan, M.; Xu, R. The

- TGF- β /Smad4 signaling pathway in pancreatic carcinogenesis and its clinical significance. *J. Clin. Med.* **2017**, *6*, 5.
64. Pizzamiglio, L.; Focchi, E.; Antonucci, F. ATM Protein Kinase: Old and New Implications in Neuronal Pathways and Brain Circuitry. *Cells* **2020**, *9*, 1969.
 65. Moffitt, R.A.; Marayati, R.; Flate, E.L.; Volmar, K.E.; Loeza, S.G.H.; Hoadley, K.A.; Rashid, N.U.; Williams, L.A.; Eaton, S.C.; Chung, A.H.; et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* **2015**, *47*, 1168.
 66. Collisson, E.A.; Sadanandam, A.; Olson, P.; Gibb, W.J.; Truitt, M.; Gu, S.; Cooc, J.; Weinkle, J.; Kim, G.E.; Jakkula, L.; et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* **2011**, *17*, 500–503.
 67. Weinstein, I.B. Cancer: Addiction to oncogenes - The Achilles heel of cancer. *Science (80-.)*. **2002**, *297*, 63–64.
 68. Bailey, P.; Chang, D.K.; Nones, K.; Johns, A.L.; Patch, A.-M.M.; Gingras, M.-C.C.; Miller, D.K.; Christ, A.N.; Bruxner, T.J.C.C.; Quinn, M.C.; et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **2016**, *531*, 47–52.
 69. Aung, K.L.; Fischer, S.E.; Denroche, R.E.; Jang, G.H.; Dodd, A.; Creighton, S.; Southwood, B.; Liang, S. Ben; Chadwick, D.; Zhang, A.; et al. Genomics-driven precision medicine for advanced pancreatic cancer: Early results from the COMPASS trial. *Clin. Cancer Res.* **2018**, *26*, 1344–1354.
 70. Puleo, F.; Nicolle, R.; Blum, Y.; Cros, J.; Marisa, L.; Demetter, P.; Quertinmont, E.; Svrcek, M.; Elarouci, N.; Iovanna, J.; et al. Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. *Gastroenterology* **2018**, *155*, 1999–2013.
 71. Collisson, E.A.; Campbell, J.D.; Brooks, A.N.; Berger, A.H.; Lee, W.; Chmielecki, J.; Beer, D.G.; Cope, L.; Creighton, C.J.; Danilova, L.; et al. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* **2014**, *511*, 543–550.
 72. Nelson, S.R.; Walsh, N. Genetic alterations featuring biological models to tailor clinical management of pancreatic cancer patients. *Cancers (Basel)*. **2020**, *12*, 1233.

73. McGarrity, T.J.; Amos, C. Peutz-Jeghers syndrome: Clinicopathology and molecular alterations. *Cell. Mol. Life Sci.* **2006**, *63*, 2135–2144.
74. van Lier, M.G.; Westerman, A.M.; Wagner, A.; Looman, C.W.; Wilson, J.H.; de Rooij, F.W.; Lemmens, V.E.; Kuipers, E.J.; Mathus-Vliegen, E.M.; van Leerdam, M.E. High cancer risk and increased mortality in patients with Peutz-Jeghers syndrome. *Gut* **2011**, *60*, 141–147.
75. Resta, N.; Pierannunzio, D.; Lenato, G.M.; Stella, A.; Capocaccia, R.; Bagnulo, R.; Lastella, P.; Susca, F.C.; Bozzao, C.; Loconte, D.C.; et al. Cancer risk associated with STK11/LKB1 germline mutations in Peutz-Jeghers syndrome patients: results of an Italian multicenter study. *Dig Liver Dis* **2013**, *45*, 606–611.
76. Launonen, V. Mutations in the human LKB1/STK11 gene. *Hum. Mutat.* **2005**, *26*, 291–297.
77. Hezel, A.F.; Gurumurthy, S.; Granot, Z.; Swisa, A.; Chu, G.C.; Bailey, G.; Dor, Y.; Bardeesy, N.; DePinho, R.A. Pancreatic Lkb1 Deletion Leads to Acinar Polarity Defects and Cystic Neoplasms. *Mol. Cell. Biol.* **2008**, *28*, 2414–2425.
78. Morton, J.P.; Jamieson, N.B.; Karim, S.A.; Athineos, D.; Ridgway, R.A.; Nixon, C.; McKay, C.J.; Carter, R.; Brunton, V.G.; Frame, M.C.; et al. LKB1 haploinsufficiency cooperates with Kras to promote pancreatic cancer through suppression of p21-dependent growth arrest. *Gastroenterology* **2010**, *139*, 586–597.
79. Schabath, M.B.; Welsh, E.A.; Fulp, W.J.; Chen, L.; Teer, J.K.; Thompson, Z.J.; Engel, B.E.; Xie, M.; Berglund, A.E.; Creelan, B.C.; et al. Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene* **2016**, *35*, 3209–3216.
80. Whitcomb, D.C.; Gorry, M.C.; Preston, R.A.; Furey, W.; Sossenheimer, M.J.; Ulrich, C.D.; Martin, S.P.; Gates Jr., L.K.; Amann, S.T.; Toskes, P.P.; et al. Hereditary pancreatitis is caused by a mutation in the cationic trypsinogen gene. *Nat Genet* **1996**, *14*, 141–145.
81. Koziel, D.; Gluszek, S.; Kowalik, A.; Chlopek, M.; Pieciak, L. Genetic mutations in SPINK1, CFTR, CTSC genes in acute pancreatitis. *BMC Gastroenterol.* **2015**, *15*, 70.

82. Dhar, P.; Kalghatgi, S.; Saraf, V. Pancreatic Cancer in Chronic Pancreatitis. *Indian J. Surg. Oncol.* **2015**, *6*, 57–62.
83. Hussain, S.P.; Hofseth, L.J.; Harris, C.C. Radical causes of cancer. *Nat. Rev. Cancer* **2003**, *3*, 276–285.
84. Whitcomb, D.; Greer, J. Germ-Line Mutations, Pancreatic Inflammation, and Pancreatic Cancer. *Clin. Gastroenterol. Hepatol.* **2009**, *7*, S29–S34.
85. Apte, M. V.; Haber, P.S.; Darby, S.J.; Rodgers, S.C.; McCaughan, G.W.; Korsten, M.A.; Pirola, R.C.; Wilson, J.S. Pancreatic stellate cells are activated by proinflammatory cytokines: Implications for pancreatic fibrogenesis. *Gut* **1999**, *44*, 534–541.
86. Mews, P.; Phillips, P.; Fahmy, R.; Korsten, M.; Pirola, R.; Wilson, J.; Apte, M. Pancreatic stellate cells respond to inflammatory cytokines: Potential role in chronic pancreatitis. *Gut* **2002**, *50*, 535–541.
87. Kolodecik, T.; Shugrue, C.; Ashat, M.; Thrower, E.C. Risk factors for pancreatic cancer: Underlying mechanisms and potential targets. *Front. Physiol.* **2014**, *4*, 415.
88. Yadav, D.; Lowenfels, A.B. The epidemiology of pancreatitis and pancreatic cancer. *Gastroenterology* **2013**, *144*, 1252–1261.
89. Eckerle Mize, D.; Bishop, M.; Resse, E.; Sluzevich, J. Familial atypical multiple mole melanoma syndrome. In *Cancer syndromes.*; Bethesda, MD: National Center for Biotechnology Information (US), 2009.
90. Gruis, N.A.; van der Velden, P.A.; Sandkuijl, L.A.; Prins, D.E.; Weaver-Feldhaus, J.; Kamb, A.; Bergman, W.; Frants, R.R. Homozygotes for CDKN2 (p16) germline mutation in Dutch familial melanoma kindreds. *Nat Genet* **1995**, *10*, 351–353.
91. de Snoo, F.A.; Bishop, D.T.; Bergman, W.; van Leeuwen, I.; van der Drift, C.; van Nieuwpoort, F.A.; Out-Luiting, C.J.; Vasen, H.F.; ter Huurne, J.A.; Frants, R.R.; et al. Increased risk of cancer other than melanoma in CDKN2A founder mutation (p16-Leiden)-positive melanoma families. *Clin Cancer Res* **2008**, *14*, 7151–7157.
92. Vasen, H.; Ibrahim, I.; Ponce, C.G.; Slater, E.P.; Matthai, E.; Carrato, A.; Earl, J.; Robbers, K.; van Mil, A.M.; Potjer, T.; et al. Benefit of Surveillance for Pancreatic Cancer in High-Risk Individuals: Outcome of Long-Term

- Prospective Follow-Up Studies From Three European Expert Centers. *J Clin Oncol* **2016**, *34*, 2010–2019.
93. Kastrinos, F.; Mukherjee, B.; Tayob, N.; Wang, F.; Sparr, J.; Raymond, V.M.; Bandipalliam, P.; Stoffel, E.M.; Gruber, S.B.; Syngal, S. Risk of pancreatic cancer in families with Lynch syndrome. *JAMA - J. Am. Med. Assoc.* **2009**, *302*, 1790–1795.
 94. Li, G.M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* **2008**, *18*, 85–98.
 95. Lee, V.; Murphy, A.; Le, D.T.; Diaz, L.A. Mismatch Repair Deficiency and Response to Immune Checkpoint Blockade. *Oncologist* **2016**, *21*, 1200.
 96. Eso, Y.; Shimizu, T.; Takeda, H.; Takai, A.; Marusawa, H. Microsatellite instability and immune checkpoint inhibitors: toward precision medicine against gastrointestinal and hepatobiliary cancers. *J. Gastroenterol.* **2020**, *55*, 15–26.
 97. Kastrinos, F.; Steyerberg, E.W.; Mercado, R.; Balmana, J.; Holter, S.; Gallinger, S.; Siegmund, K.D.; Church, J.M.; Jenkins, M.A.; Lindor, N.M.; et al. The PREMM(1,2,6) model predicts risk of MLH1, MSH2, and MSH6 germline mutations based on cancer history. *Gastroenterology* **2011**, *140*, 73–81.
 98. Breast Cancer Linkage, C. Cancer risks in BRCA2 mutation carriers. *J. Natl. Cancer Inst.* **1999**, *91*, 1310–1316.
 99. Efthimiou, E.; Crnogorac-Jurcevic, T.; Lemoine, N.R.; Brentnall, T.A. Inherited predisposition to pancreatic cancer. *Gut* **2001**, *48*, 143–147.
 100. Jones, S.; Hruban, R.H.; Kamiyama, M.; Borges, M.; Zhang, X.; Parsons, D.W.; Lin, J.C.-H.; Palmisano, E.; Brune, K.; Jaffee, E.M.; et al. Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* (80-.). **2009**, *324*, 217.
 101. Slater, E.P.; Langer, P.; Niemczyk, E.; Strauch, K.; Butler, J.; Habbe, N.; Neoptolemos, J.P.; Greenhalf, W.; Bartsch, D.K. PALB2 mutations in European familial pancreatic cancer families. *Clin Genet* **2010**, *78*, 490–494.
 102. Tomasz, M. Mitomycin C: small, fast and deadly (but very selective). *Chem. Biol.* **1995**, *2*, 575–579.
 103. Villarroel, M.C.; Rajeshkumar, N. V; Garrido-Laguna, I.; De Jesus-Acosta, A.;

- Jones, S.; Maitra, A.; Hruban, R.H.; Eshleman, J.R.; Klein, A.; Laheru, D.; et al. Personalizing cancer treatment in the age of global genomic analyses: PALB2 gene mutations and the response to DNA damaging agents in pancreatic cancer. *Mol. Cancer Ther.* **2011**, *10*, 3–8.
104. Gupta, C.; Mazzara, P.F. High-grade pancreatic intraepithelial neoplasia in a patient with familial adenomatous polyposis. *Arch. Pathol. Lab. Med.* **2005**, *129*, 1398–1400.
 105. Moussata, D.; Senouci, L.; Berger, F.; Scoazec, J.Y.; Pinson, S.; Walter, T.; Lombard-Bohas, C.; Saurin, J.C. Familial adenomatous polyposis and pancreatic cancer. *Pancreas* **2015**, *44*, 512–513.
 106. Matsumoto, T.; Lida, M.; Kabori, Y.; Mizuno, M.; Nakamura, S.; Hizawa, K.; Yao, T. Genetic predisposition to clinical manifestations in familial adenomatous polyposis with special reference to duodenal lesions. *Am J Gastroenterol* **2002**, *97*, 180–185.
 107. Schneikert, J.; Behrens, J. The canonical Wnt signalling pathway and its APC partner in colon cancer development. *Gut* **2007**, *56*, 417–425.
 108. Oshima, M.; Oshima, H.; Kitagawa, K.; Kobayashi, M.; Itakura, C.; Taketo, M. Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *95*, 4482–4486.
 109. Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **2019**, *20*, 467–484.
 110. Bush, W.S.; Moore, J.H. Genome-wide association studies. *PLoS Comput. Biol.* **2012**, *8*, e1002822.
 111. Robert, F.; Pelletier, J. Exploring the Impact of Single-Nucleotide Polymorphisms on Translation. *Front. Genet.* **2018**, *9*, 507.
 112. Amundadottir, L.; Kraft, P.; Stolzenberg-Solomon, R.Z.; Fuchs, C.S.; Petersen, G.M.; Arslan, A.A.; Bueno-de-Mesquita, H.B.; Gross, M.; Helzlsouer, K.; Jacobs, E.J.; et al. Genome-wide association study identifies variants in the ABO locus associated with susceptibility to pancreatic cancer. *Nat Genet* **2009**, *41*, 986–990.
 113. Petersen, G.M.; Amundadottir, L.; Fuchs, C.S.; Kraft, P.; Stolzenberg-

- Solomon, R.Z.; Jacobs, K.B.; Arslan, A.A.; Bueno-de-Mesquita, H.B.; Gallinger, S.; Gross, M.; et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat Genet* **2010**, *42*, 224–228.
114. Wolpin, B.M.; Rizzato, C.; Kraft, P.; Kooperberg, C.; Petersen, G.M.; Wang, Z.; Arslan, A.A.; Beane-Freeman, L.; Bracci, P.M.; Buring, J.; et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. *Nat. Genet.* **2014**, *46*, 994–1000.
115. Childs, E.J.; Mocci, E.; Campa, D.; Bracci, P.M.; Gallinger, S.; Goggins, M.; Li, D.; Neale, R.E.; Olson, S.H.; Scelo, G. - Common variation at 2p13. 3, 3q29, 7p13 and 17q25. 1 associated with susceptibility to pancreatic cancer. **2015**, *47*, 911–916.
116. Klein, A.P.; Wolpin, B.M.; Risch, H.A.; Stolzenberg-Solomon, R.Z.; Mocci, E.; Zhang, M.; Canzian, F.; Childs, E.J.; Hoskins, J.W.; Jermusyk, A.; et al. Genome-wide meta-analysis identifies five new susceptibility loci for pancreatic cancer. *Nat. Commun.* **2018**, *9*, 1–11.
117. Zhang, M.; Wang, Z.; Obazee, O.; Jia, J.; Childs, E.J.; Hoskins, J.; Figlioli, G.; Mocci, E.; Collins, I.; Chung, C.C.; et al. Three new pancreatic cancer susceptibility signals identified on chromosomes 1q32.1, 5p15.33 and 8q24.21. *Oncotarget* **2016**, *7*, 66328.
118. Low, S.K.; Kuchiba, A.; Zembutsu, H.; Saito, A.; Takahashi, A.; Kubo, M.; Daigo, Y.; Kamatani, N.; Chiku, S.; Totsuka, H.; et al. Genome-wide association study of pancreatic cancer in Japanese population. *PLoS One* **2010**, *5*, e11824.
119. Wu, C.; Miao, X.; Huang, L.; Che, X.; Jiang, G.; Yu, D.; Yang, X.; Cao, G.; Hu, Z.; Zhou, Y.; et al. Genome-wide association study identifies five loci associated with susceptibility to pancreatic cancer in Chinese populations. *Nat Genet* **2011**, *44*, 62–66.
120. Ramanan, V.K.; Shen, L.; Moore, J.H.; Saykin, A.J. Pathway analysis of genomic data: Concepts, methods, and prospects for future development. *Trends Genet.* **2012**, *28*, 323–332.
121. Li, D.; Duell, E.J.; Yu, K.; Risch, H.A.; Olson, S.H.; Kooperberg, C.; Wolpin, B.M.; Jiao, L.; Dong, X.; Wheeler, B.; et al. Pathway analysis of genome-wide

- association study data highlights pancreatic development genes as susceptibility factors for pancreatic cancer. *Carcinogenesis* **2012**, *33*, 1384–1390.
122. Walsh, N.; Zhang, H.; Hyland, P.P.L.; Yang, Q.; Mocci, E.; Zhang, M.; Childs, E.J.; Collins, I.; Wang, Z.; Arslan, A.A.; et al. Agnostic pathway/gene set analysis of genome- wide association data identifies associations for pancreatic cancer. *J. Natl. Cancer Inst.* **2018**, *111*, 557–567.
 123. Permert, J.; Ihse, I.; Jorfeldt, L.; von Schenck, H.; Arnqvist, H.J.; Larsson, J. Pancreatic cancer is associated with impaired glucose metabolism. *Eur. J. Surg.* **1993**, *159*, 101–107.
 124. Katsarou, A.; Gudbjörnsdottir, S.; Rawshani, A.; Dabelea, D.; Bonifacio, E.; Anderson, B.J.; Jacobsen, L.M.; Schatz, D.A.; Lernmark, A. Type 1 diabetes mellitus. *Nat. Rev. Dis. Prim.* **2017**, *3*, 1–17.
 125. DeFronzo, R.A. Insulin resistance: A multifaceted syndrome responsible for NIDDM, obesity, hypertension, dyslipidaemia and atherosclerosis. *Neth. J. Med.* **1997**, *14*, 173–194.
 126. Genkinger, J.M.; Spiegelman, D.; Anderson, K.E.; Bernstein, L.; Van Den Brandt, P.A.; Calle, E.E.; English, D.R.; Folsom, A.R.; Freudenheim, J.L.; Fuchs, C.S.; et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. *Int. J. Cancer* **2011**, *129*, 1708–1717.
 127. Everhart, J.; Wright, D. Diabetes Mellitus as a Risk Factor for Pancreatic Cancer: A Meta-Analysis. *JAMA J. Am. Med. Assoc.* **1995**, *273*, 1605–1609.
 128. Ben, Q.; Xu, M.; Ning, X.; Liu, J.; Hong, S.; Huang, W.; Zhang, H.; Li, Z. Diabetes mellitus and risk of pancreatic cancer: A meta-analysis of cohort studies. *Eur. J. Cancer* **2011**, *47*, 1928–1937.
 129. Hart, P.A.; Bellin, M.D.; Andersen, D.K.; Bradley, D.; Cruz-Monserrate, Z.; Forsmark, C.E.; Goodarzi, M.O.; Habtezion, A.; Korc, M.; Kudva, Y.C.; et al. Type 3c (pancreatogenic) diabetes mellitus secondary to chronic pancreatitis and pancreatic cancer. *Lancet Gastroenterol. Hepatol.* **2016**, *1*, 226–237.
 130. Amin, S.; Mhango, G.; Lin, J.; Aronson, A.; Wisnivesky, J.; Boffetta, P.; Lucas, A.L. Metformin Improves Survival in Patients with Pancreatic Ductal Adenocarcinoma and Pre-Existing Diabetes: A Propensity Score Analysis.

- Am. J. Gastroenterol.* **2016**, *111*, 1350.
131. Li, D.; Yeung, S.J.; Hassan, M.M.; Konopleva, M.; Abbruzzese, J.L. Antidiabetic Therapies Affect Risk of Pancreatic Cancer. *Gastroenterology* **2009**, *137*, 482–488.
 132. Singh, S.; Singh, P.P.; Singh, A.G.; Murad, M.H.; McWilliams, R.R.; Chari, S.T. Anti-diabetic medications and risk of pancreatic cancer in patients with diabetes mellitus: a systematic review and meta-analysis. *Am. J. Gastroenterol.* **2013**, *108*, 510–519.
 133. Pihoker, C.; Gilliam, L.K.; Ellard, S.; Dabelea, D.; Davis, C.; Dolan, L.M.; Greenbaum, C.J.; Imperatore, G.; Lawrence, J.M.; Marcovina, S.M.; et al. Prevalence, characteristics and clinical diagnosis of maturity onset diabetes of the young due to mutations in HNF1A, HNF4A, and glucokinase: Results from the SEARCH for diabetes in Youth. *J. Clin. Endocrinol. Metab.* **2013**, *98*, 4055–4062.
 134. Bishay, R.H.; Greenfield, J.R. A review of maturity onset diabetes of the young (MODY) and challenges in the management of glucokinase-MODY. *Med. J. Aust.* **2016**, *205*, 480–485.
 135. Gardner, D.S.; Tai, E.S. Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy Clinical features and treatment of maturity onset diabetes of the young (MODY). *Diabetes, Metab. Syndr. Obes. Targets Ther.* **2012**, *5*, 101.
 136. Radha, V.; Mohan, V. Genetic basis of monogenic diabetes. *Curr. Sci.* **2017**, 1277–1286.
 137. Gardner, D.S.L.; Shyong Tai, E. Clinical features and treatment of maturity onset diabetes of the young (MODY). *Diabetes, Metab. Syndr. Obes. Targets Ther.* **2012**, *5*, 101.
 138. Shields, B.M.; Hicks, S.; Shepherd, M.H.; Colclough, K.; Hattersley, A.T.; Ellard, S. Maturity-onset diabetes of the young (MODY): How many cases are we missing? *Diabetologia* **2010**, *53*, 2504–2508.
 139. Odom, D.T.; Zizlsperger, H.; Gordon, D.B.; Bell, G.W.; Rinaldi, N.J.; Murray, H.L.; Volkert, T.L.; Schreiber, J.; Rolfe, P.A.; Gifford, D.K.; et al. Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. *Science* (80-.). **2004**, *303*, 1378–1381.
 140. Sever, R.; Glass, C.K. Signaling by nuclear receptors. *Cold Spring Harb.*

- Perspect. Biol.* **2013**, *5*, a016709.
141. Boj, S.F.; Parrizas, M.; Maestro, M.A.; Ferrer, J. A transcription factor regulatory circuit in differentiated pancreatic cells. *Proc. Natl. Acad. Sci.* **2001**, *98*, 14481–14486.
 142. Brunton, H.; Caligiuri, G.; Cunningham, R.; Upstill-Goddard, R.; Bailey, U.M.; Garner, I.M.; Nourse, C.; Dreyer, S.; Jones, M.; Moran-Jones, K.; et al. HNF4A and GATA6 Loss Reveals Therapeutically Actionable Subtypes in Pancreatic Cancer. *Cell Rep.* **2020**, *31*, 107625.
 143. Bonfig, W.; Hermanns, S.; Warncke, K.; Eder, G.; Engelsberger, I.; Burdach, S.; Ziegler, A.G.; Lohse, P. GCK-MODY (MODY 2) Caused by a Novel p.Phe330Ser Mutation. *ISRN Pediatr.* **2011**, *2011*.
 144. Luo, Z.; Li, Y.; Wang, H.; Fleming, J.; Li, M.; Kang, Y.; Zhang, R.; Li, D. Hepatocyte Nuclear Factor 1A (HNF1A) as a possible tumor suppressor in pancreatic cancer. *PLoS One* **2015**, *10*, e0121082.
 145. Luni, C.; Marth, J.D.; Doyle, F.J. Computational Modeling of Glucose Transport in Pancreatic β -Cells Identifies Metabolic Thresholds and Therapeutic Targets in Diabetes. *PLoS One* **2012**, *7*, e53130.
 146. Johnson, J.D.; Ahmed, N.T.; Luciani, D.S.; Han, Z.; Tran, H.; Fujita, J.; Misler, S.; Edlund, H.; Polonsky, K.S. Increased islet apoptosis in Pdx1^{+/-} mice. *J. Clin. Invest.* **2003**, *111*, 1147–1160.
 147. Gauthier, B.R.; Brun, T.; Sarret, E.J.; Ishihara, H.; Schaad, O.; Descombes, P.; Wollheim, C.B. Oligonucleotide microarray analysis reveals PDX1 as an essential regulator of mitochondrial metabolism in rat islets. *J. Biol. Chem.* **2004**, *279*, 31121–31130.
 148. Brissova, M.; Shiota, M.; Nicholson, W.E.; Gannon, M.; Knobel, S.M.; Piston, D.W.; Wright, C.V.E.; Powers, A.C. Reduction in pancreatic transcription factor PDX-1 impairs glucose-stimulated insulin secretion. *J. Biol. Chem.* **2002**, *277*, 11225–11232.
 149. Park, J.Y.; Hong, S.M.; Klimstra, D.S.; Goggins, M.G.; Maitra, A.; Hruban, R.H. Pdx1 expression in pancreatic precursor lesions and neoplasms. *Appl. Immunohistochem. Mol. Morphol.* **2011**, *19*, 444.
 150. Liu, S.; Ballian, N.; Belaguli, N.S.; Patel, S.; Li, M.; Templeton, N.S.; Gingras, M.C.; Gibbs, R.; Fisher, W.; Brunnicardi, F.C. PDX-1 acts as a potential

- molecular target for treatment of human pancreatic cancer. *Pancreas* **2008**, 37, 210–220.
151. Annicotte, J.S.; Fayard, E.; Swift, G.H.; Selander, L.; Edlund, H.; Tanaka, T.; Kodama, T.; Schoonjans, K.; Auwerx, J. Pancreatic-duodenal homeobox 1 regulates expression of liver receptor homolog 1 during pancreas development. *Mol. Cell. Biol.* **2003**, 23, 6713–6724.
 152. Holmstrom, S.R.; Deering, T.; Swift, G.H.; Poelwijk, F.J.; Mangelsdorf, D.J.; Kliewer, S.A.; Macdonald, R.J. LRH-1 and PTF1-L coregulate an exocrine pancreas-specific transcriptional network for digestive function. *Genes Dev.* **2011**, 25, 1674–1679.
 153. De La O, J.-P.; Emerson, L.L.; Goodman, J.L.; Froebe, S.C.; Illum, B.E.; Curtis, A.B.; Murtaugh, L.C. Notch and Kras reprogram pancreatic acinar cells to ductal intraepithelial neoplasia. *Proc. Natl. Acad. Sci.* **2008**, 105, 18907–18912.
 154. Charles Murtaugh, L. Putting GWAS to the functional test: NR5A2 and pancreatic cancer risk. *Gut* **2014**, 63, 535–536.
 155. Bingham, C.; Hattersley, A.T. Renal cysts and diabetes syndrome resulting from mutations in hepatocyte nuclear factor-1 β . *Nephrol. Dial. Transplant.* **2004**, 19, 2703–2708.
 156. Bellanne-Chantelot, C.; Clauin, S.; Chauveau, D.; Collin, P.; Daumont, M.; Douillard, C.; Dubois-Laforgue, D.; Dusselier, L.; Gautier, J.F.; Jadoul, M.; et al. Large genomic rearrangements in the hepatocyte nuclear factor-1 β (TCF2) gene are the most frequent cause of maturity-onset diabetes of the young type 5. *Diabetes* **2005**, 54, 3126–3132.
 157. Hu, Y.L.; Zhong, D.; Pang, F.; Ning, Q.Y.; Zhang, Y.Y.; Li, G.; Wu, J.Z.; Mo, Z.N. HNF1 β is involved in prostate cancer risk via modulating androgenic hormone effects and coordination with other genes. *Genet. Mol. Res.* **2013**, 12, 1327–1335.
 158. Shen, H.; Fridley, B.L.; Song, H.; Lawrenson, K.; Cunningham, J.M.; Ramus, S.J.; Cicek, M.S.; Tyrer, J.; Stram, D.; Larson, M.C.; et al. Epigenetic analysis leads to identification of HNF1B as a subtype-specific susceptibility gene for ovarian cancer. *Nat. Commun.* **2013**, 4, 1–10.
 159. Cereghini, S. Liver-enriched transcription factors and hepatocyte

- differentiation. *FASEB J.* **1996**, *10*, 267–282.
160. Lau, H.H.; Ng, N.H.J.; Loo, L.S.W.; Jasmen, J.B.; Teo, A.K.K. The molecular functions of hepatocyte nuclear factors – In and beyond the liver. *J. Hepatol.* **2018**, *68*, 1033–1048.
 161. De Simone, V.; De Magistris, L.; Lazzaro, D.; Gerstner, J.; Monaci, P.; Nicosia, A.; Cortese, R. LFB3, a heterodimer-forming homeoprotein of the LFB1 family, is expressed in specialized epithelia. *EMBO J.* **1991**, *10*.
 162. Rey-Campos, J.; Chouard, T.; Yaniv, M.; Cereghini, S. vHNF1 is a homeoprotein that activates transcription and forms heterodimers with HNF1. *EMBO J.* **1991**, *10*, 1445–1457.
 163. Bach, I.; Yaniv, M. More potent transcriptional activators or a transdominant inhibitor of the HNF1 homeoprotein family are generated by alternative RNA processing. *EMBO J.* **1993**, *12*, 4229–4242.
 164. Harries, L.W.; Ellard, S.; Stride, A.; Morgan, N.G.; Hattersley, A.T.; Vaxillaire, M.; Tuomi, T.; Barbetti, E.; Njolstad, P.R.; Hansen, T.; et al. Isomers of the TCF1 gene encoding hepatocyte nuclear factor-1 alpha show differential expression in the pancreas and define the relationship between mutation position and clinical phenotype in monogenic diabetes. *Hum. Mol. Genet.* **2006**, *15*, 2216–2224.
 165. Lee, Y.-H.; Sauer, B.; Gonzalez, F.J. Laron Dwarfism and Non-Insulin-Dependent Diabetes Mellitus in the Hnf-1 α Knockout Mouse. *Mol. Cell. Biol.* **1998**, *18*, 3059–3068.
 166. Barbacci, E.; Reber, M.; Ott, M.O.; Breillat, C.; Huetz, F.; Cereghini, S. Variant Hepatocyte Nuclear Factor 1 is required for visceral endoderm specification. *Development* **1999**, *126*, 4795–4805.
 167. Kalisz, M.; Bernardo, E.; Beucher, A.; Maestro, M.A.; del Pozo, N.; Millán, I.; Haeberle, L.; Schlensog, M.; Safi, S.A.; Knoefel, W.T.; et al. HNF1A recruits KDM6A to activate differentiated acinar cell programs that suppress pancreatic cancer. *EMBO J.* **2020**.
 168. Hoskins, J.W.; Jia, J.; Flandez, M.; Parikh, H.; Xiao, W.; Collins, I.; Emmanuel, M.A.; Ibrahim, A.; Powell, J.; Zhang, L.; et al. Transcriptome analysis of pancreatic cancer reveals a tumor suppressor function for HNF1A. *Carcinogenesis* **2014**, *35*, 2670–2678.

-
169. Yu, Y.; Liang, S.; Zhou, Y.; Li, S.; Li, Y.; Liao, W. HNF1A/CASC2 regulates pancreatic cancer cell proliferation through PTEN/Akt signaling. *J. Cell. Biochem.* **2019**, *120*, 2816–2827.
170. Abel, E. V.; Goto, M.; Magnuson, B.; Abraham, S.; Ramanathan, N.; Hotaling, E.; Alaniz, A.A.; Kumar-Sinha, C.; Dziubinski, M.L.; Urs, S.; et al. HNF1A is a novel oncogene that regulates human pancreatic cancer stem cell properties. *Elife* **2018**, *7*, e33947.
171. Lu, Y.; Xu, D.; Peng, J.; Luo, Z.; Chen, C.; Chen, Y.; Chen, H.; Zheng, M.; Yin, P.; Wang, Z. HNF1A inhibition induces the resistance of pancreatic cancer cells to gemcitabine by targeting ABCB1. *EBioMedicine* **2019**, *44*, 403–418.
172. Subramani, R.; Medel, J.; Flores, K.; Perry, C.; Galvez, A.; Sandoval, M.; Rivera, S.; Pedroza, D.A.; Penner, E.; Chitti, M.; et al. Hepatocyte nuclear factor 1 alpha influences pancreatic cancer growth and metastasis. *Sci. Rep.* **2020**, *10*, 20225.
173. Nie, C.; Wang, B.B.; Wang, B.B.; Lv, N.; Zhang, E. Integrative analysis of hnf1b mrna in human cancers based on data mining. *Int. J. Med. Sci.* **2020**, *17*, 2895–2904.
174. Quilichini, E.; Fabre, M.; Dirami, T.; Stedman, A.; De Vas, M.; Ozguc, O.; Pasek, R.C.; Cereghini, S.; Morillon, L.; Guerra, C.; et al. Pancreatic Ductal Deletion of Hnf1b Disrupts Exocrine Homeostasis, Leads to Pancreatitis, and Facilitates Tumorigenesis. *Cell. Mol. Gastroenterol. Hepatol.* **2019**, *8*, 487–511.
175. Dietlein, F.; Thelen, L.; Reinhardt, H.C. Cancer-specific defects in DNA repair pathways as targets for personalized therapeutic approaches. *Trends Genet.* **2014**, *30*, 326–339.
176. Jackson, S.P.; Bartek, J. The DNA-damage response in human biology and disease. *Nature* **2010**, *461*, 1071–1078.
177. Reinhardt, H.C.; Schumacher, B. The p53 network: Cellular and systemic DNA damage responses in aging and cancer. *Trends Genet.* **2012**, *28*, 128–136.
178. R&D Systems DNA Damage Response Available online: <https://www.rndsystems.com/resources/articles/dna-damage-response>.
179. Petersen, G.M.; Chaffee, K.G.; McWilliams, R.R.; Majithia, N.; Allen, B.; Kidd,

- J.; Singh, N.; Hartman, A.-R.; Oberg, A.L. Genetic heterogeneity and survival among pancreatic adenocarcinoma (PDAC) patients with positive family history. *J. Clin. Oncol.* **2016**, 4108–4108.
180. Golan, T.; Oh, D.-Y.; Reni, M.; Macarulla, T.M.; Tortora, G.; Hall, M.J.; Reinacher-Schick, A.C.; Borg, C.; Hochhauser, D.; Walter, T.; et al. POLO: A randomized phase III trial of olaparib maintenance monotherapy in patients (pts) with metastatic pancreatic cancer (mPC) who have a germline BRCA1/2 mutation (gBRCAm). *J. Clin. Oncol.* **2016**, TPS4152–TPS4152.
181. Golan, T.; Hammel, P.; Reni, M.; Van Cutsem, E.; Macarulla, T.; Hall, M.J.; Park, J.-O.; Hochhauser, D.; Arnold, D.; Oh, D.-Y.; et al. Maintenance Olaparib for Germline BRCA -Mutated Metastatic Pancreatic Cancer. *N. Engl. J. Med.* **2019**, 381, 317–327.
182. Hanahan, D.; Weinberg, R.A. The hallmarks of cancer. *Cell* **2000**, 100, 57–70.
183. Taylor, R.C.; Cullen, S.P.; Martin, S.J. Apoptosis: Controlled demolition at the cellular level. *Nat. Rev. Mol. Cell Biol.* **2008**, 9, 231–241.
184. Martinon, F.; Tschopp, J. Inflammatory caspases: Linking an intracellular innate immune system to autoinflammatory diseases. *Cell* **2004**, 117, 561–574.
185. Jäger, R.; Zwacka, R.M. The enigmatic roles of caspases in tumor development. *Cancers (Basel)*. **2010**, 2, 1952–1979.
186. Pujana, M.A.; Han, J.D.J.; Starita, L.M.; Stevens, K.N.; Tewari, M.; Ahn, J.S.; Rennert, G.; Moreno, V.; Kirchhoff, T.; Gold, B.; et al. Network modeling links breast cancer susceptibility and centrosome dysfunction. *Nat. Genet.* **2007**, 39, 1338–1349.
187. Lucey, B.P.; Nelson-Rees, W.A.; Hutchins, G.M. Henrietta Lacks, HeLa cells, and cell culture contamination. *Arch. Pathol. Lab. Med.* **2009**, 133, 1463–1467.
188. Nelson-Rees, W.A.; Owens, R.B.; Arnstein, P.; Kniazeff, A.J. Source, alterations, characteristics and use of a new dog cell line (Cf2Th). *In Vitro* **1976**, 12, 665–669.
189. Goodspeed, A.; Heiser, L.M.; Gray, J.W.; Costello, J.C. Tumor-Derived Cell Lines as Molecular Models of Cancer Pharmacogenomics. *Mol. Cancer Res.*

- 2016, 14, 3–13.
190. Deer, E.L.; Gonzalez-Hernandez, J.; Coursen, J.D.; Shea, J.E.; Ngatia, J.; Scaife, C.L.; Firpo, M. a.; Mulvihill, S.J.; González-Hernández, J.; Coursen, J.D.; et al. Phenotype and Genotype of Pancreatic Cancer Cell Lines. *Pancreas* **2010**, 39, 425–35.
 191. Ben-David, U.; Siranosian, B.; Ha, G.; Tang, H.; Oren, Y.; Hinohara, K.; Strathdee, C.A.; Dempster, J.; Lyons, N.J.; Burns, R.; et al. Genetic and transcriptional evolution alters cancer cell line drug response. *Nature* **2018**, 560, 325–330.
 192. Boonstra, J.J.; van Marion, R.; Beer, D.G.; Lin, L.; Chaves, P.; Ribeiro, C.; Pereira, A.D.; Roque, L.; Darnton, S.J.; Altorki, N.K.; et al. Verification and Unmasking of Widely Used Human Esophageal Adenocarcinoma Cell Lines. *JNCI J. Natl. Cancer Inst.* **2010**, 102, 271–274.
 193. Horbach, S.P.J.M.; Halffman, W. The ghosts of HeLa: How cell line misidentification contaminates the scientific literature. *PLoS One* **2017**, 12, e0186281.
 194. Oldroyd, N.J.; Urquhart, A.J.; Kimpton, C.P.; Millican, E.S.; Watson, S.K.; Downes, T.; Gill, P.D. A highly discriminating octoplex short tandem repeat polymerase chain reaction system suitable for human individual identification. *Electrophoresis* **1995**, 16, 334–7.
 195. Reid, Y.; Storts, D.; Riss, T.; Minor, L. Authentication of human cell lines by STR DNA profiling analysis. In *Assay Guidance Manual [Internet]*; Eli Lilly & Company and the National Center for Advancing Translational Sciences, 2013.
 196. Masters, J.R.; Thomson, J.A.; Daly-Burns, B.; Reid, Y.A.; Dirks, W.G.; Packer, P.; Toji, L.H.; Ohno, T.; Tanabe, H.; Arlett, C.F.; et al. Short tandem repeat profiling provides an international reference standard for human cell lines. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 8012–7.
 197. Janik, K.; Popeda, M.; Peciak, J.; Rosiak, K.; Smolarz, M.; Treda, C.; Rieske, P.; Stoczynska-Fidelus, E.; Ksiazkiewicz, M. Efficient and simple approach to in vitro culture of primary epithelial cancer cells. *Biosci. Rep.* **2016**, 36.
 198. Zieba, J.; Ksiazkiewicz, M.; Janik, K.; Banaszczyk, M.; Peciak, J.; Piaskowski, S.; Lipinski, M.; Olczak, M.; Stoczynska-Fidelus, E.; Rieske, P. Sensitivity of

- neoplastic cells to senescence unveiled under standard cell culture conditions. *Anticancer Res.* **2015**, 35, 2759–68.
199. Kodack, D.P.; Farago, A.F.; Dastur, A.; Held, M.A.; Dardaei, L.; Friboulet, L.; von Flotow, F.; Damon, L.J.; Lee, D.; Parks, M.; et al. Primary Patient-Derived Cancer Cells and Their Potential for Personalized Cancer Patient Care. *Cell Rep.* **2017**, 21, 3298–3309.
 200. Jaeyun Jung, Hyang Sook Seol, S.C. The Generation and Application of Patient-Derived Xenograft Model for Cancer Research. *Cancer Res. Treat.* **2018**, 50, 1.
 201. Garrido-Laguna, I.; Uson, M.; Rajeshkumar, N. V.; Tan, A.C.; de Oliveira, E.; Karikari, C.; Villaroel, M.C.; Salomon, A.; Taylor, G.; Sharma, R.; et al. Tumor Engraftment in Nude Mice and Enrichment in Stroma- Related Gene Pathways Predict Poor Survival and Resistance to Gemcitabine in Patients with Pancreatic Cancer. *Clin. Cancer Res.* **2011**, 17, 5793–5800.
 202. Jimeno, A.; Rubio-Viqueira, B.; Rajeshkumar, N. V; Chan, A.; Solomon, A.; Hidalgo, M. A Fine-Needle Aspirate-Based Vulnerability Assay Identifies Polo-Like Kinase 1 as a Mediator of Gemcitabine Resistance in Pancreatic Cancer. *Mol. Cancer Ther.* **2010**, 9, 311–318.
 203. Festing, S.; Wilkinson, R. The ethics of animal research. Talking Point on the use of animals in scientific research. *EMBO Rep.* **2007**, 8, 526–530.
 204. Herreros-Villanueva, M.; Hijona, E.; Cosme, A.; Bujanda, L. Mouse models of pancreatic cancer. *World J. Gastroenterol.* **2012**, 18, 1286.
 205. Schneeberger, V.E.; Allaj, V.; Gardner, E.E.; Poirier, J.T.; Rudin, C.M. Quantitation of murine stroma and selective purification of the human tumor component of patient-derived xenografts for genomic analysis. *PLoS One* **2016**, 11, e0160587.
 206. Weiswald, L.B.; Bellet, D.; Dangles-Marie, V. Spherical cancer models in tumor biology. *Neoplasia* **2015**, 17, 1–15.
 207. Thoma, C.R.; Zimmermann, M.; Agarkova, I.; Kelm, J.M.; Krek, W. 3D cell culture systems modeling tumor growth determinants in cancer target discovery. *Adv. Drug Deliv. Rev.* **2014**, 69, 29–41.
 208. Zanoni, M.; Piccinini, F.; Arienti, C.; Zamagni, A.; Santi, S.; Polico, R.; Bevilacqua, A.; Tesei, A. 3D tumor spheroid models for in vitro therapeutic

- screening: a systematic approach to enhance the biological relevance of data obtained. *Sci. Rep.* **2016**, *6*, 19103.
209. Kim, J. Bin Three-dimensional tissue culture models in cancer biology. *Semin. Cancer Biol.* **2005**, *15*, 365–377.
 210. Smith, S.J.; Wilson, M.; Ward, J.H.; Rahman, C. V.; Peet, A.C.; Macarthur, D.C.; Rose, F.R.A.J.; Grundy, R.G.; Rahman, R. Recapitulation of Tumor Heterogeneity and Molecular Signatures in a 3D Brain Cancer Model with Decreased Sensitivity to Histone Deacetylase Inhibition. *PLoS One* **2012**, *7*.
 211. Loessner, D.; Stok, K.S.; Lutolf, M.P.; Hutmacher, D.W.; Clements, J.A.; Rizzi, S.C. Bioengineered 3D platform to explore cell-ECM interactions and drug resistance of epithelial ovarian cancer cells. *Biomaterials* **2010**, *31*, 8494–8506.
 212. Shaw, K.R.; Wrobel, C.N.; Brugge, J.S. Use of three-dimensional basement membrane cultures to model oncogene-induced changes in mammary epithelial morphogenesis. *J. Mammary Gland Biol. Neoplasia* **2004**, *9*, 297–310.
 213. Anton, D.; Burckel, H.; Josset, E.; Noel, G. Three-dimensional cell culture: A breakthrough in vivo. *Int. J. Mol. Sci.* **2015**, *16*, 5517–5527.
 214. Adcock, A.F.; Trivedi, G.; Edmondson, R.; Yang, C.S. and L. Three-Dimensional (3D) Cell Cultures in Cell-based Assays for in-vitro Evaluation of Anticancer Drugs. *J. Anal. Bioanal. Tech.* **2015**, *6*, 1.
 215. Durand, R.E.; Olive, P.L. Resistance of tumor cells to chemo- and radiotherapy modulated by the three-dimensional architecture of solid tumors and spheroids. *Methods Cell Biol.* **2001**, *64*, 211–233.
 216. Li, L.; Knutsdottir, H.; Hui, K.; Weiss, M.J.; He, J.; Philosophe, B.; Cameron, A.M.; Wolfgang, C.L.; Pawlik, T.M.; Ghiaur, G.; et al. Human primary liver cancer organoids reveal intratumor and interpatient drug response heterogeneity. *JCI Insight* **2019**, *4*.
 217. Sato, T.; Vries, R.G.; Snippert, H.J.; van de Wetering, M.; Barker, N.; Stange, D.E.; van Es, J.H.; Abo, A.; Kujala, P.; Peters, P.J.; et al. Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* **2009**, *459*, 262–265.
 218. Lancaster, M.A.; Knoblich, J.A. Organogenesis in a dish: Modeling

- development and disease using organoid technologies. *Science* (80-.). **2014**, 345.
219. Clevers, H. Modeling Development and Disease with Organoids. *Cell* **2016**, 165, 1586–1597.
 220. Jiang, F.X.; Harrison, L.C. Laminin-1 and epidermal growth factor family members co-stimulate fetal pancreas cell proliferation and colony formation. *Differentiation* **2005**, 73, 45–49.
 221. Bleijs, M.; Wetering, M.; Clevers, H.; Drost, J. Xenograft and organoid model systems in cancer research. *EMBO J.* **2019**, 38, e101654.
 222. Masters, J.R.W. Human cancer cell lines: Fact and fantasy. *Nat. Rev. Mol. Cell Biol.* **2000**, 1, 233–236.
 223. Qiu, W.; Su, G.H. Development of orthotopic pancreatic tumor mouse models. *Methods Mol. Biol.* **2013**, 215–223.
 224. Tiriach, H.; Bucobo, J.C.; Tzimas, D.; Grewel, S.; Lacombe, J.F.; Rowehl, L.M.; Nagula, S.; Wu, M.; Kim, J.; Sasson, A.; et al. Successful creation of pancreatic cancer organoids by means of EUS-guided fine-needle biopsy sampling for personalized cancer treatment. *Gastrointest. Endosc.* **2018**, 87, 1474–1480.
 225. Walsh, A.J.; Cook, R.S.; Sanders, M.E.; Arteaga, C.L.; Skala, M.C. Drug response in organoids generated from frozen primary tumor tissues. *Nat. Publ. Gr.* **2015**, 6, 1–11.
 226. Boj, S.F.; Hwang, C. Il; Baker, L.A.; Chio, I.I.C.; Engle, D.D.; Corbo, V.; Jager, M.; Ponz-Sarvisé, M.; Tiriach, H.; Spector, M.S.; et al. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **2015**, 160, 324–338.
 227. Olive, K.P.; Jacobetz, M.A.; Davidson, C.J.; Gopinathan, A.; McIntyre, D.; Honess, D.; Madhu, B.; Goldgraben, M.A.; Caldwell, M.E.; Allard, D.; et al. Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a mouse model of pancreatic cancer. *Science* **2009**, 324, 1457–61.
 228. Greggio, C.; De Franceschi, F.; Figueiredo-Larsen, M.; Gobaa, S.; Ranga, A.; Semb, H.; Lutolf, M.; Grapin-Botton, A. Artificial three-dimensional niches deconstruct pancreas development in vitro. *Development* **2013**, 140, 4452–62.
 229. Huch, M.; Bonfanti, P.; Boj, S.F.; Sato, T.; Loomans, C.J.M.; van de Wetering, M.; Sojoodi, M.; Li, V.S.W.; Schuijers, J.; Gracanin, A.; et al. Unlimited in vitro

- expansion of adult bi-potent pancreas progenitors through the Lgr5/R-spondin axis. *EMBO J.* **2013**, *32*, 2708–21.
230. Broutier, L.; Andersson-Rolf, A.; Hindley, C.J.; Boj, S.F.; Clevers, H.; Koo, B.K.; Huch, M. Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nat. Protoc.* **2016**, *11*, 1724.
 231. Georgakopoulos, N.; Prior, N.; Angres, B.; Mastrogiovanni, G.; Cagan, A.; Harrison, D.; Hindley, C.J.; Arnes-Benito, R.; Liao, S.S.; Curd, A.; et al. Long-term expansion, genomic stability and in vivo safety of adult human pancreas organoids. *BMC Dev. Biol.* **2020**, *20*, 1–20.
 232. Hou, S.; Tiriach, H.; Sridharan, B.P.; Scampavia, L.; Madoux, F.; Seldin, J.; Souza, G.R.; Watson, D.; Tuveson, D.; Spicer, T.P. Advanced Development of Primary Pancreatic Organoid Tumor Models for High-Throughput Phenotypic Drug Screening. *SLAS Discov. Adv. life Sci. R D* **2018**, *23*, 574–584.
 233. Huang, L.; Holtzinger, A.; Jagan, I.; Begora, M.; Lohse, I.; Ngai, N.; Nostro, C.; Wang, R.; Muthuswamy, L.B.; Crawford, H.C.; et al. Ductal pancreatic cancer modeling and drug screening using human pluripotent stem cell- and patient-derived tumor organoids. *Nat. Med.* **2015**, *21*, 1364–1371.
 234. Romero-Calvo, I.; Weber, C.R.; Ray, M.; Brown, M.; Kirby, K.; Nandi, R.K.; Long, T.M.; Sparrow, S.M.; Ugolkov, A.; Qiang, W.; et al. Human organoids share structural and genetic features with primary pancreatic adenocarcinoma tumors. *Mol. Cancer Res.* **2019**, *17*, 70–83.
 235. Frappart, P.O.; Walter, K.; Gout, J.; Beutel, A.K.; Morawe, M.; Arnold, F.; Breunig, M.; Barth, T.F.E.; Marienfeld, R.; Schulte, L.; et al. Pancreatic cancer-derived organoids – a disease modeling tool to predict drug response. *United Eur. Gastroenterol. J.* **2020**, *8*, 594–606.
 236. Tiriach, H.; Belleau, P.; Engle, D.D.; Plenker, D.; Deschênes, A.; Somerville, T.D.D.; Froeling, F.E.M.; Burkhart, R.A.; Denroche, R.E.; Jang, G.-H.; et al. Organoid Profiling Identifies Common Responders to Chemotherapy in Pancreatic Cancer. *Cancer Discov.* **2018**, *8*, 1112–1129.
 237. Von Bernstorff, W.; Voss, M.; Freichel, S.; Schmid, A.; Vogel, I.; Jöhnk, C.; Henne-Bruns, D.; Kremer, B.; Kalthoff, H. Systemic and local

- immunosuppression in pancreatic cancer patients. *Clin. Cancer Res.* **2001**, *7*, 925s-932s.
238. Tsai, S.; McOlash, L.; Palen, K.; Johnson, B.; Duris, C.; Yang, Q.; Dwinell, M.B.; Hunt, B.; Evans, D.B.; Gershan, J.; et al. Development of primary human pancreatic cancer organoids, matched stromal and immune cells and 3D tumor microenvironment models. *BMC Cancer* **2018**, *18*, 1–13.
 239. Menon, S.; Shin, S.; Dy, G. Advances in cancer immunotherapy in solid tumors. *Cancers (Basel)*. **2016**, *8*, 106.
 240. Öhlund, D.; Handly-Santana, A.; Biffi, G.; Elyada, E.; Almeida, A.S.; Ponz-Sarvise, M.; Corbo, V.; Oni, T.E.; Hearn, S.A.; Lee, E.J.; et al. Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J. Exp. Med.* **2017**, *214*, 579–596.
 241. Huang, L.; Bockorny, B.; Paul, I.; Akshinthala, D.; Gandarilla, O.; Bose, A.; Sanchez-Gonzalez, V.; Rouse, E.E.; Lehoux, S.D.; Pandell, N.; et al. Pancreatic tumor organoids for modeling in vivo drug response and discovering clinically-actionable biomarkers. *bioRxiv* **2019**, 513267.
 242. Hwang, W.Y.; Fu, Y.; Reyon, D.; Maeder, M.L.; Tsai, S.Q.; Sander, J.D.; Peterson, R.T.; Yeh, J.R.J.; Joung, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 227–229.
 243. Takeda, H.; Kataoka, S.; Nakayama, M.; Ali, M.A.E.; Oshima, H.; Yamamoto, D.; Park, J.-W.; Takegami, Y.; An, T.; Jenkins, N.A.; et al. CRISPR-Cas9-mediated gene knockout in intestinal tumor organoids provides functional validation for colorectal cancer driver genes. *Proc. Natl. Acad. Sci.* **2019**, *116*, 15635–15644.
 244. Niu, D.; Wei, H.J.; Lin, L.; George, H.; Wang, T.; Lee, I.H.; Zhao, H.Y.; Wang, Y.; Kan, Y.; Shrock, E.; et al. Inactivation of porcine endogenous retrovirus in pigs using CRISPR-Cas9. *Science (80-.)*. **2017**, *357*, 1303–1307.
 245. Yang, H.; Wang, H.; Jaenisch, R. Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. *Nat. Protoc.* **2014**, *9*, 1956.
 246. Driehuis, E.; Clevers, H. CRISPR/Cas 9 genome editing and its applications in organoids. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **2017**, *312*, G257–G265.
 247. Lee, J.; Snyder, E.R.; Liu, Y.; Gu, X.; Wang, J.; Flowers, B.M.; Kim, Y.J.; Park,

- S.; Szot, G.L.; Hruban, R.H.; et al. Reconstituting development of pancreatic intraepithelial neoplasia from primary human pancreas duct cells. *Nat. Commun.* **2017**, *8*, 14686.
248. Seino, T.; Kawasaki, S.; Shimokawa, M.; Tamagawa, H.; Toshimitsu, K.; Fujii, M.; Ohta, Y.; Matano, M.; Nanki, K.; Kawasaki, K.; et al. Human Pancreatic Tumor Organoids Reveal Loss of Stem Cell Niche Factor Dependence during Disease Progression. *Cell Stem Cell* **2018**, *22*, 454–467.
249. Aberle, M.R.; Burkhart, R.A.; Tiriach, H.; Olde Damink, S.W.M.; Dejong, C.H.C.; Tuveson, D.A.; van Dam, R.M. Patient-derived organoid models help define personalized management of gastrointestinal cancer. *Br. J. Surg.* **2018**, *105*, e48–e60.
250. Russell, W.M.S.; Burch, R.L. *The Principles of Humane Experimental Technique*; Methuen, 1959;
251. Beer, M.; Kuppalu, N.; Stefanini, M.; Becker, H.; Schulz, I.; Manoli, S.; Schuette, J.; Schmees, C.; Casazza, A.; Stelzle, M.; et al. A novel microfluidic 3D platform for culturing pancreatic ductal adenocarcinoma cells: Comparison with in vitro cultures and in vivo xenografts. *Sci. Rep.* **2017**, *7*, 1–12.
252. Young, E.W.K. Cells, tissues, and organs on chips: challenges and opportunities for the cancer tumor microenvironment. *Integr. Biol.* **2013**, *5*, 1096–1109.
253. Kersten, K.; de Visser, K.E.; van Miltenburg, M.H.; Jonkers, J. Genetically engineered mouse models in oncology research and cancer medicine. *EMBO Mol. Med.* **2017**, *9*, 137–153.
254. Hingorani, S.R.; Wang, L.; Multani, A.S.; Combs, C.; Deramaudt, T.B.; Hruban, R.H.; Rustgi, A.K.; Chang, S.; Tuveson, D.A. Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. *Cancer Cell* **2005**, *7*, 469–483.
255. Wiedenheft, B.; Sternberg, S.H.; Doudna, J.A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **2012**, *482*, 331–338.
256. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial

- immunity. *Science* (80-.). **2012**, 337, 816–821.
257. Pennisi, E. The CRISPR craze. *Science* (80-.). **2013**, 833–836.
 258. Ran, F.A.; Hsu, P.D.; Lin, C.Y.; Gootenberg, J.S.; Konermann, S.; Trevino, A.E.; Scott, D.A.; Inoue, A.; Matoba, S.; Zhang, Y.; et al. Double nicking by RNA-guided CRISPR cas9 for enhanced genome editing specificity. *Cell* **2013**, 154, 1380–1389.
 259. Horvath, P.; Barrangou, R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* (80-.). **2010**, 327, 167–170.
 260. Schouls, L.M.; Reulen, S.; Duim, B.; Wagenaar, J.A.; Willems, R.J.L.; Dingle, K.E.; Colles, F.M.; Van Embden, J.D.A. Comparative genotyping of *Campylobacter jejuni* by amplified fragment length polymorphism, multilocus sequence typing, and short repeat sequencing: Strain diversity, host range, and recombination. *J. Clin. Microbiol.* **2003**, 41, 15–26.
 261. Deveau, H.; Garneau, J.E.; Moineau, S. CRISPR/Cas system and its role in phage-bacteria interactions. *Annu. Rev. Microbiol.* **2010**, 64, 475–493.
 262. Lintner, N.G.; Frankel, K.A.; Tsutakawa, S.E.; Alsbury, D.L.; Copié, V.; Young, M.J.; Tainer, J.A.; Lawrence, C.M. The structure of the CRISPR-associated protein csa3 provides insight into the regulation of the CRISPR/Cas system. *J. Mol. Biol.* **2011**, 405, 939–955.
 263. Cribbs, A.P.; Perera, S.M.W. Science and bioethics of CRISPR-CAS9 gene editing: An analysis towards separating facts and fiction. *Yale J. Biol. Med.* **2017**, 90, 625–634.
 264. Raha, D.; Hong, M.; Snyder, M. ChIP-Seq: A method for global identification of regulatory elements in the genome. *Curr. Protoc. Mol. Biol.* **2010**, 91, 21–19.
 265. Hainer, S.J.; Fazzio, T.G. High-Resolution Chromatin Profiling Using CUT&RUN. *Curr. Protoc. Mol. Biol.* **2019**, 126, e85.
 266. Baranello, L.; Kouzine, F.; Sanford, S.; Levens, D. ChIP bias as a function of cross-linking time. *Chromosom. Res.* **2016**, 24, 175–181.
 267. Van Steensel, B.; Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered Dam methyltransferase. *Nat. Biotechnol.* **2000**, 18, 424–428.
 268. Schmid, M.; Durussel, T.; Laemmli, U.K. ChIC and ChEC: Genomic mapping

- of chromatin proteins. *Mol. Cell* **2004**, *16*, 147–157.
269. Skene, P.J.; Henikoff, J.G.; Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **2018**, *13*, 1006.
270. Policastro, R.A.; Zentner, G.E. Enzymatic methods for genome-wide profiling of protein binding sites. *Brief. Funct. Genomics* **2018**, *17*, 138–145.
271. Addgene CUT&RUN: An Improved Method for Studying Protein-DNA Interactions Available online: <https://blog.addgene.org/cutrun-a-improved-method-for-studying-protein-dna-interactions>.
272. Hannah, R.R.; Wood, K. V Dual-Luciferase™ Reporter Assay: An Advanced Co-Reporter Technology Integrating Firefly and Renilla Luciferase Assays. *Promega Notes Mag. Number* **1996**, *57*, 3–10.
273. De Wet, J.R.; Wood, K. V.; Helinski, D.R.; DeLuca, M. Cloning of firefly luciferase cDNA and the expression of active luciferase in *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82*, 7870–7873.
274. Wood, K. V.; de Wet, J.R.; Dewji, N.; DeLuca, M. Synthesis of active firefly luciferase by in vitro translation of RNA obtained from adult lanterns. *Biochem. Biophys. Res. Commun.* **1984**, *124*, 592–596.
275. Matthews, J.C.; Hori, K.; Cormier, M.J. Purification and Properties of Renilla reniformis Luciferase. *Biochemistry* **1977**, *16*, 85–91.
276. Loughran, G.; Zhdanov, A. V.; Mikhaylova, M.S.; Rozov, F.N.; Datskevich, P.N.; Kovalchuk, S.I.; Serebryakova, M. V.; Kiniry, S.J.; Michel, A.M.; O'Connor, P.B.F.; *et al.* Unusually efficient CUG initiation of an overlapping reading frame in POLG mRNA yields novel protein POLGARF. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*.
277. Gaspar, P.; Lopes, P.; Oliveira, J.; Santos, R.; Dagleish, R.; Oliveira, J.L. Variobox: Automatic detection and annotation of human genetic variants. *Hum. Mutat.* **2014**, *35*, 202–207.
278. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21.
279. Ge, S.X.; Son, E.W.; Yao, R. iDEP: An integrated web application for differential expression and pathway analysis of RNA-Seq data. *BMC Bioinformatics* **2018**, *19*, 1–24.

-
280. Langmead, B.; Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **2012**, *9*, 357.
281. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nussbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, 1–9.
282. Heinz, S.; Benner, C.; Spann, N.; Bertolino, E.; Lin, Y.C.; Laslo, P.; Cheng, J.X.; Murre, C.; Singh, H.; Glass, C.K. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **2010**, *38*, 576–589.
283. Robinson, J.T.; Thorvaldsdóttir, H.; Winckler, W.; Guttman, M.; Lander, E.S.; Getz, G.; Mesirov, J.P. Integrative genomics viewer. *Nat. Biotechnol.* **2011**, *29*, 24–26.
284. Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; Kaul, R.; et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57.
285. Kulakovskiy, I. V.; Vorontsov, I.E.; Yevshin, I.S.; Sharipov, R.N.; Fedorova, A.D.; Rumynskiy, E.I.; Medvedeva, Y.A.; Magana-Mora, A.; Bajic, V.B.; Papatsenko, D.A.; et al. HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **2018**, *46*, D252–D259.
286. Vorontsov, I.E.; Kulakovskiy, I. V.; Khimulya, G.; Nikolaeva, D.D.; Makeev, V.J. PERFECTOS-APE: Predicting regulatory functional effect of SNPs by approximate P-value estimation. *Proc. Int. Jt. Conf. Biomed. Eng. Syst. Technol.* **2015**, *3*.
287. Gillet, J.-P.; Varma, S.; Gottesman, M.M. The clinical relevance of cancer cell lines. *J. Natl. Cancer Inst.* **2013**, *105*, 452–458.
288. Miyoshi, H.; Stappenbeck, T.S. In vitro expansion and genetic modification of gastrointestinal stem cells in spheroid culture. *Nat. Protoc.* **2013**, *8*, 2471–2482.
289. Veeman, M.T.; Slusarski, D.C.; Kaykas, A.; Louie, S.H.; Moon, R.T. Zebrafish prickles, a modulator of noncanonical Wnt/Fz signaling, regulates gastrulation movements. *Curr. Biol.* **2003**, *13*, 680–685.
290. Ono, K.; Satoh, M.; Yoshida, T.; Ozawa, Y.; Kohara, A.; Takeuchi, M.;

- Mizusawa, H.; Sawada, H. Species identification of animal cells by nested PCR targeted to mitochondrial DNA. *Vitr. Cell. Dev. Biol. - Anim.* **2007**, *43*, 168–175.
291. Heurtier, V.; Owens, N.; Gonzalez, I.; Mueller, F.; Proux, C.; Mornico, D.; Clerc, P.; Dubois, A.; Navarro, P. The molecular logic of Nanog-induced self-renewal in mouse embryonic stem cells. *Nat. Commun.* **2019**, *10*, 1–15.
292. Shi, G.; Jin, Y. Role of Oct4 in maintaining and regaining stem cell pluripotency. *Stem Cell Res. Ther.* **2010**, *1*, 1–9.
293. Schaefer, T.; Lengerke, C. SOX2 protein biochemistry in stemness, reprogramming, and cancer: the PI3K/AKT/SOX2 axis and beyond. *Oncogene* **2020**, *39*, 278–292.
294. Roche, S.; O'Neill, F.; Murphy, J.; Swan, N.; Meiller, J.; Conlon, N.T.; Geoghegan, J.; Conlon, K.; McDermott, R.; Rahman, R.; *et al.* Establishment and characterisation by expression microarray of patient derived xenograft panel of human pancreatic adenocarcinoma patients. *Int. J. Mol. Sci.* **2020**, *21*, 962.
295. European Council; European Parliament Caring for animals aiming for better science. *Off. J. Eur. Union* **2010**.
296. Walsh, A.J.; Castellanos, J.A.; Nagathihalli, N.S.; Merchant, N.B.; Skala, M.C. Optical Imaging of Drug-Induced Metabolism Changes in Murine and Human Pancreatic Cancer Organoids Reveals Heterogeneous Drug Response. *Pancreas* **2016**, *45*, 863.
297. Vidal, V.P.I.; Motamedi, F.J.; Rekima, S.; Gregoire, E.P.; Szenker-Ravi, E.; Leushacke, M.; Reversade, B.; Chaboissier, M.C.; Schedl, A. R-spondin signalling is essential for the maintenance and differentiation of mouse nephron progenitors. *Elife* **2020**, *9*, e53895.
298. Sato, T.; Stange, D.E.; Ferrante, M.; Vries, R.G.J.; Van Es, J.H.; Van Den Brink, S.; Van Houdt, W.J.; Pronk, A.; Van Gorp, J.; Siersema, P.D.; *et al.* Long term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology* **2011**, *141*, 1762–1772.
299. Hilkens, J.; Timmer, N.C.; Boer, M.; Ikink, G.J.; Schewe, M.; Sacchetti, A.; Koppens, M.A.J.; Song, J.Y.; Bakker, E.R.M. RSPO3 expands intestinal stem

- cell and niche compartments and drives tumorigenesis. *Gut* **2017**, *66*, 1095–1105.
300. Choi, S. Il; Jeon, A.R.; Kim, M.K.; Lee, Y.S.; Im, J.E.; Koh, J.W.; Han, S.S.; Kong, S.Y.; Yoon, K.A.; Koh, Y.H.; *et al.* Development of Patient Derived Preclinical Platform for Metastatic Pancreatic Cancer: PDOX and a Subsequent Organoid Model System Using Percutaneous Biopsy Samples. *Front. Oncol.* **2019**, *9*, 875.
 301. Fong, E.L.S.; Toh, T.B.; Lin, X.; Liu, Z.; Hooi, L.; Mohd Abdul Rashid, M.B.; Benoukraf, T.; Chow, E.K.H.; Huynh, T.H.; Yu, H. Generation of matched patient derived xenograft in vitro-in vivo models using 3D macroporous hydrogels for the study of liver cancer. *Biomaterials* **2018**, *159*, 229–240.
 302. Beshiri, M.L.; Tice, C.M.; Tran, C.; Nguyen, H.M.; Sowalsky, A.G.; Agarwal, S.; Jansson, K.H.; Yang, Q.; McGowen, K.M.; Yin, J.J.; *et al.* A PDX/Organoid biobank of advanced prostate cancers captures genomic and phenotypic heterogeneity for disease modeling and therapeutic screening. *Clin. Cancer Res.* **2018**, *24*, 4332–4345.
 303. Guillen, K.P.; Fujita, M.; Butterfield, A.J.; Scherer, S.D.; Bailey, M.H.; Chu, Z.; Derosé, Y.S.; Zhao, L.; Cortes-Sanchez, E.; Yang, C.-H.; *et al.* A breast cancer patient derived xenograft and organoid platform for drug discovery and precision oncology. *bioRxiv* **2021**.
 304. Finlay-Schultz, J.; Jacobsen, B.M.; Riley, D.; Paul, K. V.; Turner, S.; Ferreira-Gonzalez, A.; Harrell, J.C.; Kabos, P.; Sartorius, C.A. New generation breast cancer cell lines developed from patient derived xenografts. *Breast Cancer Res.* **2020**, *22*, 1–12.
 305. Borodovsky, A.; McQuiston, T.J.; Stetson, D.; Ahmed, A.; Whitston, D.; Zhang, J.; Grondine, M.; Lawson, D.; Challberg, S.S.; Zinda, M.; *et al.* Generation of stable PDX derived cell lines using conditional reprogramming. *Mol. Cancer* **2017**, *16*, 1–7.
 306. Nelson, S.R.; Zhang, C.; Roche, S.; O'Neill, F.; Swan, N.; Luo, Y.; Larkin, A.M.; Crown, J.; Walsh, N. Modelling of pancreatic cancer biology: transcriptomic signature for 3D PDX derived organoids and primary cell line organoid development. *Sci. Rep.* **2020**, *10*, 1–12.
 307. Hidalgo, M.; Amant, F.; Biankin, A. V.; Budinská, E.; Byrne, A.T.; Caldas, C.;

- Clarke, R.B.; de Jong, S.; Jonkers, J.; Mælandsmo, G.M.; *et al.* Patient derived Xenograft models: An emerging platform for translational cancer research. *Cancer Discov.* **2014**, *4*, 998–1013.
308. D'Agosto, S.; Andreani, S.; Scarpa, A.; Corbo, V. Preclinical modelling of PDA: Is organoid the new black? *Int. J. Mol. Sci.* **2019**, *20*, 2766.
309. Bartfeld, S.; Clevers, H. Organoids as model for infectious diseases: Culture of human and murine stomach organoids and microinjection of helicobacter pylori. *J. Vis. Exp.* **2015**, 105.
310. Xiang, Y.; Cakir, B.; Park, I.-H. Generation of Regionally Specified Human Brain Organoids Resembling Thalamus Development. *STAR Protoc.* **2020**, *1*, 100001.
311. Kim, M.P.; Fleming, J.B.; Wang, H.; Abbruzzese, J.L.; Choi, W.; Kopetz, S.; McConkey, D.J.; Evans, D.B.; Gallick, G.E. ALDH Activity Selectively Defines an Enhanced Tumor-Initiating Cell Population Relative to CD133 Expression in Human Pancreatic Adenocarcinoma. *PLoS One* **2011**, *6*, e20636.
312. Tomita, H.; Tanaka, K.; Tanaka, T.; Hara, A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget* **2016**, *7*, 11018.
313. Sun, J.S.; Zhang, X.L.; Yang, Y.J.; Nie, Z.G.; Zhang, Y. Hypoxia promotes C-X-C chemokine receptor type 4 expression through microRNA-150 in pancreatic cancer cells. *Oncol. Lett.* **2015**, *10*, 835–840.
314. Erkan, M.; Kurtoglu, M.; Kleeff, J. The role of hypoxia in pancreatic cancer: A potential therapeutic target? *Expert Rev. Gastroenterol. Hepatol.* **2016**, *10*, 301–316.
315. Hammond, E.M.; Asselin, M.C.; Forster, D.; O'Connor, J.P.B.; Senra, J.M.; Williams, K.J. The Meaning, Measurement and Modification of Hypoxia in the Laboratory and the Clinic. *Clin. Oncol.* **2014**, *26*, 277–288.
316. Skoda, J.; Hermanova, M.; Loja, T.; Nemec, P.; Neradil, J.; Karasek, P.; Veselska, R. Co-expression of cancer stem cell markers corresponds to a pro-tumorigenic expression profile in pancreatic adenocarcinoma. *PLoS One* **2016**, *11*, e0159255.
317. Meng, Y.; Xu, B.Q.; Fu, Z.G.; Wu, B.; Xu, B.; Chen, Z.N.; Li, L. Cytoplasmic EpCAM over-expression is associated with favorable clinical outcomes in pancreatic cancer patients with hepatitis B virus negative infection. *Int. J.*

- Clin. Exp. Med.* **2015**, *8*, 22204.
318. Dassaye, R.; Naidoo, S.; Cerf, M.E. Transcription factor regulation of pancreatic organogenesis, differentiation and maturation. *Islets* **2016**, *8*, 13–34.
319. Roy, N.; Takeuchi, K.K.; Ruggeri, J.M.; Bailey, P.; Chang, D.; Li, J.; Leonhardt, L.; Puri, S.; Hoffman, M.T.; Gao, S.; *et al.* PDX1 dynamically regulates pancreatic ductal adenocarcinoma initiation and maintenance. *Genes Dev.* **2016**, *30*, 2669–2683.
320. Hohwieler, M.; Müller, M.; Frappart, P.O.; Heller, S. Pancreatic progenitors and organoids as a prerequisite to model pancreatic diseases and cancer. *Stem Cells Int.* **2019**.
321. Kapałczyńska, M.; Kolenda, T.; Przybyła, W.; Zajączkowska, M.; Teresiak, A.; Filas, V.; Ibbs, M.; Bliźniak, R.; Łuczewski, Ł.; Lamperska, K. 2D and 3D cell cultures – a comparison of different types of cancer cell cultures. *Arch. Med. Sci.* **2018**, *14*, 910.
322. Wen, Z.; Liao, Q.; Hu, Y.; You, L.; Zhou, L.; Zhao, Y. A spheroid-based 3-D culture model for pancreatic cancer drug testing, using the acid phosphatase assay. *Brazilian J. Med. Biol. Res.* **2013**, *46*, 634–642.
323. Miyamoto, H.; Murakami, T.; Tsuchida, K.; Sugino, H.; Miyake, H.; Tashiro, S. Tumor-Stroma Interaction of Human Pancreatic Cancer: Acquired Resistance to Anticancer Drugs and Proliferation Regulation Is Dependent on Extracellular Matrix Proteins. *Pancreas* **2004**, *28*, 38–44.
324. Williams, T.M.; Medina, F.; Badano, I.; Hazan, R.B.; Hutchinson, J.; Muller, W.J.; Chopra, N.G.; Scherer, P.E.; Pestell, R.G.; Lisanti, M.P. Caveolin-1 gene disruption promotes mammary tumorigenesis and dramatically enhances lung metastasis in vivo: Role of Cav-1 in cell invasiveness and matrix metalloproteinase (MMP-2/9) secretion. *J. Biol. Chem.* **2004**, *479*, 51630–51646.
325. Cervantes-Madrid, D.L.; Nagi, S.; Gustafsson, A.A. FosB transcription factor regulates COX-2 expression in colorectal cancer cells without affecting PGE2 expression. *Oncol. Lett.* **2017**, *13*, 1411–1416.
326. Lu, D.; Wang, J.; Shi, X.; Yue, B.; Hao, J. AHNK2 is a potential prognostic biomarker in patients with PDAC. *Oncotarget* **2017**, *8*, 31775.

327. Mrug, M.; Zhou, J.; Yang, C.; Aronow, B.J.; Cui, X.; Schoeb, T.R.; Siegal, G.P.; Yoder, B.K.; Guay-Woodford, L.M. Genetic and informatic analyses implicate Kif12 as a candidate gene within the Mpkd2 locus that modulates renal cystic disease severity in the Cyslcpk mouse. *PLoS One* **2015**, *10*, e0135678.
328. Cano, D.A.; Sekine, S.; Hebrok, M. Primary Cilia Deletion in Pancreatic Epithelial Cells Results in Cyst Formation and Pancreatitis. *Gastroenterology* **2006**, *131*, 1856–1869.
329. Li, P.; He, Y.; Cai, G.; Xiao, F.; Yang, J.; Li, Q.; Chen, X. CCDC114 is mutated in patient with a complex phenotype combining primary ciliary dyskinesia, sensorineural deafness, and renal disease. *J. Hum. Genet.* **2019**, *64*, 39–48.
330. Goetz, S.C.; Anderson, K. V. The primary cilium: A signalling centre during vertebrate development. *Nat. Rev. Genet.* **2010**, *11*, 331–344.
331. diIorio, P.; Rittenhouse, A.R.; Bortell, R.; Jurczyk, A. Role of cilia in normal pancreas function and in diseased states. *Birth Defects Res. Part C - Embryo Today Rev.* **2014**, *102*, 126–138.
332. Powell, A.E.; Wang, Y.; Li, Y.; Poulin, E.J.; Means, A.L.; Washington, M.K.; Higginbotham, J.N.; Juchheim, A.; Prasad, N.; Levy, S.E.; *et al.* The pan-ErbB negative regulator Irlg1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell* **2012**, *149*, 146–158.
333. Ioannou, M.; Serafimidis, I.; Arnes, L.; Sussel, L.; Singh, S.; Vasiliou, V.; Gavalas, A. ALDH1B1 is a potential stem/progenitor marker for multiple pancreas progenitor pools. *Dev. Biol.* **2013**, *374*, 153–163.
334. Chatterjee, M.; Ben-Josef, E.; Thomas, D.G.; Morgan, M.A.; Zalupski, M.M.; Khan, G.; Andrew Robinson, C.; Griffith, K.A.; Chen, C.S.; Ludwig, T.; *et al.* Caveolin-1 is associated with tumor progression and confers a multi-modality resistance phenotype in pancreatic cancer. *Sci. Rep.* **2015**, *5*, 1–15.
335. Tsai, M.S.; Bogart, D.F.; Castañeda, J.M.; Li, P.; Lupu, R. Cyr61 promotes breast tumorigenesis and cancer progression. *Oncogene* **2002**, *21*, 8178–8185.
336. Harris, N.L.E.; Vennin, C.; Conway, J.R.W.; Vine, K.L.; Pinese, M.; Cowley, M.J.; Shearer, R.F.; Lucas, M.C.; Herrmann, D.; Allam, A.H.; *et al.* SerpinB2 regulates stromal remodelling and local invasion in pancreatic cancer.

- Oncogene* **2017**, *36*, 4288–4298.
337. Wang, L.; Yu, J.; Ni, J.; Xu, X.M.; Wang, J.; Ning, H.; Pei, X.F.; Chen, J.; Yang, S.; Underhill, C.B.; *et al.* Extracellular matrix protein 1 (ECM1) is over-expressed in malignant epithelial tumors. *Cancer Lett.* **2003**, *200*, 57–67.
 338. Saito, T.; Kasamatsu, A.; Ogawara, K.; Miyamoto, I.; Saito, K.; Iyoda, M.; Suzuki, T.; Endo-Sakamoto, Y.; Shiiba, M.; Tanzawa, H.; *et al.* Semaphorin7A promotion of tumoral growth and metastasis in human oral cancer by regulation of g1 cell cycle and matrix metalloproteases: Possible contribution to tumoral angiogenesis. *PLoS One* **2015**, *10*, e0137923.
 339. Bressy, C.; Lac, S.; Nigri, J.E.; Leca, J.; Roques, J.; Lavaut, M.N.; Secq, V.; Guillaumond, F.; Bui, T.T.; Pietrasz, D.; *et al.* LIF drives neural remodeling in pancreatic cancer and offers a new candidate biomarker. *Cancer Res.* **2018**, *78*, 909–921.
 340. Roche, J.; Boldog, F.; Robinson, M.; Robinson, L.; Varella-Garcia, M.; Swanton, M.; Waggoner, B.; Fishel, R.; Franklin, W.; Gemmill, R.; *et al.* Distinct 3p21.3 deletions in lung cancer and identification of a new human semaphorin. *Oncogene* **1996**, *12*, 1289–1297.
 341. Kasashima, H.; Duran, A.; Cid-Diaz, T.; Kudo, Y.; Diaz-Meco, M.T.; Moscat, J. An Orthotopic Implantation Mouse Model of Hepatocellular Carcinoma with Underlying Liver Steatosis. *STAR Protoc.* **2020**, *1*, 100185.
 342. Benton, G.; Arnaoutova, I.; George, J.; Kleinman, H.K.; Koblinski, J. Matrigel: From discovery and ECM mimicry to assays and models for cancer research. *Adv. Drug Deliv. Rev.* **2014**, *79*, 3–18.
 343. Hayashi, Y.; Emoto, T.; Futaki, S.; Sekiguchi, K. Establishment and characterization of a parietal endoderm-like cell line derived from Engelbreth-Holm-Swarm tumor (EHSPEL), a possible resource for an engineered basement membrane matrix. *Matrix Biol.* **2004**, *23*, 47–62.
 344. Eirew, P.; Steif, A.; Khattra, J.; Ha, G.; Yap, D.; Farahani, H.; Gelmon, K.; Chia, S.; Mar, C.; Wan, A.; *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **2015**, *518*, 422–426.
 345. VanLiere, J.M.; Rosenberg, N.A. Mathematical properties of the r^2 measure of linkage disequilibrium. *Theor. Popul. Biol.* **2008**, *74*, 130–137.
 346. Martinez, E.; Silvy, F.; Fina, F.; Bartoli, M.; Krahn, M.; Barlesi, F.; Figarella-

- Branger, D.; Iovanna, J.; Laugier, R.; Ouaisi, M.; *et al.* Rs488087 single nucleotide polymorphism as predictive risk factor for pancreatic cancers. *Oncotarget* **2015**, *6*, 39855.
347. Ardlie, K.G.; DeLuca, D.S.; Segrè, A. V.; Sullivan, T.J.; Young, T.R.; Gelfand, E.T.; Trowbridge, C.A.; Maller, J.B.; Tukiainen, T.; Lek, M.; *et al.* The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* (80-.). **2015**, *348*, 648–660.
348. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **2019**, *47*, W556–W560.
349. Tang, H.; Wei, P.; Chang, P.; Li, Y.; Yan, D.; Liu, C.; Hassan, M.; Li, D. Genetic polymorphisms associated with pancreatic cancer survival: a genome-wide association study. *Int. J. Cancer* **2017**, *141*, 678–686.
350. Rizzato, C.; Campa, D.; Talar-Wojnarowska, R.; Halloran, C.; Kupcinkas, J.; Butturini, G.; Mohelníková-Duchoňová, B.; Sperti, C.; Tjaden, C.; Ghaneh, P.; *et al.* Association of genetic polymorphisms with survival of pancreatic ductal adenocarcinoma patients. *Carcinogenesis* **2016**, *37*, 957–964.
351. Baker, L.; Muir, P.; Sample, S.J. Genome-wide association studies and genetic testing: Understanding the science, success, and future of a rapidly developing field. *J. Am. Vet. Med. Assoc.* **2019**, *255*, 1126–1136.
352. Vavouri, T.; Lehner, B. Human genes with CpG island promoters have a distinct transcription-associated chromatin organization. *Genome Biol.* **2012**, *13*, 1–12.
353. Gates, L.A.; Foulds, C.E.; O'Malley, B.W. Histone Marks in the 'Driver's Seat': Functional Roles in Steering the Transcription Cycle. *Trends Biochem. Sci.* **2017**, *42*, 977–989.
354. Nilsson, J.; Bläckberg, L.; Carlsson, P.; Enerbäck, S.; Hernell, O.; Bjursell, G. cDNA cloning of human-milk bile-salt-stimulated lipase and evidence for its identity to pancreatic carboxylic ester hydrolase. *Eur. J. Biochem.* **1990**, *192*, 543–550.
355. Lidberg, U.; Nilsson, J.; Strömberg, K.; Stenman, G.; Sahlin, P.; Enerbäck, S.; Bjursell, G. Genomic organization, sequence analysis, and chromosomal localization of the human carboxyl ester lipase (CEL) gene and a CEL-like

- (CELL) gene. *Genomics* **1992**, *13*, 630–640.
356. Fjeld, K.; Weiss, F.U.; Lasher, D.; Rosendahl, J.; Chen, J.M.; Johansson, B.B.; Kirsten, H.; Ruffert, C.; Masson, E.; Steine, S.J.; *et al.* A recombined allele of the lipase gene CEL and its pseudogene CELP confers susceptibility to chronic pancreatitis. *Nat. Genet.* **2015**, *47*, 518–522.
 357. Ræder, H.; Johansson, S.; Holm, P.I.; Haldorsen, I.S.; Mas, E.; Sbarra, V.; Nermoen, I.; Eide, S.Å.; Grevle, L.; Bjørkhaug, L.; *et al.* Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. *Nat. Genet.* **2006**, *38*, 54–62.
 358. Raimondi, S.; Lowenfels, A.B.; Morselli-Labate, A.M.; Maisonneuve, P.; Pezzilli, R. Pancreatic cancer in chronic pancreatitis; Aetiology, incidence, and early detection. *Best Pract. Res. Clin. Gastroenterol.* **2010**, *24*, 349–358.
 359. El-Brolosy, M.A.; Stainier, D.Y.R. Genetic compensation: A phenomenon in search of mechanisms. *PLoS Genet.* **2017**, *13*, e1006780.
 360. Wang, Z.N.; Bassett, M.; Rainey, W.E. Liver receptor homologue-1 is expressed in the adrenal and can regulate transcription of 11 β -hydroxylase. *J. Mol. Endocrinol.* **2001**, *27*, 255–258.
 361. Martin, L.J.; Taniguchi, H.; Robert, N.M.; Simard, J.; Tremblay, J.J.; Viger, R.S. GATA factors and the nuclear receptors, steroidogenic factor 1/liver receptor homolog 1, are key mutual partners in the regulation of the human 3 β -hydroxysteroid dehydrogenase type 2 promoter. *Mol. Endocrinol.* **2005**, *19*, 2358–2370.
 362. Müller, S.; Raulefs, S.; Bruns, P.; Afonso-Grunz, F.; Plötner, A.; Thermann, R.; Jäger, C.; Schlitter, A.M.; Kong, B.; Regel, I.; *et al.* Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. *Mol. Cancer* **2015**, *14*, 1–18.
 363. Yang, X.; Song, J.H.; Cheng, Y.; Wu, W.; Bhagat, T.; Yu, Y.; Abraham, J.M.; Ibrahim, S.; Ravich, W.; Roland, B.C.; *et al.* Long non-coding RNA HNF1A-AS1 regulates proliferation and migration in oesophageal adenocarcinoma cells. *Gut* **2014**, *63*, 881–890.
 364. Liu, Z.; Wei, X.; Zhang, A.; Li, C.; Bai, J.; Dong, J. Long non-coding RNA HNF1A-AS1 functioned as an oncogene and autophagy promoter in hepatocellular carcinoma through sponging hsa-miR-30b-5p. *Biochem.*

- Biophys. Res. Commun.* **2016**, *473*, 1268–1275.
365. Pressler, H.; Sissung, T.M.; Venzon, D.; Price, D.K.; Figg, W.D. Expression of OATP family members in hormone-related cancers: Potential markers of progression. *PLoS One* **2011**, *6*, e20372.
366. Kloesch, B.; Ionasz, V.; Paliwal, S.; Hruschka, N.; Martinez De Villarreal, J.; Öllinger, R.; Mueller, S.; Dienes, H.P.; Schindl, M.; Gruber, E.S.; *et al.* A GATA6-centred gene regulatory network involving HNFs and Δ Np63 controls plasticity and immune escape in pancreatic cancer. *Gut* **2021**.
367. Yao, W.; Maitra, A.; Ying, H. Recent insights into the biology of pancreatic cancer. *EBioMedicine* **2020**, *53*, 102655.
368. Anderson, K.G.; Stromnes, I.M.; Greenberg, P.D. Obstacles Posed by the Tumor Microenvironment to T cell Activity: A Case for Synergistic Therapies. *Cancer Cell* **2017**, *31*, 311–325.
369. Kidder, B.L.; Hu, G.; Zhao, K. ChIP-Seq: Technical considerations for obtaining high-quality data. *Nat. Immunol.* **2011**, *12*, 918–922.
370. Sancar, A.; Lindsey-Boltz, L.A.; Ünsal-Kaçmaz, K.; Linn, S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annu. Rev. Biochem.* **2004**, *73*, 39–85.
371. Ko, H.L.; Ren, E.C. Functional aspects of PARP1 in DNA repair and transcription. *Biomolecules* **2012**, *2*, 524–548.
372. Bagnolini, G.; Milano, D.; Manerba, M.; Schipani, F.; Ortega, J.A.; Gioia, D.; Falchi, F.; Balboni, A.; Farabegoli, F.; De Franco, F.; *et al.* Synthetic Lethality in Pancreatic Cancer: Discovery of a New RAD51-BRCA2 Small Molecule Disruptor That Inhibits Homologous Recombination and Synergizes with Olaparib. *J. Med. Chem.* **2020**, *63*, 2588–2619.
373. Jamieson, E.R.; Lippard, S.J. Structure, recognition, and processing of cisplatin-DNA adducts. *Chem. Rev.* **1999**, *99*, 2467–2498.
374. Chen, Y.; Zhang, L.; Hao, Q. Olaparib: A promising PARP inhibitor in ovarian cancer therapy. *Arch. Gynecol. Obstet.* **2013**, *288*, 367–374.
375. Liu, X.; Qian, D.; Liu, H.; Abbruzzese, J.L.; Luo, S.; Walsh, K.M.; Wei, Q. Genetic variants of the peroxisome proliferator-activated receptor (PPAR) signaling pathway genes and risk of pancreatic cancer. *Mol. Carcinog.* **2020**, *59*, 930–939.

376. Feng, Y.; Liu, H.; Duan, B.; Liu, Z.; Abbruzzese, J.; Walsh, K.M.; Zhang, X.; Wei, Q. Potential functional variants in SMC2 and TP53 in the AURORA pathway genes and risk of pancreatic cancer. *Carcinogenesis* **2019**, *40*, 521–528.
377. Wang, M.Y.; Zhu, M.L.; He, J.; Shi, T.Y.; Li, Q.X.; Wang, Y.N.; Li, J.; Zhou, X.Y.; Sun, M.H.; Wang, X.F.; *et al.* Potentially Functional Polymorphisms in the CASP7 Gene Contribute to Gastric Adenocarcinoma Susceptibility in an Eastern Chinese Population. *PLoS One* **2013**, *8*, e74041.
378. Zheng, Z.; Liu, S.; Wang, C.; Wang, C.; Tang, D.; Shi, Y.; Han, X. Association of genetic polymorphisms in CASP7 with risk of ischaemic stroke. *Sci. Rep.* **2019**, *9*, 1–8.
379. García-Lozano, J.R.; Torres, B.; Fernández, O.; Orozco, G.; Álvarez-Márquez, A.; García, A.; González-Gay, M.A.; García, A.; Núñez-Roldán, A.; Martín, J.; *et al.* Caspase 7 influences susceptibility to rheumatoid arthritis. *Rheumatology* **2007**, *46*, 1243–1247.
380. Babu, S.R.; Bao, F.; Roberts, C.M.; Martin, A.K.; Gowan, K.; Eisenbarth, G.S.; Fain, P.R. Caspase 7 is a positional candidate gene for IDDM 17 in a bedouin arab family. *Ann. N. Y. Acad. Sci.* **2003**, *1005*, 340–343.
381. Paull, T.T.; Cortez, D.; Bowers, B.; Elledge, S.J.; Gellert, M. Direct DNA-binding by Brcal. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6086–6091.
382. Yu, R.; Li, C.; Lin, X.; Chen, Q.; Li, J.; Song, L.; Lin, L.; Liu, J.; Zhang, Y.; Kong, W.; *et al.* Clinicopathologic features and prognostic implications of MYBL2 protein expression in pancreatic ductal adenocarcinoma. *Pathol. Res. Pract.* **2017**, *213*, 964–968.
383. Xiong, Y.C.; Wang, J.; Cheng, Y.; Zhang, X.Y.; Ye, X.Q. Overexpression of MYBL2 promotes proliferation and migration of non-small-cell lung cancer via upregulating NCAPH. *Mol. Cell. Biochem.* **2020**, 1–9.
384. Soung, Y.H.; Lee, J.W.; Kim, H.S.; Park, W.S.; Kim, S.Y.; Lee, J.H.; Park, J.Y.; Cho, Y.G.; Kim, C.J.; Park, Y.G.; *et al.* Inactivating mutations of CASPASE-7 gene in human cancers. *Oncogene* **2003**, *22*, 8048–8052.
385. Palmerini, F.; Devilard, E.; Jarry, A.; Birg, F.; Xerri, L. Caspase 7 downregulation as an immunohistochemical marker of colonic carcinoma. *Hum. Pathol.* **2001**, *32*, 461–467.

386. Yoo, N.J.; Lee, J.W.; Kim, Y.J.; Soung, Y.H.; Kim, S.Y.; Nam, S.W.; Park, W.S.; Lee, J.Y.; Lee, S.H. Loss of caspase-2, -6 and -7 expression in gastric cancers. *APMIS* **2004**, *112*, 330–335.
387. Chaudhary, S.; Madhukrishna, B.; Adhya, A.K.; Keshari, S.; Mishra, S.K. Overexpression of caspase 7 is ER α dependent to affect proliferation and cell growth in breast cancer cells by targeting p21Cip. *Oncogenesis* **2016**, *5*, e219–e219.
388. Lindner, A.U.; Lucantoni, F.; Varešlija, D.; Resler, A.; Murphy, B.M.; Gallagher, W.M.; Hill, A.D.K.; Young, L.S.; Prehn, J.H.M. Low cleaved caspase-7 levels indicate unfavourable outcome across all breast cancers. *J. Mol. Med.* **2018**, *96*, 1025–1037.
389. Murakami-Tonami, Y.; Kishida, S.; Takeuchi, I.; Katou, Y.; Maris, J.M.; Ichikawa, H.; Kondo, Y.; Sekido, Y.; Shirahige, K.; Murakami, H.; *et al.* Inactivation of SMC2 shows a synergistic lethal response in MYCN-amplified neuroblastoma cells. *Cell Cycle* **2014**, *13*, 1115–1131.
390. Dávalos, V.; Suárez-López, L.; Castaño, J.; Messent, A.; Abasolo, I.; Fernandez, Y.; Guerra-Moreno, A.; Espín, E.; Armengol, M.; Musulen, E.; *et al.* Human SMC2 protein, a core subunit of human condensin complex, is a novel transcriptional target of the Wnt signaling pathway and a new therapeutic target. *J. Biol. Chem.* **2012**, *287*, 43472–43481.
391. Zhang, T.; Si-Hoe, S.L.; Hudson, D.F.; Surana, U. Condensin recruitment to chromatin is inhibited by Chk2 kinase in response to DNA damage. *Cell Cycle* **2016**, *15*, 3454–3470.
392. Perkhofer, L.; Gout, J.; Roger, E.; Kude De Almeida, F.; Baptista Simões, C.; Wiesmüller, L.; Seufferlein, T.; Kleger, A. DNA damage repair as a target in pancreatic cancer: State-of-the-art and future perspectives. *Gut* **2021**, *70*, 606–617.
393. Wattenberg, M.M.; Asch, D.; Yu, S.; O'Dwyer, P.J.; Domchek, S.M.; Nathanson, K.L.; Rosen, M.A.; Beatty, G.L.; Siegelman, E.S.; Reiss, K.A. Platinum response characteristics of patients with pancreatic ductal adenocarcinoma and a germline BRCA1, BRCA2 or PALB2 mutation. *Br. J. Cancer* **2020**, *122*, 333–339.
394. O'Reilly, E.M.; Lee, J.W.; Zalupski, M.; Capanu, M.; Park, J.; Golan, T.;

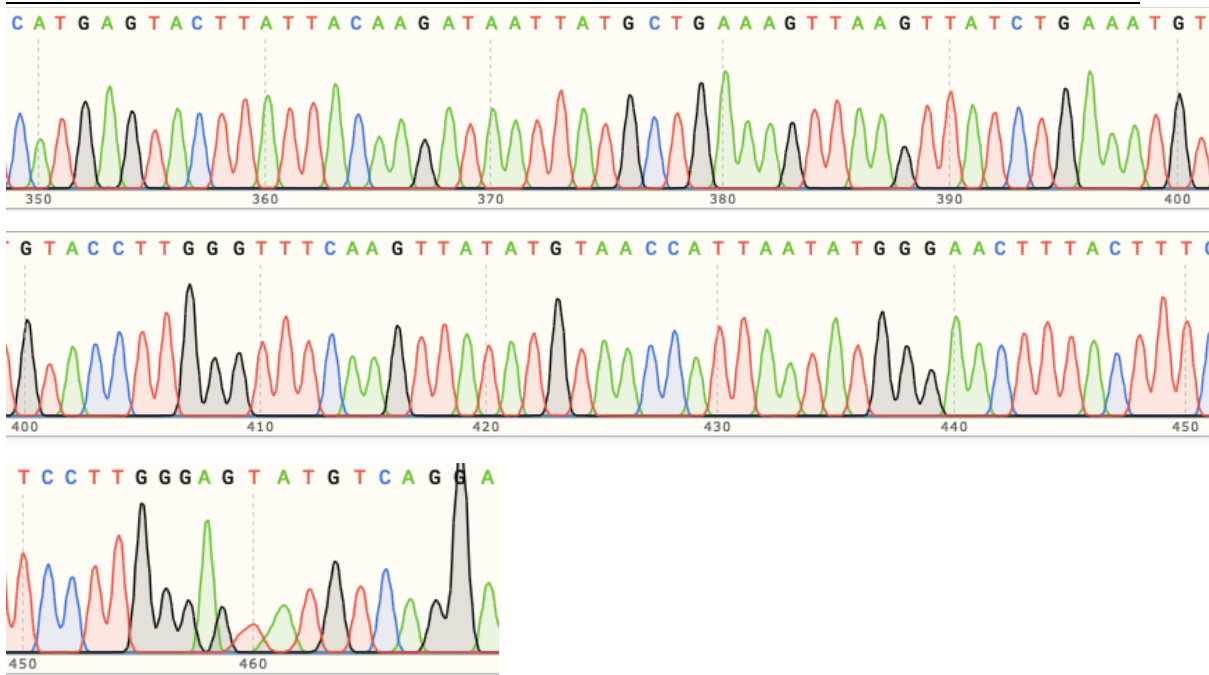
- Tahover, E.; Lowery, M.A.; Chou, J.F.; Sahai, V.; *et al.* Randomized, multicenter, phase II trial of gemcitabine and cisplatin with or without veliparib in patients with pancreas adenocarcinoma and a germline BRCA/PALB2 mutation. *J. Clin. Oncol.* **2020**, *38*, 1378.
395. Golan, T.; Kanji, Z.S.; Epelbaum, R.; Devaud, N.; Dagan, E.; Holter, S.; Aderka, D.; Paluch-Shimon, S.; Kaufman, B.; Gershoni-Baruch, R.; *et al.* Overall survival and clinical characteristics of pancreatic cancer in BRCA mutation carriers. *Br. J. Cancer* **2014**, *111*, 1132–1138.
396. O'Reilly, E.M.; Lee, J.W.; Lowery, M.A.; Capanu, M.; Stadler, Z.K.; Moore, M.J.; Dhani, N.; Kindler, H.L.; Estrella, H.; Maynard, H.; *et al.* Phase 1 trial evaluating cisplatin, gemcitabine, and veliparib in 2 patient cohorts: Germline BRCA mutation carriers and wild-type BRCA pancreatic ductal adenocarcinoma. *Cancer* **2018**, *127*, 1374–1382.
397. Hegewisch-Becker, S.; Aldaoud, A.; Wolf, T.; Krammer-Steiner, B.; Linde, H.; Scheiner-Sparna, R.; Hamm, D.; Jänicke, M.; Marschner, N. Results from the prospective German TPK clinical cohort study: Treatment algorithms and survival of 1,174 patients with locally advanced, inoperable, or metastatic pancreatic ductal adenocarcinoma. *Int. J. Cancer* **2019**, *144*, 981–990.
398. Nepomuceno, T.C.; De Gregoriis, G.; de Oliveira, F.M.B.; Suarez-Kurtz, G.; Monteiro, A.N.; Carvalho, M.A. The role of PALB2 in the DNA damage response and cancer predisposition. *Int. J. Mol. Sci.* **2017**, *18*, 1886.
399. Cruz, C.; Castroviejo-Bermejo, M.; Gutiérrez-Enríquez, S.; Llop-Guevara, A.; Ibrahim, Y.H.; Gris-Oliver, A.; Bonache, S.; Morancho, B.; Bruna, A.; Rueda, O.M.; *et al.* RAD51 foci as a functional biomarker of homologous recombination repair and PARP inhibitor resistance in germline BRCA-mutated breast cancer. *Ann. Oncol.* **2018**, *29*, 1203–1210.
400. Mladenov, E.; Anachkova, B.; Tsaneva, I. Sub-nuclear localization of Rad51 in response to DNA damage. *Genes to Cells* **2006**, *11*, 513–524.
401. Yuan, S.S.F.; Lee, S.Y.; Chen, G.; Song, M.; Tomlinson, G.E.; Lee, E.Y.H.P. BRCA2 is required for ionizing radiation-induced assembly of Rad51 complex in vivo. *Cancer Res.* **1999**, *59*, 3547–3551.

[This page is intentionally left blank]

Chapter 8. Appendix

8.1. Appendix A





Supplementary Figure 8.1: Representative chromatogram, image shown is KRAS gene in patient PDM106. Highlighted in blue is the G to T change, resulting in a G12V KRAS mutation.

8.2. Appendix B

Supplementary Table 8.1: Representative raw data of RT-qPCR experiments. All RT-qPCR experiments performed are of the same standard.

Well	Sample Name	Target Name	Reporter	Quencher	Ct	Ct Mean	Ct SD	Ct Threshold
A1	PT291 cell line P6	SOX	FAM	NFQ-MGB	24.8785999	24.7075999	0.1689116	0.0586177
A2	PT291 cell line P6	SOX	FAM	NFQ-MGB	24.7033424	24.7075999	0.1689116	0.0586177
A3	PT291 cell line P6	SOX	FAM	NFQ-MGB	24.5408573	24.7075999	0.1689116	0.0586177
A4	PT291 cell line P6	OCT	FAM	NFQ-MGB	26.9192219	26.6592770	0.3086243	0.1158579
A5	PT291 cell line P6	OCT	FAM	NFQ-MGB	26.7404232	26.6592770	0.3086243	0.1158579
A6	PT291 cell line P6	OCT	FAM	NFQ-MGB	26.3181877	26.6592770	0.3086243	0.1158579
A7	PT291 cell line P6	NANOG	FAM	NFQ-MGB	29.9890175	30.1819916	0.2596854	0.1912728
A8	PT291 cell line P6	NANOG	FAM	NFQ-MGB	30.4772434	30.1819916	0.2596854	0.1912728
A9	PT291 cell line P6	NANOG	FAM	NFQ-MGB	30.0797157	30.1819916	0.2596854	0.1912728
A10	PT291 cell line P6	18S	FAM	NFQ-MGB	8.8272219	8.8648494	0.1841387	0.1632130
A11	PT291 cell line P6	18S	FAM	NFQ-MGB	9.0648956	8.8648494	0.1841387	0.1632130

A12	PT291 cell line P6	18S	FAM	NFQ-MGB	8.7024307	8.8648494	0.1841387	0.1632130
B1	PT291 organoid P7	SOX	FAM	NFQ-MGB	32.4316254	32.2615821	0.3685893	0.0586177
B2	PT291 organoid P7	SOX	FAM	NFQ-MGB	32.5144539	32.2615821	0.3685893	0.0586177
B3	PT291 organoid P7	SOX	FAM	NFQ-MGB	31.8386669	32.2615821	0.3685893	0.0586177
B4	PT291 organoid P7	OCT	FAM	NFQ-MGB	26.2870865	25.9501724	0.3499882	0.1158579
B5	PT291 organoid P7	OCT	FAM	NFQ-MGB	25.9750004	25.9501724	0.3499882	0.1158579
B6	PT291 organoid P7	OCT	FAM	NFQ-MGB	25.5884323	25.9501724	0.3499882	0.1158579
B7	PT291 organoid P7	NANOG	FAM	NFQ-MGB	29.6711693	29.4440403	0.2267883	0.1912728
B8	PT291 organoid P7	NANOG	FAM	NFQ-MGB	29.4433632	29.4440403	0.2267883	0.1912728
B9	PT291 organoid P7	NANOG	FAM	NFQ-MGB	29.2175941	29.4440403	0.2267883	0.1912728
B10	PT291 organoid P7	18S	FAM	NFQ-MGB	8.0631304	7.7883134	0.4453900	0.1632130
B11	PT291 organoid P7	18S	FAM	NFQ-MGB	8.0273752	7.7883134	0.4453900	0.1632130
B12	PT291 organoid P7	18S	FAM	NFQ-MGB	7.2744365	7.7883134	0.4453900	0.1632130
C1	PT291 CLO P7	SOX	FAM	NFQ-MGB	25.8890095	25.6462307	0.2109226	0.0586177

C2	PT29I CLO P7	SOX	FAM	NFQ-MGB	25.5416393	25.6462307	0.2109226	0.0586177
C3	PT29I CLO P7	SOX	FAM	NFQ-MGB	25.5080433	25.6462307	0.2109226	0.0586177
C4	PT29I CLO P7	OCT	FAM	NFQ-MGB	26.2791519	25.9315281	0.3435091	0.1158579
C5	PT29I CLO P7	OCT	FAM	NFQ-MGB	25.9231453	25.9315281	0.3435091	0.1158579
C6	PT29I CLO P7	OCT	FAM	NFQ-MGB	25.5922871	25.9315281	0.3435091	0.1158579
C7	PT29I CLO P7	NANOG	FAM	NFQ-MGB	29.5342834	29.2291599	0.2647564	0.1912728
C8	PT29I CLO P7	NANOG	FAM	NFQ-MGB	29.0930500	29.2291599	0.2647564	0.1912728
C9	PT29I CLO P7	NANOG	FAM	NFQ-MGB	29.0601463	29.2291599	0.2647564	0.1912728
C10	PT29I CLO P7	I8S	FAM	NFQ-MGB	11.9179487	12.1283541	0.4469663	0.1632130
C11	PT29I CLO P7	I8S	FAM	NFQ-MGB	12.6416941	12.1283541	0.4469663	0.1632130
C12	PT29I CLO P7	I8S	FAM	NFQ-MGB	11.8254194	12.1283541	0.4469663	0.1632130
D1	Negative control	SOX	FAM	NFQ-MGB	Undetermined			0.0586177
D2	Negative control	SOX	FAM	NFQ-MGB	Undetermined			0.0586177
D3	Negative control	SOX	FAM	NFQ-MGB	Undetermined			0.0586177

D4	Negative control	OCT	FAM	NFQ-MGB	Undetermined	0.1158579
D5	Negative control	OCT	FAM	NFQ-MGB	Undetermined	0.1158579
D6	Negative control	OCT	FAM	NFQ-MGB	Undetermined	0.1158579
D7	Negative control	NANOG	FAM	NFQ-MGB	Undetermined	0.1912728
D8	Negative control	NANOG	FAM	NFQ-MGB	Undetermined	0.1912728
D9	Negative control	NANOG	FAM	NFQ-MGB	Undetermined	0.1912728
D10	Negative control	18S	FAM	NFQ-MGB	Undetermined	0.1632130
D11	Negative control	18S	FAM	NFQ-MGB	Undetermined	0.1632130
D12	Negative control	18S	FAM	NFQ-MGB	Undetermined	0.1632130

8.3. Appendix C

Supplementary Table 8.2: Genes, source of gene set from <https://maayanlab.cloud/Harmonizome/> and number of genes included in pathway for MODY gene RNA-seq analysis.

Gene	Source of gene set	Number of genes
<i>HNF1A</i>	TRANSFAC Curated Transcription Factor Targets	1131
<i>HNF1B</i>	GEO Signatures of Differentially Expressed Genes for Gene Perturbations	451
<i>HNF4A</i>	TRANSFAC Curated Transcription Factor Targets	887
<i>HNF4G</i>	ENCODE Transcription Factor Targets	4121
<i>GATA6</i>	TRANSFAC Curated Transcription Factor Targets	252
<i>NR5A2</i>	TRANSFAC Predicted Transcription Factor Targets	4968
<i>PDX1</i>	CHEA Transcription Factor Targets	638

8.4. Appendix D

Supplementary Table 8.3: Gene list – genes upregulated in both PT127 PDX and CLO when compared with PT127 cell line (fold change ≥ 2 , $p\text{-adj} < 0.001$).

<i>Gene Name</i>	<i>Log2Fold Change</i>	<i>p-adj</i>	<i>Gene Name</i>	<i>Log2Fold Change</i>	<i>p-adj</i>
<i>KIF12</i>	5.422	1.62E-62	<i>SEMA6A</i>	3.16	1.90E-06
<i>CAI2</i>	5.302	1.44E-19	<i>MEX3A</i>	3.125	2.64E-08
<i>LRIG1</i>	5.286	5.37E-57	<i>DEPTOR</i>	3.115	2.43E-52
<i>LTB</i>	5.263	8.48E-05	<i>EPHX2</i>	3.031	2.45E-18
<i>FAM3B</i>	5.004	6.01E-51	<i>LARGE2</i>	2.996	5.92E-06
<i>TNFRSF11B</i>	4.76	5.34E-09	<i>TMEM229B</i>	2.979	9.77E-13
<i>SLC5A1</i>	4.749	1.95E-50	<i>100288152</i>	2.975	3.29E-08
<i>LCN2</i>	4.551	4.24E-12	<i>MAP2K6</i>	2.963	4.38E-10
<i>PLEKHB1</i>	4.505	4.08E-13	<i>ATPI0B</i>	2.953	7.15E-42
<i>DEGS2</i>	4.455	1.86E-07	<i>GPX2</i>	2.941	1.01E-31
<i>FAM222A</i>	4.176	1.96E-08	<i>SLCIA4</i>	2.927	1.58E-16
<i>NECTIN3</i>	3.92	4.77E-14	<i>CRACR2A</i>	2.925	1.25E-05
<i>SELENBP1</i>	3.873	1.29E-12	<i>RNASEI</i>	2.913	1.44E-08
<i>GJB1</i>	3.831	1.06E-12	<i>SELENOM</i>	2.889	1.28E-21
<i>FILIPIL</i>	3.795	1.92E-14	<i>RXFP4</i>	2.882	2.53E-09
<i>MYH7B</i>	3.776	5.46E-07	<i>RGMB</i>	2.871	3.57E-18
<i>PLBI</i>	3.748	2.16E-24	<i>SEZ6L2</i>	2.867	1.46E-07
<i>NROB2</i>	3.738	4.10E-09	<i>SLC4A8</i>	2.855	6.12E-06
<i>PLEKHSI</i>	3.669	3.63E-28	<i>MLLT3</i>	2.845	1.31E-19
<i>PLA2G4F</i>	3.638	1.05E-24	<i>BDH2</i>	2.839	3.17E-10
<i>CLDN2</i>	3.594	4.09E-79	<i>EPHB3</i>	2.809	6.31E-11
<i>SLC9A3</i>	3.539	9.80E-12	<i>CRISPLD2</i>	2.797	0.00023
<i>C9orf152</i>	3.527	7.66E-12	<i>PDZK1</i>	2.786	6.55E-13
<i>SMIM24</i>	3.516	3.49E-09	<i>TPPP3</i>	2.755	3.95E-06
<i>KLK1</i>	3.488	8.79E-23	<i>SP5</i>	2.753	0.00019
<i>CCDC40</i>	3.452	1.39E-11	<i>CCDC170</i>	2.751	1.43E-12
<i>RSPH1</i>	3.418	1.59E-45	<i>MDK</i>	2.746	1.34E-10
<i>PSTPIP2</i>	3.394	4.18E-08	<i>PAQR6</i>	2.738	0.00074
<i>ITPR2</i>	3.386	5.38E-07	<i>MAPILC3A</i>	2.73	2.32E-14
<i>BRSK2</i>	3.365	0.00015	<i>RAB30</i>	2.718	1.91E-05
<i>PTGER4</i>	3.344	1.79E-19	<i>SOWAHA</i>	2.699	2.31E-14
<i>DAPK2</i>	3.338	6.65E-10	<i>NEDD9</i>	2.694	2.88E-05
<i>LINC01124</i>	3.277	1.35E-07	<i>ACSSI</i>	2.682	5.28E-20
<i>ZSWIM5</i>	3.245	5.34E-20	<i>MYCL</i>	2.673	5.40E-09
<i>AMACR</i>	3.235	1.17E-18	<i>BCL2L14</i>	2.646	0.0002
<i>IGSF9</i>	3.175	2.46E-18	<i>HNFA-ASI</i>	2.643	5.24E-06
<i>GPC2</i>	2.64	1.22E-06	<i>ALDH1A1</i>	2.259	2.75E-46
<i>SIAE</i>	2.64	4.95E-06	<i>GATM</i>	2.258	4.67E-09

<i>100124692</i>	2.634	5.42E-34	<i>ABAT</i>	2.244	2.45E-10
<i>KREMENI</i>	2.604	4.31E-06	<i>SOX9-ASI</i>	2.227	8.80E-05
<i>TSPOAPI</i>	2.579	1.86E-08	<i>FOXD4</i>	2.209	1.25E-05
<i>SELENOW</i>	2.579	1.13E-18	<i>ESPN</i>	2.19	2.55E-09
<i>INPP5J</i>	2.543	7.16E-07	<i>MYLIP</i>	2.189	1.25E-25
<i>IFI6</i>	2.534	2.67E-05	<i>CCSERI</i>	2.187	0.00064
<i>MCF2L</i>	2.533	2.81E-10	<i>VIPRI</i>	2.178	7.25E-14
<i>RAB26</i>	2.519	4.82E-05	<i>RN7SK</i>	2.167	1.06E-08
<i>MGAM2</i>	2.517	3.09E-07	<i>SERPINA4</i>	2.162	1.75E-07
<i>INKA2</i>	2.505	2.42E-14	<i>IFI27</i>	2.151	0.00094
<i>ORMDL3</i>	2.505	2.02E-09	<i>CEMIP2</i>	2.148	1.16E-07
<i>MAPK8IP1</i>	2.498	2.49E-06	<i>TESC</i>	2.138	2.12E-06
<i>TNFAIP2</i>	2.496	1.92E-08	<i>RBBP8NL</i>	2.134	4.95E-17
<i>GRAMD2A</i>	2.49	5.35E-16	<i>MMP1I</i>	2.129	3.08E-08
<i>TGFB3</i>	2.452	0.00093	<i>CES3</i>	2.122	1.97E-11
<i>PDE4D</i>	2.445	1.06E-08	<i>AKRIC3</i>	2.118	2.76E-16
<i>TOX3</i>	2.44	0.00029	<i>STXBP6</i>	2.107	3.33E-11
<i>ALDHIB1</i>	2.438	8.55E-53	<i>MLEC</i>	2.102	1.18E-11
<i>HS6ST1</i>	2.417	6.87E-08	<i>QSOX1</i>	2.095	1.20E-14
<i>SLC29A4</i>	2.4	0.00014	<i>NFIA</i>	2.086	2.62E-12
<i>AIFM3</i>	2.394	6.63E-10	<i>CHST6</i>	2.07	2.07E-11
<i>ARHGEF40</i>	2.394	3.02E-05	<i>ARRB2</i>	2.067	3.82E-16
<i>SCAMP5</i>	2.376	1.26E-12	<i>STK31</i>	2.063	9.70E-06
<i>FAAH</i>	2.375	5.63E-10	<i>CTSS</i>	2.063	1.18E-08
<i>KCTDI7</i>	2.36	1.83E-06	<i>ZNF233</i>	2.05	5.37E-05
<i>FOXA3</i>	2.36	1.43E-12	<i>LRRC73</i>	2.048	1.69E-05
<i>ETS2</i>	2.352	2.60E-32	<i>MYB</i>	2.047	1.51E-14
<i>GSDMB</i>	2.327	0.00041	<i>RAB11FIP3</i>	2.044	3.49E-05
<i>SGSH</i>	2.324	4.02E-08	<i>100129046</i>	2.029	3.29E-14
<i>PAPSS2</i>	2.323	0.00042	<i>TEAD2</i>	2.025	2.91E-17
<i>PYCARD</i>	2.312	1.55E-06	<i>PLCB4</i>	2.025	0.000444
<i>ASRGL1</i>	2.276	7.24E-08	<i>COROIA</i>	2.019	1.13E-31
<i>FKBP7</i>	2.267	0.00013	<i>ACCS</i>	2.006	7.84E-05
<i>PLPP3</i>	2.266	0.00066	<i>SYK</i>	2.002	1.96E-08

Supplementary Table 8.4: Gene list – genes downregulated in both PT127 PDX and CLO when compared with PT127 cell line (fold change ≥ 2 , p-adj < 0.001).

<i>Gene List</i>	<i>log2Fold Change</i>	<i>p-adj</i>	<i>Gene List</i>	<i>log2Fold Change</i>	<i>p-adj</i>
<i>FOSB</i>	-7.241	1.82E-55	<i>HBEGF</i>	-3.295	1.56E-07
<i>CAVI</i>	-7.054	1.16E-101	<i>MYBL1</i>	-3.27	5.46E-06
<i>AXL</i>	-6.625	6.32E-262	<i>CTGF</i>	-3.252	4.80E-14
<i>RAB3B</i>	-6.322	1.39E-60	<i>ERRF1</i>	-3.243	6.42E-14
<i>FOS</i>	-5.132	1.13E-31	<i>MICB</i>	-3.203	1.24E-11
<i>PAD1</i>	-4.767	1.24E-51	<i>SMAD7</i>	-3.144	2.10E-10
<i>DUSP1</i>	-4.67	2.09E-42	<i>TSC22D3</i>	-3.136	1.27E-12
<i>EHD2</i>	-4.651	2.54E-09	<i>THSD4</i>	-3.083	2.45E-21
<i>EGRI</i>	-4.637	3.22E-50	<i>KRT80</i>	-3.059	9.81E-55
<i>ATF3</i>	-4.383	4.50E-09	<i>SLC7A7</i>	-3.044	1.05E-16
<i>TGM2</i>	-4.348	3.35E-176	<i>PMAIP1</i>	-3.024	3.98E-15
<i>CYR61</i>	-4.301	1.62E-62	<i>652995</i>	-3.02	5.44E-17
<i>OXTR</i>	-4.236	1.98E-10	<i>ID2</i>	-3.014	2.31E-11
<i>PHACTR3</i>	-4.21	1.13E-31	<i>RGCC</i>	-2.952	1.82E-29
<i>NR4A1</i>	-4.174	6.61E-11	<i>LICAM</i>	-2.873	2.27E-07
<i>FOSL1</i>	-3.986	2.47E-35	<i>PLAUR</i>	-2.868	5.74E-06
<i>KRT86</i>	-3.876	1.44E-19	<i>GADD45B</i>	-2.86	8.18E-09
<i>SUSD2</i>	-3.862	4.44E-26	<i>SLC16A4</i>	-2.814	8.15E-26
<i>AKAP12</i>	-3.796	1.86E-64	<i>ARHGAP29</i>	-2.813	2.58E-08
<i>ADGRF1</i>	-3.731	9.83E-27	<i>LAT2</i>	-2.792	3.03E-32
<i>LCK</i>	-3.681	2.45E-21	<i>PEAR1</i>	-2.791	2.23E-05
<i>CAVIN1</i>	-3.66	1.43E-78	<i>FGFBP1</i>	-2.787	0.00016
<i>PLK2</i>	-3.623	2.30E-40	<i>ANO1</i>	-2.774	1.11E-36
<i>LGALS1</i>	-3.598	4.75E-45	<i>ZFP36</i>	-2.761	1.81E-05
<i>284454</i>	-3.584	1.52E-09	<i>WNT7B</i>	-2.753	2.44E-11
<i>ARHGD1B</i>	-3.58	2.63E-34	<i>NCEH1</i>	-2.711	4.27E-23
<i>EDN1</i>	-3.579	1.43E-64	<i>NTN4</i>	-2.633	6.66E-05
<i>UNC13D</i>	-3.472	1.51E-27	<i>MIR4435-2HG</i>	-2.628	7.08E-32
<i>MMP7</i>	-3.463	1.17E-08	<i>STEAP4</i>	-2.623	3.44E-21
<i>ELFN2</i>	-3.455	1.30E-06	<i>SNORC</i>	-2.621	9.24E-14
<i>ANXA1</i>	-3.415	4.82E-31	<i>CAVIN3</i>	-2.618	5.03E-16
<i>CACNG4</i>	-3.391	2.59E-18	<i>ANXA2R</i>	-2.603	4.38E-10
<i>PLAU</i>	-3.355	5.27E-60	<i>CYTOR</i>	-2.583	5.81E-09
<i>KLF2</i>	-3.347	1.90E-17	<i>RNF39</i>	-2.571	1.95E-19

<i>CAV2</i>	-3.346	4.90E-24	<i>KLF6</i>	-2.552	2.57E-05
<i>GPAT3</i>	-3.297	3.55E-28	<i>ANLN</i>	-2.544	3.89E-17
<i>KLK7</i>	-2.532	0.0007	<i>CDKN2B</i>	-2.212	1.41E-05
<i>ID3</i>	-2.517	2.16E-05	<i>100506906</i>	-2.206	0.00088
<i>ERFE</i>	-2.493	6.50E-08	<i>EVAIC</i>	-2.205	7.97E-24
<i>100287314</i>	-2.489	7.83E-16	<i>DUSP5</i>	-2.19	0.00055
<i>SGKI</i>	-2.486	7.57E-21	<i>LAT</i>	-2.182	1.07E-14
<i>TMPRSS13</i>	-2.476	0.00031	<i>BCAR3</i>	-2.176	9.14E-32
<i>CLCFI</i>	-2.457	2.64E-08	<i>ITGB2</i>	-2.171	1.16E-07
<i>ATP2B4</i>	-2.452	1.11E-13	<i>CXCL2</i>	-2.162	1.06E-06
<i>DUSP10</i>	-2.448	1.77E-39	<i>GRHL3</i>	-2.145	2.98E-23
<i>IL6R</i>	-2.393	3.62E-24	<i>TTYHI</i>	-2.144	1.43E-05
<i>OLFML2A</i>	-2.376	4.05E-07	<i>CITED2</i>	-2.143	1.99E-29
<i>FAM122B</i>	-2.376	2.04E-12	<i>HMGAI</i>	-2.14	1.89E-15
<i>MSMOI</i>	-2.363	3.79E-09	<i>PTAFR</i>	-2.139	4.88E-19
<i>SMURF2</i>	-2.356	6.07E-16	<i>B3GNTL1</i>	-2.132	6.95E-11
<i>TNNT1</i>	-2.347	2.70E-08	<i>AMIGO2</i>	-2.13	1.33E-17
<i>DCBLD2</i>	-2.337	1.31E-05	<i>KRT7</i>	-2.11	0.0006
<i>221584</i>	-2.336	9.97E-10	<i>RBMS2</i>	-2.096	2.03E-10
<i>CHAC1</i>	-2.331	4.38E-07	<i>FAM3C</i>	-2.078	7.55E-22
<i>DUSP6</i>	-2.327	1.22E-19	<i>SLC20A1</i>	-2.068	1.79E-07
<i>LIF</i>	-2.319	4.58E-05	<i>HESI</i>	-2.066	2.02E-15
<i>MOSPD1</i>	-2.318	2.88E-09	<i>DENND3</i>	-2.063	1.00E-06
<i>ARNTL2</i>	-2.299	4.73E-07	<i>KIAA0040</i>	-2.047	0.00049
<i>ZNFI65</i>	-2.291	1.88E-33	<i>HDAC9</i>	-2.046	3.15E-13
<i>FAM122C</i>	-2.291	5.25E-15	<i>VSIR</i>	-2.044	1.01E-09
<i>643977</i>	-2.289	1.15E-13	<i>THBS1</i>	-2.041	8.25E-18
<i>TIMP2</i>	-2.286	9.19E-49	<i>SI00A2</i>	-2.015	4.64E-06
<i>ARSK</i>	-2.263	1.54E-05	<i>KIF20B</i>	-2.011	0.00038
<i>F3</i>	-2.255	1.34E-07	<i>HMGCSI</i>	-2.005	2.24E-07
<i>SEMA7A</i>	-2.223	2.26E-06	<i>ETSI</i>	-2	2.01E-05

Supplementary Table 8.5: Gene list – genes upregulated in both PT291 PDX and CLO when compared with PT291 cell line (fold change ≥ 2 , p-adj < 0.001).

<i>Gene Name</i>	<i>log2Fold Change</i>	<i>p-adj</i>	<i>Gene Name</i>	<i>log2Fold Change</i>	<i>p-adj</i>
<i>SERPINA5</i>	4.212	2.17E-14	<i>RCBTB2</i>	2.434	2.55E-11
<i>CCDC114</i>	4.17	4.30E-22	<i>SEMA3F</i>	2.362	2.31E-23
<i>PTK7</i>	4.061	2.95E-27	<i>STXBP6</i>	2.343	1.99E-11
<i>FCGBP</i>	3.931	1.87E-29	<i>TMEM97</i>	2.316	0.0007
<i>CCDC170</i>	3.607	4.53E-13	<i>LRRC73</i>	2.315	2.94E-11
<i>ANPEP</i>	3.575	1.43E-09	<i>ITPR2</i>	2.306	1.32E-05
<i>TFF3</i>	3.512	3.51E-11	<i>NAALADL2</i>	2.272	5.79E-14
<i>NECTIN3</i>	3.469	4.17E-09	<i>RASL11A</i>	2.272	7.44E-05
<i>NDRG1</i>	3.461	1.94E-11	<i>APBA1</i>	2.271	4.59E-08
<i>SERPINA4</i>	3.414	2.62E-13	<i>FHADI</i>	2.271	6.05E-10
<i>PPIL6</i>	3.215	9.37E-07	<i>FAM227A</i>	2.262	5.65E-10
<i>FOXJ1</i>	3.037	2.97E-10	<i>CLU</i>	2.239	1.20E-05
<i>TGFB3</i>	3.015	8.47E-19	<i>RNASEI</i>	2.23	4.96E-05
<i>SLC29A4</i>	2.986	3.75E-07	<i>DEGS2</i>	2.214	1.16E-08
<i>PLEKHBI</i>	2.89	2.69E-14	<i>RSPHI</i>	2.202	2.63E-08
<i>CFAP73</i>	2.84	4.88E-07	<i>HNFI A-AS1</i>	2.194	9.64E-07
<i>DEPTOR</i>	2.811	2.36E-09	<i>PTPRN2</i>	2.192	3.33E-13
<i>EGLN3</i>	2.794	1.90E-22	<i>SYNM</i>	2.179	1.02E-05
<i>GCNT1</i>	2.778	1.82E-09	<i>C4A</i>	2.176	7.21E-08
<i>PPMIL</i>	2.771	3.59E-12	<i>PPPIR1B</i>	2.151	1.64E-20
<i>LRIG1</i>	2.709	8.57E-11	<i>USP2</i>	2.146	6.50E-06
<i>FKBP9P1</i>	2.685	5.22E-05	<i>TMEM231</i>	2.144	1.60E-05
<i>PRSS33</i>	2.646	0.00026	<i>C4B</i>	2.131	8.57E-08
<i>TRPM5</i>	2.596	0.00055	<i>ST6GALNAC1</i>	2.128	4.31E-05
<i>100506211</i>	2.581	3.46E-09	<i>MESPI</i>	2.083	6.24E-09
<i>CCDC40</i>	2.58	4.78E-08	<i>SNEDI</i>	2.082	2.69E-13
<i>WDR78</i>	2.576	2.32E-06	<i>ALDOC</i>	2.066	0.00069
<i>GNAZ</i>	2.512	8.39E-10	<i>PROM1</i>	2.05	1.70E-12
<i>FAM222A</i>	2.509	1.38E-07	<i>HIST1H4K</i>	2.043	1.65E-06
<i>LINC00261</i>	2.5	4.31E-05	<i>SLC4A8</i>	2.042	2.91E-07
<i>ASRGL1</i>	2.444	2.60E-06	<i>FAM110C</i>	2.018	1.47E-06

Supplementary Table 8.6: Gene list – genes downregulated in both PT291 PDX and CLO when compared with PT291 cell line (fold change ≥ 2 , $p\text{-adj} < 0.001$).

<i>Gene Name</i>	<i>log2Fold Change</i>	<i>p-adj</i>	<i>Gene Name</i>	<i>log2Fold Change</i>	<i>p-adj</i>
<i>AHNAK2</i>	-5.316	1.28E-110	<i>GPAT3</i>	-2.828	6.09E-09
<i>AXL</i>	-5.301	2.00E-44	<i>PHACTR3</i>	-2.787	5.07E-06
<i>COL17A1</i>	-4.747	0.0002	<i>EMPI</i>	-2.774	6.03E-25
<i>CDHR2</i>	-4.682	4.55E-11	<i>GJB3</i>	-2.763	2.63E-34
<i>ECM1</i>	-4.628	9.92E-49	<i>PLAUR</i>	-2.757	3.59E-17
<i>FOSB</i>	-4.491	2.55E-11	<i>MDFI</i>	-2.757	2.68E-12
<i>ID3</i>	-4.263	8.26E-17	<i>GEM</i>	-2.736	2.77E-18
<i>SEMA7A</i>	-3.985	5.67E-39	<i>IL18</i>	-2.683	1.84E-10
<i>IL23A</i>	-3.926	1.92E-29	<i>652995</i>	-2.681	2.51E-05
<i>AKAP12</i>	-3.887	3.40E-11	<i>DCBLD2</i>	-2.678	8.37E-11
<i>TM4SF4</i>	-3.865	0.00055	<i>ELFN2</i>	-2.639	1.03E-07
<i>FLNC</i>	-3.848	1.75E-08	<i>SMIM5</i>	-2.634	0.00094
<i>CDHI7</i>	-3.757	0.00044	<i>SDR16C5</i>	-2.617	9.01E-06
<i>LAT2</i>	-3.757	2.11E-46	<i>OXTR</i>	-2.611	9.83E-06
<i>CAVIN1</i>	-3.704	8.30E-24	<i>PCED1B</i>	-2.605	7.03E-05
<i>UNC13D</i>	-3.629	8.24E-24	<i>CYP3A5</i>	-2.588	1.02E-05
<i>SNCG</i>	-3.594	4.14E-08	<i>RASA3</i>	-2.555	1.78E-06
<i>SULT2BI</i>	-3.551	1.25E-11	<i>GPSMI</i>	-2.514	0.00032
<i>LGALS1</i>	-3.437	1.46E-26	<i>UNC5B</i>	-2.485	0.0006
<i>BTBD11</i>	-3.424	6.91E-09	<i>PLEKHNI</i>	-2.468	2.97E-10
<i>RAB3B</i>	-3.422	1.57E-10	<i>P2RY2</i>	-2.443	3.57E-07
<i>FOSL1</i>	-3.414	2.08E-10	<i>ETSI</i>	-2.436	8.63E-08
<i>PAD11</i>	-3.414	1.48E-10	<i>CRYBG2</i>	-2.398	4.15E-06
<i>SDCBP2</i>	-3.338	1.90E-08	<i>ARL4C</i>	-2.394	4.33E-09
<i>CAVI</i>	-3.321	3.16E-08	<i>TNFAIP3</i>	-2.394	8.60E-05
<i>IDI</i>	-3.267	3.22E-07	<i>KLK11</i>	-2.391	3.57E-07
<i>AREG</i>	-3.23	2.10E-48	<i>CXCL8</i>	-2.391	9.14E-10
<i>DUSP1</i>	-3.224	2.53E-16	<i>DUSP10</i>	-2.375	1.28E-07
<i>727738</i>	-3.172	2.10E-48	<i>IL32</i>	-2.373	1.11E-06
<i>F3</i>	-3.047	9.10E-15	<i>8511</i>	-2.352	8.08E-05
<i>CPNE7</i>	-2.962	2.07E-06	<i>ADAMTSL5</i>	-2.34	1.17E-08
<i>ABLIM3</i>	-2.933	2.95E-27	<i>PTPRR</i>	-2.336	5.46E-09
<i>CYP4F3</i>	-2.925	0.00035	<i>SYNGR3</i>	-2.31	0.00077
<i>PLAU</i>	-2.918	1.10E-17	<i>PSORSIC1</i>	-2.296	1.16E-06

<i>SMAD7</i>	-2.908	3.01E-18	<i>PHLDB2</i>	-2.277	1.48E-09
<i>AMIGO2</i>	-2.875	3.99E-21	<i>MMP7</i>	-2.247	1.24E-06
<i>CD55</i>	-2.236	1.34E-13	<i>RND1</i>	-2.089	0.0003
<i>MMP23B</i>	-2.226	7.11E-05	<i>TNFRSF11B</i>	-2.084	8.31E-11
<i>ITGA2</i>	-2.221	7.93E-26	<i>SERPINB5</i>	-2.083	8.15E-05
<i>STOM</i>	-2.217	2.66E-07	<i>SLC22A3</i>	-2.081	7.58E-15
<i>KIAA0040</i>	-2.189	1.50E-06	<i>SI00A2</i>	-2.07	3.58E-06
<i>MTMRII</i>	-2.186	1.81E-05	<i>DUSP5</i>	-2.058	4.04E-08
<i>PERM1</i>	-2.178	4.58E-06	<i>ARL14.00</i>	-2.058	0.0002
<i>PROCR</i>	-2.177	9.52E-11	<i>FSCN1</i>	-2.054	0.00066
<i>HEPH</i>	-2.163	4.92E-05	<i>284454</i>	-2.053	2.58E-10
<i>IRS2</i>	-2.145	1.60E-15	<i>KLK10</i>	-2.043	4.77E-07
<i>KLK7</i>	-2.129	0.00032	<i>CLCF1</i>	-2.041	1.10E-06
<i>SH3TC2</i>	-2.12	1.15E-19	<i>LMTK3</i>	-2.035	0.00051
<i>PLCXD2</i>	-2.119	0.0003	<i>SMIM6</i>	-2.019	2.29E-05
<i>PTPRB</i>	-2.114	1.28E-24	<i>CAV2</i>	-2.012	1.42E-14
<i>LIF</i>	-2.103	1.34E-34	<i>TNFRSF6B</i>	-2.006	9.01E-05
<i>PRDMI</i>	-2.102	2.46E-07	<i>EPHB6</i>	-2.004	2.91E-07
<i>PIMI</i>	-2.094	2.07E-09			