

INVESTIGATING THE EFFECTS OF BIOMARKERS CHRDL2 AND POLD3 ON COLORECTAL CANCER RISK

A thesis submitted for the degree of Doctor of
Philosophy

by Eloise Clarkson

Department of Life Sciences, Brunel University London

Acknowledgments

First and foremost, I would like to thank my PhD supervisor, Dr Annabelle Lewis. You have been a wonderful mentor, boss, and friend to me throughout my PhD. You gave me the freedom and confidence to take charge of my own research, supporting me throughout my time at Brunel. Not only did you push me to achieve things I wouldn't have done without your encouragement, your kindness and understanding made my PhD an amazing experience. You have taught me so much in the past 3 years, and I will be forever grateful to have had you as my PhD supervisor.

I would also like to thank Dr Cristina Pina, for her excellent advice and theoretical knowledge which made my research the best it could be. You pushed me to have a vigorous research method, and without your guidance much of my research would not have been possible.

To my fellow lab group members, Samantha Jerin and Emine Efendi, I would like to thank you for your help and support. I would also like to thank all my wonderful colleagues at Brunel, your advice on experiments helped my research, but more importantly your friendships allowed my PhD to be full of laughter and helped to keep me sane during stressful periods.

Thank you to Dr Terry Roberts and Sarah Aldulaimi for their guidance on c-circle and telomere analysis, and Dr Hayley Belnoue-Davis for her hand in helping me develop our organoid models.

Finally, I would like to thank my parents and family. You have always supported me, both financially, emotionally, and through providing career advice. Especially to my mum, who at my age was finishing her PhD while also welcoming me into the world. I could have never accomplished what I have without you pushing me to be the best I could be, and through showing me first-hand what can be achieved. You have always believed in my abilities even when I have not. I have looked up to you both more than you will ever realise. I hope I have made you proud.

Abstract

Colorectal cancer (CRC) is one of the leading causes of cancer globally. Genome-wide association studies (GWAS) have identified several loci with single nucleotide polymorphisms that may increase CRC risk, including adjacent regions near the *POLD3* and *CHRD2* genes. This thesis investigates the functional role these genes and their encoded proteins play in CRC development.

Our research utilized cancer cell lines grown as 2D cultures engineered to overexpress *CHRD2* and knockdown *POLD3* expression, alongside 3D organoid models treated with *CHRD2*, supplemented by RNA sequencing.

Chordin-like-2 (*CHRD2*) is a secreted BMP antagonist with overexpression linked to poor CRC prognosis. However, its functional role remains unclear. BMP signalling promotes differentiation and counters WNT signalling, which sustains stem-cell maintenance in the intestinal epithelium. Reduced BMP and elevated WNT signalling leads to an increased stem-cell phenotype, a hallmark of CRC. We have found that *CHRD2* overexpression reduces BMP signalling and elevates WNT signalling, enhancing cancer cell survival during chemotherapy and irradiation by activation of DNA damage response pathways. Treated cells showed increased stem-cell markers and reduced differentiation, suggesting *CHRD2*'s potential to intensify the cancer stem-cell phenotype and influence therapy response.

The second gene investigated, *POLD3*, is the third subunit of Pol δ and plays a critical role in DNA synthesis and repair, which are vital for cancer cells. Knock-down of *POLD3* in CRC cells reduced their proliferative ability, causing cell cycle arrest in S phase, leading to apoptosis and cell death. Furthermore, *POLD3* was shown to be essential for ALT telomere maintenance, with knockdown of *POLD3* reducing telomere length, a key factor affecting cancer cell longevity. Additionally, *POLD3* knockdown induced DNA damage and instability, and increased sensitivity to chemotherapy and irradiation.

This research highlights the roles of *POLD3* and *CHRD2* dysregulation in CRC, offering new insights into their potential as therapeutic targets and CRC biomarkers.

Table of contents

Acknowledgments	2
Abstract.....	3
Table of contents	3
Chapter 1: Introduction.....	8
Colorectal cancer	8
Colorectal cancer biology	9
Clinical presentation and diagnosis	9
Diagnosis	9
Treatments	10
Colorectal cancer pathophysiology	11
Consensus of molecular subtypes (CMS)	13
Genetic and epigenetic pathways of colorectal cancer	13
Microsatellite instability / mismatch repair	14
CpG island methylator phenotype	15
WNT pathway mutations and Chromosomal instability	15
The WNT/BMP signalling axis	20
BMP signalling in colorectal cancer	23
WNT and BMP signalling and Cancer stem cells	24
BMP antagonists	25
DNA polymerases and Colorectal cancer	30
Metastasis	35
Tumour microenvironment and immune landscape	36
GWAS studies in CRC patients	39
Project Background and rationale	41
Genome wide association studies	41
CHRDL2	42
POLD3	43
Project aims	45
Chapter 2: Methods	46
Cell culture and maintenance	46
Overexpression of CHRDL2	46
Knockdown of POLD3	47
.....	48
Vector expansion and purification	48
Oligo preparation	48
Plasmid Ligation	49

Transformation	49
Mini prep and restriction digest of ligated clones	50
Colony PCR	50
Plasmid sequencing	50
Wound healing assay	51
Migration assay	52
Clonogenic assay	52
Drug dose response curve	53
Organoid preparation and maintenance	55
Organoid immunofluorescence staining	56
Flow cytometry	57
Western blot	57
<i>Sample preparation</i>	57
<i>Bicinchoninic acid (BCA) assay</i>	58
<i>Protein separation and membrane transfer</i>	58
<i>Blocking and antibody staining</i>	58
<i>Imaging</i>	59
<i>Loading control</i>	59
Immunofluorescence	59
<i>Coverslip preparation</i>	59
<i>Cell preparation</i>	59
<i>Immunofluorescence staining</i>	59
Comet assay	61
<i>Sample preparation</i>	62
<i>RNA extraction</i>	62
<i>cDNA synthesis</i>	62
<i>Real-time polymerase chain reaction (qPCR)</i>	63
Telomere Length analysis	64
Statistical analysis	65
Bioinformatics	65
Analysis of RNA-seq data	65
Data visualisation and R	66
GSEA	66
Chapter 3: Effects of CHRDL2 overexpression on 2D CRC cell lines and 3D models	67
Introduction	67
Establishing CHRDL2 cell lines	69

CHRDL2 overexpression increases WNT signalling through β-catenin nuclear localisation.....	73
Overexpression of CHRDL2 decreases cellular proliferation and slows cell cycle progression.....	74
.....	77
CHRDL2 decreases colony size and increases migration.....	78
CHRDL2 increases migration and adherence markers.....	82
The effects of CHRDL2 on 3D organoid models.....	86
Discussion.....	91
Limitations and future work.....	95
Conclusion.....	96
Chapter 4: CHRDL2 and chemotherapy resistance.....	96
Introduction.....	96
CHRDL2 increases resistance to chemotherapy.....	99
CHRDL2 overexpression reduces cell cycle stalling during chemotherapy.....	102
Secreted CHRDL2 increases chemotherapy resistance.....	104
CHRDL2 addition increases chemotherapy resistance in intestinal organoids.....	106
CHRDL2 increases radiation resistance.....	108
CHRDL2 overexpressing cells have high BMP antagonism following chemotherapy treatment.....	109
CHRDL2 decreases DNA damage through upregulation of DNA repair pathways during chemotherapy.....	110
CHRDL2 upregulates DNA repair mechanisms.....	114
Discussion.....	119
Limitations and future work.....	123
Conclusion.....	123
Chapter 5: Gene expression analysis and mechanisms of CHRDL2.....	124
Introduction.....	124
CHRDL2 predicts poor survival in late-stage CRC.....	125
CHRDL2 induces cancer-related genes during overexpression.....	128
Identification and validation of differentially expressed genes by CHRDL2.....	132
Upregulated genes.....	137
Downregulated genes.....	138
Conflicting genes.....	140
CHRDL2 regulates cancer pathways during overexpression.....	141
Discussion.....	145
Limitations and future work.....	147

Conclusion	147
Chapter 6: The role of POLD3 knockdown in CRC	148
Introduction	148
Establishing POLD3 knockdown cell lines	151
POLD3 knockdown reduces proliferation and clonogenic ability	154
POLD3 knockdown induces senescence and apoptosis	156
POLD3 knockdown sensitises cells to chemotherapy	158
POLD3 knockdown causes DNA damage	160
POLD3 knockdown induced DNA damage results in upregulation of DNA repair ..	162
POLD3 knockdown reduces telomere length	164
POLD3 expression in patient tumours and survivability	166
Discussion	168
Limitations and future work	171
Conclusion	171
Chapter 7: The effects of POLD3 and CHRDL2 in Colorectal cancer	173
POLD3 and CHRDL2 do not correlate in expression	174
POLD3 knockdown and CHRDL2 overexpression results in a reduction of cell proliferation	176
Discussion	178
Limitations and future work	179
Conclusion	179
Chapter 8: Thesis Discussion.....	180
Conclusion	187
References	188
Appendices.....	209
Linux scripts for RNAseq analysis and data processing	209
R studio analysis for data visualization	210

Chapter 1: Introduction

Colorectal cancer

Colorectal cancer (CRC) represents a major global health challenge. Ranking as the third most common form of cancer, it is responsible for over 1.5 million new cases annually [1]. Accounting for 10% of all new cancer cases, CRC is the second leading cause of cancer related death, at over 900,000 incidences in 2020 [1]. CRC is a type of gastrointestinal malignancy that forms in either the colon or rectum. More than 90% of CRCs present as adenocarcinomas, originating from epithelial cells in the mucosal layer. Other rare types of CRCs include neuroendocrine (<1%), squamous cell (0.1-2.5%), Adenosquamous (0.06-0.18%), spindle cell (<1%) and undifferentiated carcinomas (8%) [2].

CRC aetiology is multifactorial, encompassing both somatic and hereditary forms. Risk factors for CRC include age, smoking, diet, obesity, and physical inactivity, which significantly increases susceptibility [3]. 10-16% of cases are caused by hereditary conditions such as Lynch syndrome, Hereditary mixed polyposis syndrome (HMPS), familial adenomatous polyposis (FAP), and MUTYH-associated polyposis [4]. However, the majority of CRC cases emerge sporadically through oncogenic mutations in intestinal epithelial cells, precipitating hyperproliferation, polyp formation, and ultimately tumour development.

A family history of CRC can also increase the likelihood of development, with 10% of all patients having a close family member, with varying risk depending on relatedness of individual and number of relatives affected [5]. However, it has been proposed that this number may be much higher, as twin and family studies on CRC have projected up to 40% of sporadic cases carry some genetic risk [6]. It has been shown that genetic risks may be caused by single nucleotide polymorphisms (SNPs), which are single base mutations in the genome. This project will focus on two genes, *POLD3* and *CHRD2*, that have been highlighted through genome-wide association studies (GWAS) as containing a high number of risk SNPs, and therefore may increase the risk for the development of CRC [148]. These genes lie next to each other on chromosome 11, so risk SNPs in this area may affect the expression of either gene. Therefore, we wished to study these genes together, to elucidate any potential effects either gene may play separately or in combination, in CRC development.

Colorectal cancer biology

The large intestine, where CRC occurs, encompasses the cecum, colon and the rectum. CRC is a heterogeneous disease which can arise from numerous mutations and varying pathways of genetic instability. The pathogenesis of CRC can also vary according to the anatomical location of the tumours and can differ between left sided (distal colon) and right sided (proximal colon which includes transverse colon, ascending colon, and the cecum). Differences in these tumours can be attributed to cause of malignancy, as well as carcinogenic factors including differing bacterial populations and exposure to bile [7]. Left and right sided tumours exhibit different histology, with right-sided tumours showing flatter, sessile serrated adenomas or mucinous adenocarcinomas, and left-sided tumours typically exhibiting polypoid, tubular, and villous adenocarcinomas [8][9]. These morphological differences result in left-sided tumours being more easily detected, whereas right sided tumours are harder to detect and tend to be larger, more advanced, and poorly differentiated.

Clinical presentation and diagnosis

In clinical cases, presentation of CRC symptoms includes abdominal pain, changes in type and frequency of bowel movements, rectal bleeding, involuntary weight loss, anaemia, nausea, vomiting, malaise, anorexia, and abdominal distention [10]. Changes in bowel movements are more common symptoms for left sided cancers caused by a progressive narrowing of the bowel lumen, with diarrhoea, a change in stool form, and eventually intestinal obstruction. Patients who have iron deficiency anaemia have a 10% likelihood of CRC, which usually presents as right sided [10].

Diagnosis

Diagnosis for CRC usually begins when a patient exhibits symptoms. Initial tests to look to the presence of blood in the stool is performed, known as a Faecal Immunochemical Test (FIT), with assessment of symptoms presented. Blood tests may also be performed to assess functions of the kidneys, and low levels of red blood cells which may indicate colon bleeding. After initial testing, a colonoscopy will be performed in which a long, flexible and slender tube is inserted into the rectum to look for the formation of polyps, and biopsies may be taken of abnormal areas. Pathological screening will then be performed of biopsies to confirm diagnosis. This will involve identification of molecular characteristics of the biopsy, identification of tumour cells, and the determination of cancer type. . Molecular profiling will

then be performed to assess the molecular subtype [11]. Presence of metastases is also examined for, and is often determined by a CT scan, PET or MRI of the thoracic cavity.

Screening for CRC has also been introduced to improve diagnosis at earlier stage of the disease and is recommended for adults from 50-75 years. Screening is performed by colonoscopy (every 10 years), high sensitivity guaiac faecal occult blood test (FOBT_ or FIT every year.

Treatments

Treatment for colorectal cancer relies of the initial profiling of cancer stage. The tumour-node-metastasis system is first used to identify the depth of tumour invasion in the bowel wall, extent of regional lymph node involvement, and presence of distant sites of disease (metastasis) .

TNM staging can be identified as follows:

T _{is}	Carcinoma in situ
T ₁	Tumour invades submucosa
T ₂	Tumour invades muscularis propria
T ₃	Tumour invades through the muscularis propria into the subserosa
T ₄	Tumour directly invades other organs or structures, or perforates visceral peritoneum

Treatment will then be decided by assessing the health of the patient, and apparent staging of the tumour. In most cases, surgery excision will be performed. Surgery methods include endoscopic mucosal resection or endoscopic submucosal dissection. In instances of localised cancer, complete resection of the diseased tissues by partial colectomy can be performed. Adjuvant chemotherapy may also be administered in instances of TNM stage 2 and above. Chemotherapy may also be administered as neoadjuvant, in which it is given before surgery in order to shrink the tumour size, enabling an easier resection.

Within the UK, chemotherapy for the treatment of CRC relies on a FOLFOX or FOLFIRI regime. 5-flourouracil is used in combination with Oxaliplatin (FOLFOX) or Irinotecan (FOLFIRI). Radiotherapy may also be used in cases of rectal cancer; however, it is not used in the colon due to the high sensitivity of radiotherapy in the bowel [13].

Immunotherapy may also be used in cases of CRC with mis-match-repair deficiency and POLE and POLD1 proof-reading deficient cancers. Immunotherapy utilises immune checkpoint inhibitors in order to limit cell growth. Common immunotherapies include Dostarlimab, Nivolumab, and Pembrolizumab, which target the D-1/PD-L1 pathway for mismatch-repair deficient tumours, as well as Ipilimumab which targets the CTLA-4 pathway in colorectal cancer that has high microsatellite instability. Immune checkpoint inhibitors have also recently been found to produce positive responses in patients with POLE/POLD1 proofreading mutations, which are characterised as a rare group of ultra mutated cancers. These patients were treated with PD_1 and CTLA-4 and had a significant increase in survival [13]. Other immunotherapies, such as Bevacizumab and Ramucirumab, target the VEGF/VEGFR pathway to inhibit tumour blood vessel growth. In cases of EGFR positive cancers, Cetuximab or Panitumumab may also be used [14].

However, mutated KRAS lowers the efficacy of these treatments, due to reduced immune infiltration [15]. Furthermore, the use of therapies that target KRAS mutations has been explored, but they lack efficacy due to high resistance exhibited. For instance, drugs Sotorasib and Adagrasib, which target the GTPase binding socket of G12C KRAS mutants has been approved for use, but have shown high levels of resistance, making KRAS mutations “undruggable” [16].

Colorectal cancer pathophysiology

Colorectal cancer typically develops in a step wise manner known as the adenoma-carcinoma sequence [39]. CRC can result from a combination of initial mutations resulting from Chromosomal Instability (CIN) or Microsatellite Instability (MSI) which activates oncogenic pathways such as APC and WNT, transforming growth factor beta (TGF- β), and Epidermal growth factor receptor (EGFR), which drive malignant transformation in epithelial stem cells [40]. Other mutations may also be present such as p53 which occurs in 70% of cases and causes resistance to apoptosis. These early mutations lead to the development of colorectal adenomas or polyps, which progress to carcinoma through mutations in genes such as *KRAS* (occurs in 50% of cases and regulates cell proliferation), *PIK3CA* and *SMAD4* [39][41][42] (Diagram 1). In many CRC cases up to 80 mutated genes can be found, however in this model it is believed that >15 genes are true drivers of tumorigenesis [42]. However, this model is yet to be fully elucidated and recent genome wide association studies have revealed over 100 common genetic variants that can increase CRC risk, with 53 new risk loci associated across

all ethnicities. This indicates that multiple genes may initiate and increase the risk to developing CRC, which are not yet fully understood [43].

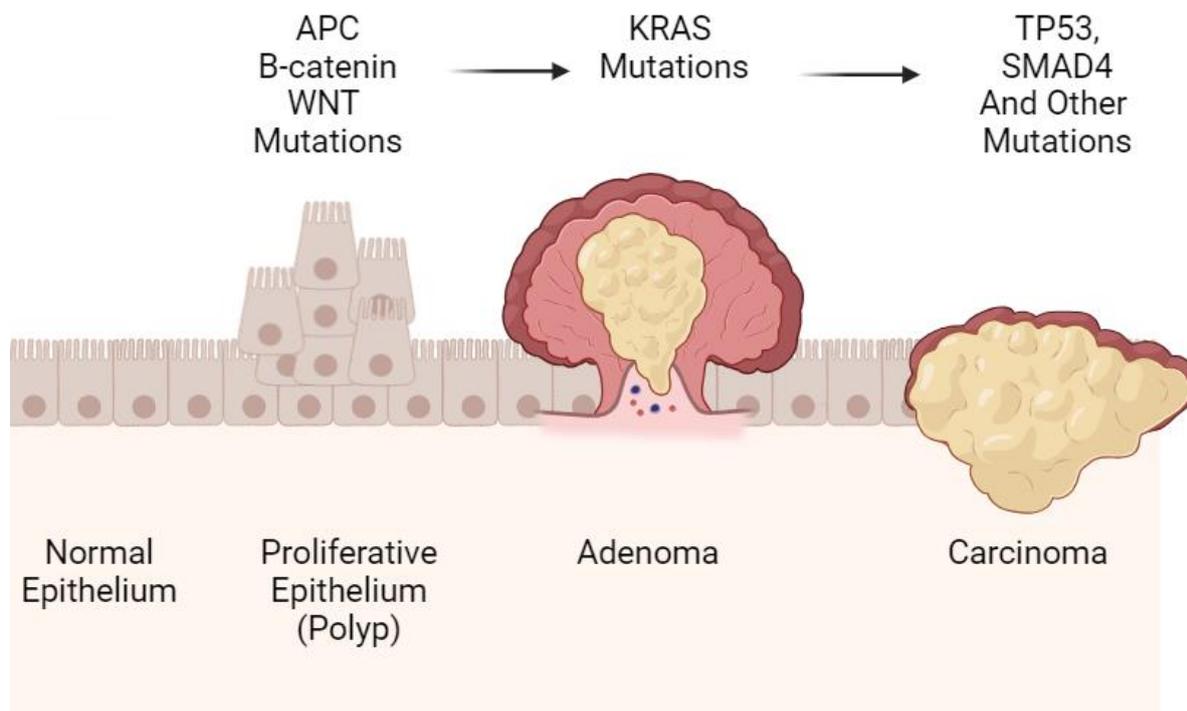


Diagram 1: Conventional pathway of Adenoma-carcinoma sequence in colorectal cancer. The adenoma-carcinoma sequence describes the progressive transformation of normal colonic epithelium into malignant carcinoma through a series of histological and genetic changes. This process typically involves transformation of the normal epithelium where tissue starts in a healthy, non-neoplastic state. Early mutations in *APC*, *Beta-catenin* and *WNT* cause hyperproliferation of epithelial cells to form a polyp. Further mutations in genes such as *KRAS* lead to the formation of benign adenomatous polyps, characterized by glandular proliferation. Further genetic alterations promote increased dysplasia, larger size, and potential villous features, with a higher risk of progression. Accumulation of additional mutations such as TP53 results in invasive carcinoma, characterized by malignant epithelial cells capable of invading surrounding tissues and metastasizing.

Consensus of molecular subtypes (CMS)

CRC tumours can be classified by their genomic profile. An international consortium grouped CRCs through large scale data sharing and analytics across multiple groups, leading to the development of consensus of molecular subtypes (CMS), which classifies four molecular subtypes with distinct features [44]. CMS1 (14%) is described by microsatellite instability (MSI) and strong upregulation of immune genes [45]. CMS2 (37%) is categorised as canonical and follows the adenoma-carcinoma sequence. CMS2 tumours have an over expression of epidermal growth factor receptor (EGFR) and commonly have mutations in APC, p53, and KRAS genes. CMS2 tumours are epithelial, with marked WNT and MYC signalling activation. CMS3 (13%) is described as epithelial with evident metabolic dysregulation and higher activity in glutaminolysis and lipogenesis [44]. CMS4 (23%), is mesenchymal with prominent transforming growth factor- β activation, stromal invasion, and angiogenesis [44]. This classification also confirmed previous findings that right-sided CRC patients tend to exhibit MSI-high tumours, left-sided CRC patients tend to have chromosomal instability-high (CIN-high) tumours [46].

Genetic and epigenetic pathways of colorectal cancer

Common CRC mutations develop through multiple different pathways of genetic and epigenetic instability. These well studied pathways are characterised by their molecular characteristics. These include, mutations in the Mismatch-repair pathway which causes Microsatellite instability (MSI), the CpG island methylator phenotype (CIMP) pathway which is characterised by global genome hypermethylation and gene silencing, and mutations in and overexpression of the WNT signalling pathway [41]. Each of these pathways exhibit independent molecular phenotypes and pathological features, which dictate the mechanisms of carcinogenesis and tumour development. Elucidating ways in which these pathways of genetic instability arise and are modulated is key to uncovering and categorising the molecular basis of CRC.

Microsatellite instability / mismatch repair

The MSI pathway is a unique mechanism in tumour development, and characterises a form of genomic instability that is found in approximately 15% of sporadic cases [47]. MSI arises from deficiency in the DNA mismatch repair system (MMR), and can be defined by the presence of alternate sized repetitive DNA sequences or short tandem repeats (STRs)[48]. The MMR system comprises of multiple proteins which act as a proof-reading mechanism in DNA replication and identify and repair mismatched nucleotides that have evaded detection by DNA polymerases. In eukaryotic cells this consist of key MMR players MutSa and MutLa, as well as many other protein components such as PCNA and RPA [49]. Inactivity of the MMR system causes 100-fold increase in the mutation rate of colorectal mucosal cells [49]. MSI tumours consequently often contain many SNPs and alterations in these repetitive microsatellite sequences. MSI leads to a range of mutated genes which are present in 20% of sporadic cases and 80% of hereditary cases [50]. In sporadic cases, MSI+ CRCs are typically caused by epigenetic silencing of the *MLH1* promoter gene, which leads to inactivation of target tumour suppressor genes which contain microsatellites in their coding regions [51][52].

MSI+ tumours are less likely to be invasive, and less likely to have mutations in *KRAS* or *p53* compared with other tumours. MSI+ tumours are more likely to appear poorly differentiated and serrated, and come from younger patients [53]. Lynch syndrome occurs when there are germline defaults in the MMR pathways. Lynch syndrome is inherited in an autosomal dominated manner, and patients are born with 1 functional and 1 non-functional allele of the MMR gene. These tumours occur in younger patients, and can have *KRAS* mutations, but never *BRAF* mutations, and are associated with a better prognosis than non-MSI tumours. However somatic MMR deficiency occurs in older individuals and tends to have *BRAF* mutations in about half of the cases. These tumours are associated with a background of CIMP and are associated with a reduced mortality[53].

Tumours that are MSI+ are often located in the proximal colon, and more poorly differentiated compared to CIN+ CRC tumours. However, patients that are MSI+ tend to have a better prognosis and survival compared to their CIN+ counterparts [54].

CpG island methylator phenotype

The CIMP or “Serrated” pathway is the second pathway in which CRC may progress. It consists of epigenetic hypermethylation of CpG dinucleotide sequences within promoter regions of genes involved in cell cycle regulation, DNA repair, invasion, adhesion, angiogenesis, and apoptosis [47] [55]. These sequences of CpG dinucleotides are referred to as CpG islands, and it has been shown that 50-60% of genes contain these CpG islands in their 5' region, which are typically maintained in an unmethylated state [56]. Many cases of CRC show aberrant hypermethylation of CpG islands, which cause gene repression or loss of gene expression entirely [57]. For example, hyper methylation of the *MLH1* gene (encodes MutL protein) is found in >80% of sporadic MSI colorectal cancers, and demethylation of the *MLH1* gene restored MLH1 protein function [58]. Therefore, this pathway of CRC development often occurs in conjunction with the MSI pathway, and has many overlapping pathologies.

CRCs with the CIMP phenotype can be classified further into two groups: CIMP-high and CIMP-low. The CIMP-high phenotype is characterised by high methylation levels, and often identified with (>90%) the *BRAF* oncogene, which causes increased proliferation, increased carcinogenesis, and high mortality rates [47][59]. However, the CIMP-high phenotype is often associated with low mortality rates. CIMP-high phenotype is often also associated with older age and smoking, which may explain these discrepancies [59].

WNT pathway mutations and Chromosomal instability

Almost all CRC tumours show hyperactivity of WNT signalling, a downstream pathway from the canonical APC/beta-catenin pathway [60]. Mutations most commonly occur in stem-cells in the villi crypt of the intestine. A 2012 report from The Cancer Genome Atlas (TCGA) consortium showed that 92% of CRC cases had at least one WNT pathway mutation, and recent studies elucidating further regulators of the WNT pathway have brought this figure closer to 100% [61]. WNT is necessary for normal intestinal function and regulates self-renewal of intestinal stem-cells in the villi crypt [62]. WNT binds to extracellular Frizzled-G-protein coupled receptors, which trigger binding to the β -catenin destruction complex, preventing β -catenin from degradation. Thereby leading to an increase in β -catenin signalling within the nucleus (Diagram 2). β -catenin within the nucleus mediates T-cell factor (TCF) signalling which activates known stem-like and oncogenic pathways [61]. However,

aberrant WNT signalling can drive stem-like features in intestinal epithelial cells outside of the villi-crypt, leading to the formation of malignant tumours.

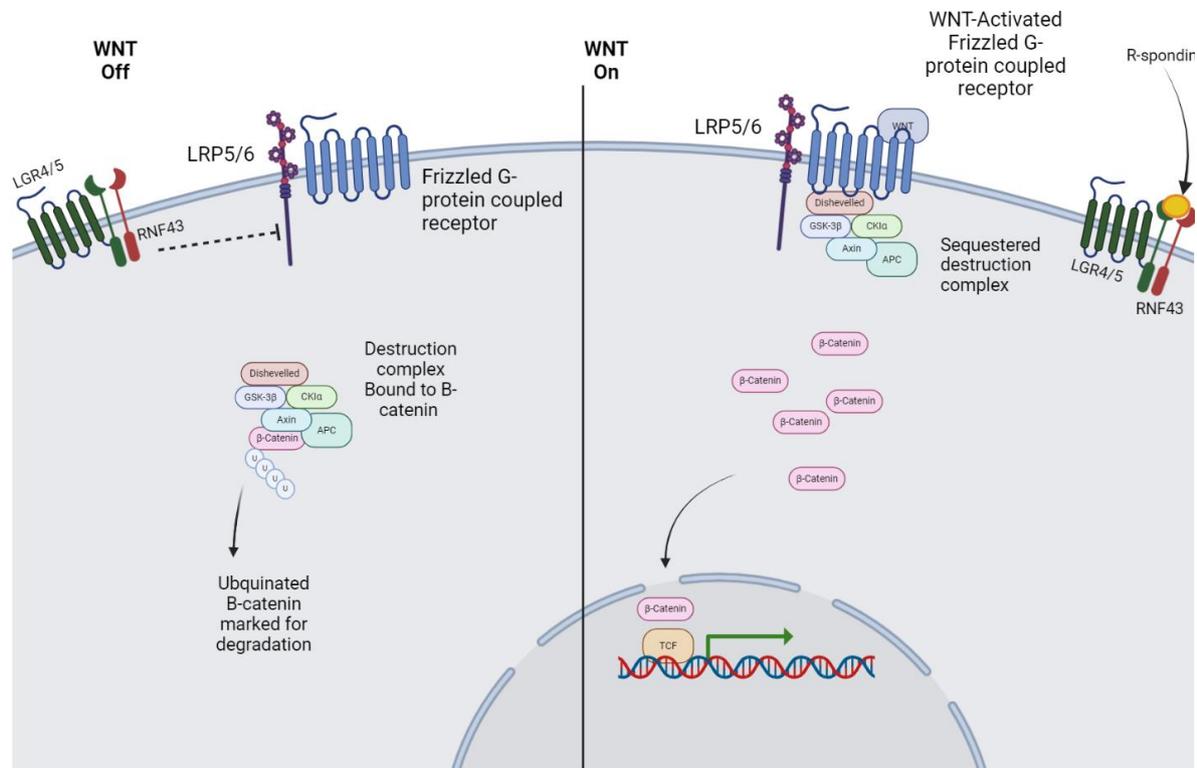


Diagram 2: WNT/APC/Beta-catenin pathway. WNT binds to G-protein coupled receptors to phosphorylate LRP. Phosphorylated LRP induces translocation of the destruction complex to the membrane receptors, activating Dishevelled, leading to the inhibition of the destruction complex. Inhibition of the destruction complex prevents ubiquitination of Beta-catenin, leading to an increase of cellular Beta-catenin levels. Increased Beta-catenin levels leads to increased Beta-catenin translocation into the nucleus, which binds to TCF and causes TCF mediated transcription of target genes.

WNT pathway activation is often “driven” by key mutations in the truncation of the *Adenomatous Polyposis Coli (APC)* tumour suppressor gene, which is found in approximately of 80% of sporadic cases [63]. The APC protein forms part of the Beta-catenin destruction complex as part of the WNT/ β -catenin/TCF pathway, with inactivating mutations in APC leading to increased WNT signalling.

Mutations in APC prevent the degradation of β -catenin by the destruction complex, leading to increased β -catenin translocation within the nucleus, where it disrupts Groucho-mediated inhibition of downstream TCF targets leading to increased proliferation, differentiation, and migration of colorectal cells [47]. In sporadic cases, mutations in β -catenin are also commonly found, with 48% of tumours without APC mutations containing β -catenin mutations [64]. Cells with APC mutations show extensive chromosome and spindle aberrations, which has been shown to cause CIN and further downstream mutations as a result [65]. CIN is the most well

characterised pathway in CRC and is defined by widespread imbalances in chromosome number (aneuploidy), loss of heterozygosity (LOH), and sub-chromosomal genomic amplifications [47]. It is highly debated whether these mutations in the WNT pathway cause CIN or vice versa. However, once CIN has been activated, this leads to fast accumulation of mutations in oncogenic genes.

Additional mutations that perturb this APC/ β -catenin/TCF pathway have been found in CRC cases. For example, cyclin dependent kinase-8 (CDK8) gene amplification has been found in approximately 70% of CRC cases [66]. CDK8 is a member of the mediator complex and is necessary for β -catenin driven transformation and the expression of several β -catenin transcriptional targets. Through this, CDK8 acts as an oncogene by stimulating β -Catenin as well as NOTCH1, leading to increased proliferation and migration [67]. Other mutations common in the APC/WNT pathway are Cyclin D1 (CCND1), which is stimulated by APC mutations and leads to the development of colonic neoplasia allowing the cell to escape apoptosis, and *catenin beta-1* (CTNNB1) which encode the β -catenin protein [68].

An additional layer of WNT regulation is found in the form of E3 ubiquitin ligases ring finger protein 43 (RNF43) and zinc and ring finger 3 (ZNRFF3). RNF43 and ZNRFF3 degrade frizzled to repress WNT signalling. However, R-spondin ligands can bind to leucine-rich repeat-containing G protein-coupled (LGR) receptors, inhibiting RNF43/ZNRFF3 and substantially amplifying WNT signalling (Diagram 2)[69]. Therefore, mutations in RNF43 and R-spondin can influence WNT expression but require exogenous WNT ligand to activate signalling (ligand-dependent). Mutations in the R-spondin gene *RSPO*, usually induce ligand overexpression from epithelial cells, or as *RSPO* fusion genes. R-spondin gain-of-function mutations are observed in *RSPO2* and *RSPO3* and enhance WNT signalling *in vivo*. Fusion genes in *RSPO3* result in the replacement of the promoter and exon 1 with that of a gene with higher basal expression, resulting in a functional epithelial-expressed protein [70]. *RNF43* mutations commonly occur as frameshift mutation and result in a truncated gene product [71]. Mutations tend to occur at microsatellite tandem repeats and are therefore commonly found in tumours with MMR deficiency [69].

After initial mutations in APC, the adenoma to carcinoma transition is often facilitated by mutations in the *KRAS* gene. *KRAS* is a proto-oncogene that encodes for the GTPase protein involved in the transduction and propagation of extracellular signals. Mutations of *KRAS* allow the cell to avoid apoptosis and increase proliferation, leading to a permanently active state [47]. Furthermore, *KRAS* mutated cells show WNT activity that is connected to increased MAPK signalling, leading to further increased growth and resistance to apoptosis [72].

Several other mutations commonly occur after mutations in APC. For example, *TP53* is a tumour suppressor gene encoding the P53 protein and is commonly found mutated in CRC [73]. P53 loss of function is frequently found in later stages of colorectal tumorigenesis. P53 is known as the “guardian of the genome” due to its role in protecting genome stability, regulation of the cell cycle, and regulation of apoptosis [47]. P53 directly regulates WAF1, which binds to G1-S/cyclin dependent kinase (CDK), which holds the cell cycle and will not permit cell cycle continuation in tumour cell growth. However, when P53 is mutated, the protective role of WAF1 is not expressed, and proliferation is uncontrolled through lack of cell cycle progression inhibitors [74]. P53 is also thought to regulate energy balance through modulation of the AMPK pathway [75]. AMPK activates catabolic pathways and regulates the glucose-dependent checkpoint at the G1-S boundary of proliferating cells. P53 loss of function and subsequent deregulation of the AMPK therefore allows proliferating tumour cells to bypass this checkpoint [75]. P53 also interacts with Cyclooxygenase-2 (COX-2), which promotes inflammation and proliferation in CRC [76].

PI3KCA mutations are also commonly found to occur simultaneously to mutations in the *APC* gene. The PI3KCA-AKT pathway triggers the activation of different nuclear transcriptional factors through a kinase cascade. PI3KCA mutations stimulate cell proliferation and production of fatty acid synthases via the AKT pathway [77]. PI3KCA also interreacts with mTOR and KRAS, which are central regulators of cell growth and metabolism. The PI3KCA-AKT pathway is also regulated by PTEN, with mutations or loss of expression found in 40% of CRC cases [10].

MYC is a family of regulator genes and proto-oncogenes which are encoded for by the *c-MYC* gene, as well as *I-MYC* and *n-MYC*. MYC has been implicated in many forms of cancer, and acts as a transcription factor to regulate many cellular processes, such as proliferation, cell cycle progression, and apoptosis [78]. Activated MYC is found in many types of cancer, however constitutively overexpressed in normal cell lines does not exert any effects on its own, and requires other genetic abnormalities, such as the activation of CIN [78]. In CRC, MYC expression is often deregulated, and it has been shown that aberrant WNT/ β -catenin signalling induced overexpression of *c-MYC*, which in turn enhanced cell proliferation, and inhibited differentiation [79]. MYC overexpression may also induce WNT signalling, as MYC target genes includes *LEF1*, a co-transcription factor of β -catenin. MYC transcription of *LEF1* signalling may therefore act in a positive feedback loop to sustain WNT [80].

Following WNT activation, several downstream processes are affected. For example, genes involved in the proliferation, migration, and survival of cancer cells are upregulated upon WNT

activation. A list of WNT target genes that can be affected by dysregulation of WNT signalling can be found below:

Biological function	Gene	Full name	Direct/indirect target
Initiation	LGR5	LEUCINE-RICHH REPEAT-CONTAINING G-COUPLED RECEPTOR 6	DIRECT
	DCLK1	DOUBLECORTIN-LIKE KINASE	DIRECT
	KLF5	KRUEPPEL-LIKE FACTOR 5	DIRECT
	EDN1	ENDOTHELIN 1	DIRECT
	ASCL2	ACHARTE-SCUTE HOMOLG 2	DIRECT
Persistence	FRA1	FOX-RELATED ANTIGEN 1	DIRECT
	MYC	MYC PRTO-ONCOGENE PROTEIN	DIRECT
	CCND1	CYCLIN D1	DIRECT
	TCF1	T-CELL FACTOR 1	DIRECT
	ABCB1	ABC MULTIDRUG TRASNPORTER	DIRECT
	MMP7	MATRIX METALLOPROTEASE 7	DIRECT
	LEF1	LYMPHOID ENHANCING FCATOR 1	DIRECT
Migration	HAS2	HYALURONAN SYNTHASE-2	DIRECT
	CD44	CLUSTER OF DIFFERNTIATION 44	INDIRECT
	CXCL12	C-X-C MOTIF CHEMOKINE LOGAND 12	DIRECT
	CXCR4	CHEMOKINE REPCTOR TYPE 4	DIRECT
	CXCL12	C-X-C MOTIF CHEMOKINE LIGAND 12	DIRECT
	CXCR4	CHEMOKINE REPCTOR TYPE 4	DIRECT
	DKK1	DICKKOPF-RELATED RPTOEIN 1	DIRECT
Invasion	CLDN1	CLAUDIN-1	DIRECT
	CD44v6	CLUSTER OF DIFFERENTIATED 44 VARIANT EXON 6	INDIRECT
	FN1	FIBRONECTIN	DIRECT
	COX2	CYCLOOXYGENASE-2	DIRECT

Other common mutations in CRC include in genes *ARID1A*, *DCC*, *FAM123B*, *FBXW7*, *SMAD4*, *TGFBR2*, *BRAF*, and *SOX9* [10].

The WNT/BMP signalling axis

It is well-known that the majority of sporadic CRC cases occur from mutations within the WNT signalling pathway. WNT signalling, along with BMP signalling, is the main driving force behind the classic intestinal architecture and is crucial for maintaining intestinal homeostasis. The intestinal tract epithelium is comprised of crypts and villi, with the small intestine characterized by villi protruding into the lumen and crypts invaginating into the mucosa. The large intestine, encompassing the caecum and colon, shares this architecture but lacks prominent villi. Epithelial cells line the intestinal tube and originate from stem cells situated in mitotic zones at the crypt base. These stem cells facilitate the continuous renewal of epithelial cells, which are crucial for maintaining intestinal epithelial integrity. WNT signalling, which helps maintain intestinal stem-cells, works in a counter gradient to pro-differentiation BMP signalling, which is found in the villi of the intestinal epithelium (Diagram 3).

The WNT signalling pathway is central to the creation and maintenance of stem -cell crypts. Stem cells within the crypt respond and are maintained by WNT signalling, and are capable of self-renewal, passively competing within their niche to provide a steady production of epithelial cells to maintain the intestinal epithelial lining [81]. Stem cells in the crypt are identified by the presence of the LGR5+ marker, which is a downstream WNT target. These LGR5+ cells are actively cycling to be capable of producing all lineages of epithelial cells [82]. Cells generated from LGR5+ cells exit the crypt through passive displacement and migrate along the villi-crypt axis towards the villi-tip in a conveyor belt-like movement. As epithelial cells migrate along the villi-crypt axis, they are exposed to BMP signalling which promotes differentiation into mature cell lineages before shedding at the villi-tip [81][83].

Oncogenic hyperactivity of the WNT signalling pathway causes aberrant uncontrolled proliferation of epithelial cells, leading to the development of CRC [84]. Mutations in the WNT signalling pathway can be inherited or acquired, and most likely occur within the intestinal stem cell crypt [85]. Epithelial cells with hyperactivity in the WNT signalling pathway, such as mutated APC, exhibit stem-like properties such as hyper-proliferation, increased migration, and resistance to common chemotherapies [86].

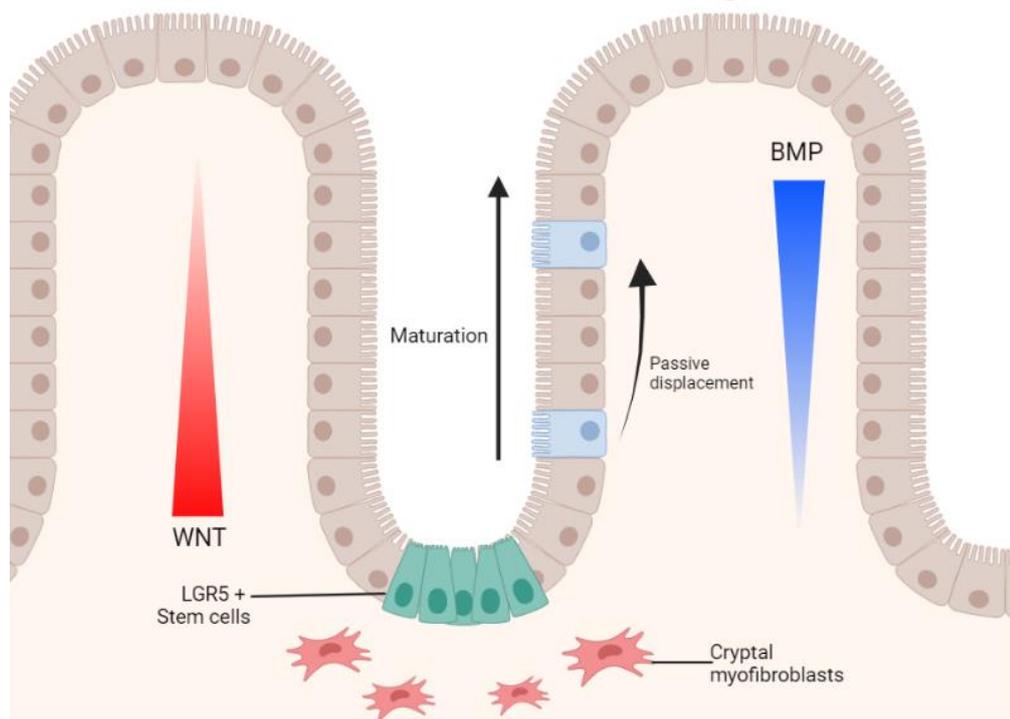


Diagram 3: WNT/BMP signalling axis in the intestinal epithelium. WNT signalling is localised in the intestinal crypt, and produced by Paneth cells maintain the pool of LGR5+ stem cells. WNT expression is reduced in epithelial cells as they exit the crypt in a gradient pattern. BMP signalling expression is strongly localised at villus-tips and is gradually reduced down the villus-crypt axis where it is blocked by BMP antagonists produced by cryptal myofibroblasts localised in the crypt base. As epithelial cells exit the crypt through passive displacement, they mature in response to increasing BMP signalling before shedding the villus tip.

WNT signalling in the intestinal epithelium is therefore heterogeneous, not acting as an “on/off” switch, but as a highly regulated pathway, with spatial and temporal variation effecting signalling intensity, which depends on a multitude of factors [87]. This was first demonstrated in 1998, where nuclear staining of Beta-catenin was found to be heterogeneous in colorectal cancers, despite all tumour cells harbouring *APC* mutations. Variations in WNT signalling can occur from external and internal factors. For example, macrophage-derived TNF- α and *H. pylori* infection can both activate the WNT pathway via frizzled, highlighting the role for the tumour microenvironment in WNT activation [88,89]. Furthermore, the level of *APC* mutation can affect the severity of WNT signalling. For example, deleted *APC* results in a more severe phenotype, but truncated *APC* mutants are still transcribed and are able to partially function. Additionally, the *APC* promoter may be methylated, effectively downregulating *APC* expression [87]. As mentioned previously, *RNF43* mutations may also occur in *APC* mutant cells, occurring in approximately 5.5% on colonic cases, compounding the already activated WNT signalling pathway [90]. On the other hand, gastric cancers have *RNF43* mutations in

approximately 37% of cases, suggesting gastric and colorectal cancers differ in their required WNT activation.

This developed the “just right” theory, in which *APC* mutations are selected for sub-maximal levels of WNT signalling to provide a WNT signal that is sufficient to transform cells, but not excessive and cytotoxic. This theory stems from the “two-hit” hypothesis by Knudson in 1971, in which tumour suppressor genes require both alleles to be inactivated, through either mutation or epigenetic silencing. However, it has been shown that *APC* does not strictly follow this model, as the type and position of the second hit in the *APC* gene depends on the initial hit [91]. This results in a selection bias for *APC* genotypes that retain some level of activity in downregulating β -catenin for optimal tumour formation.

This has led to the theory of cancer cells harbouring multiple mutations and epigenetic modifications in the WNT pathway for optimal expression. This appears to constrain WNT signalling within an optimal “just right” spectrum, where WNT signalling is required in a context dependent manner, such as at the tumour front [92,93]. This is not just limited to *APC* mutations but includes intracellular and extracellular members of the WNT signalling pathway which modulate its expression. For example, WNT signalling can be further stimulated by inhibiting GSK3 or addition of WNT3A in *APC*-mutant colon cancer cells [93][92].

In contrast to WNT signalling, BMP signalling is the main differentiation pathway of epithelial cells in the villi, and loss of BMP signalling is often part of oncogenic activation in CRC [94]. This connects to the “just right” theory, in which loss of WNT-suppressive pro-differentiation BMP signalling may lead to oncogenic activation and enhancement of WNT signalling [95] [96].

In order to differentiate stem epithelial cells into their mature lineages, BMP signalling works in a counter gradient to WNT signalling and is present outside of the crypt along the villi. Epithelial cells such as absorptive enterocytes and mucous-secreting goblet cells exit the crypt and mature in response to BMP signalling, and migrate toward the villi-tip, where they differentially express zoned genes in response to BMP signalling [97]. It is this balance between differentiative BMP signalling and stem WNT signalling which maintains homeostasis. However, disruption of the balanced gradient may therefore lead to oncogenic transformation of epithelial cells and create the “just right” WNT signalling activation for tumour initiation. .

BMP signalling in colorectal cancer

BMPs are a group of cytokines and metabologens that are widely recognised for their important role in morphogenic signalling and are responsible for orchestrating multiple tissue architectures throughout the body. Originally identified through their ability to initiate and direct bone and bone-cartilage formation [98], they have subsequently been shown to influence a wide-range of biological processes, including cancer progression.

BMP signalling in cancer appears to be paradoxical, as BMP signalling has been shown to be both oncogenic and tumour suppressive in nature [99][100][101][102]. BMPs belong to the TGF-B superfamily and bind to a complex of transmembrane serine threonine kinase receptors I and II (BMPRs I and II) [103]. This initiates phosphorylation of the type I receptor by the type II receptors, triggering phosphorylation of a receptor-associated SMAD that subsequently complexes with SMAD4, resulting in translocation to the nucleus to regulate gene transcription [104] (Diagram 4). While epithelial cells and mesenchymal cells express BMPs and their receptors, BMP antagonists are primarily found in the mesenchyme and are expressed by intestinal cryptal myofibroblasts and smooth muscle cells. These antagonists block BMP signalling in the stem-cell compartment, allowing WNT signalling to maintain stem progenitors [105].

Mutations in the BMP signalling pathway have been well established as risk factors for hereditary CRC. For example, *SMADH4*, which encodes the SMAD4 protein that transduces the BMP signal, often has copy number deletions in ~30% of CRC samples [106]. Other reports are that approximately 5-24% of colorectal cancer have inactivating mutations in SMAD4, collectively making it one of the most frequently mutated genes in CRC. *BMPR1A* is also frequently lost in juvenile polyposis syndrome, with 17-38% of cases having disease causing variants within the gene [107].

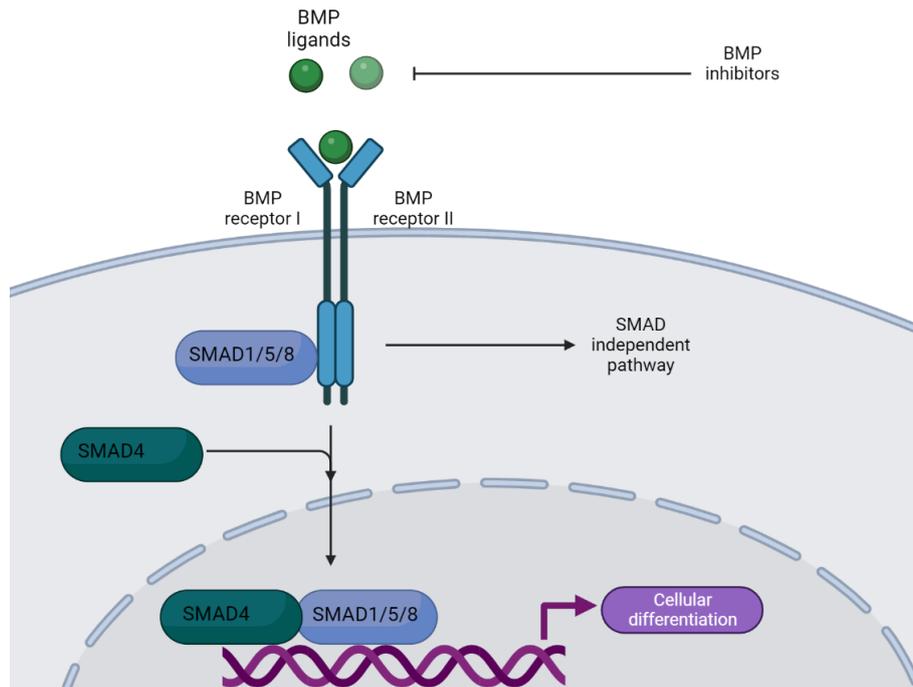


Diagram 4: BMP signalling through SMAD 1/5/8 pathway. Extracellular BMP ligands bind the BMP receptors I and II which dimerise to recruit and phosphorylate receptor-associated intracellular SMAD1/5/8. SMAD1/5/8 subsequently complexes with SMAD4, resulting in translocation to the nucleus to regulate gene transcription and cellular differentiation.

WNT and BMP signalling and Cancer stem cells

The WNT and BMP signalling pathways have been shown to influence populations of Cancer Stem cells (CSCs) in colorectal cancer. CSCs are described as a population of self-renewing malignant and highly tumorigenic cells that contain high levels of WNT signalling and drive tumour initiation and progression. In recent years our view of the tumour landscape has changed from that of a homogenous pool of dividing cells, to a more complex hierarchy, with CSCs at the top of this hierarchy, giving rise to lineages of cancer cells that complex to form the tumour [108]. CSCs are often termed “immortal” due to their ability to self-renew and their resistance to conventional chemo and radiotherapies. This makes populations of CSCs able to drive tumour initiation from rapid self-renewal, but also increase the risk for cancer recurrence by evading therapies [109]. CSCs have enhanced DNA repair mechanisms through enhanced DDR engagement, activation of cell cycle checkpoints and longer residence in G₀. CSCs are able to use this enhanced DNA repair ability, along with the ability to maintain and extend telomeres, to avoid senescence and differentiation-imposed cell death to become immortal, which is not found in differentiated cancer cells. The cell of origin of cancer stem cells is controversial, with some speculating that cancer cells adapt gain-of function mutations to

de-differentiate into a stem-like state, and others theorising that stem-cells gain oncogenic proteins to give rise to cancer stem cells which drive tumour formation [292] [293][294].

In the intestinal tract, BMPs have been shown to promote differentiation, apoptosis, and chemo-sensitization, and are therefore an opposing factor to cancer stem cells. CSCs in the intestinal tract have been shown to lack BMP expression compared to differentiated cancer cells, and addition of BMP4 induces CSC differentiation and chemo sensitization [110]. Furthermore, BMP4 blocks transplantation ability of CSCs, and inhibits tumour growth in the absence of SMAD4 loss [111].

A screen for genes that were required for maintaining the tumour stem cell phenotype identified the zinc-finger transcription factor GATA6 as a key regulator of the WNT and BMP pathways in CRC. It was found that GATA6 directly drove the expression of LGR5 in adenoma stem cells but restricted BMP signalling to differentiated tumour cells. GATA6 deletion in mouse colon adenomas increases the levels of BMP factors, thus blocking the self-renewal capacity of tumour stem cells [112].

BMP antagonists

It is clear that loss of BMP signalling and hyperactivity of WNT signalling is imperative to cancer progression. While loss of key BMP transducers is well studied, another modulator of BMP activity is found in the form of BMP antagonists. These antagonists are expressed by mesenchymal cells in the stroma of the crypt, which block BMP signalling to maintain WNT signalling. BMP antagonists can be categorized into 3 classes: ligand antagonists which bind directly to BMPs, BMP pro-regions which complex back with mature BMPs, and receptor antagonists, which bind to BMPRs to prevent BMPs from binding to their cognate cell surface receptors [113]. Similarly to BMPs, BMP antagonists contain cysteine knots and typically form homo or hetero dimers. Some well-studied BMP antagonists include Noggin, the Gremlins (GREM1 and 2), and the Chordin family of proteins, including Chordin, Chordin-like 1 (CHRDL1) and Chordin-like 2 (CHRDL2) [114][115][116][117].

Noggin

Noggin, also known as Nog, and encoded by the *NOG* gene, is a BMP antagonist that is highly homologous between human, Rat and Mouse, and therefore has earned its place as a potent BMP antagonist that is widely used in cell culture to stimulate BMP inhibition. Noggin was first discovered in *Xenopus* due to its ability to restore normal dorsal-ventral body axis patterning in embryos [118], and since gained significance in its ability to form germ layer-specific derivation of specialized cells. Noggin activity is crucial for the formation of early neural tissues, including the notochord and eye structures. The activity of Noggin in the mesoderm aides the formation of cartilage, bone and muscle growth, and in the endoderm, noggin is involved in the development of the lungs [119].

The function of noggin as a BMP inhibitor has been well established and has vast implications in the context of cancer progression. Noggin is important in maintaining intestinal stem-cells, with Noggin expression relied on as a “second signal” in co-occurrence with WNT signalling to release PTEN inactivation of Beta-catenin and therefore coordinate WNT signalling to activated stem cells [95]. Treatment with Noggin increases the levels of phospho-PTEN and P-Akt which are known proto-oncogenic pathways [95]. During Noggin overexpression, abnormal branching and budding of the intestinal epithelium can be observed, with crypt dilatation and reactive inflammatory changes resulting in the epithelium [120].

Due to its inhibitory nature of BMP signalling, Noggin has been implicated in a wide variety of cancers, including breast, gastric, colorectal and skin tumorigenesis. Within the context of skin cancer, BMP inhibition by Noggin in mice showed a marked increase in proto-oncogenic signalling, with genes encoding adhesion/extracellular matrix molecules, cell cycle/apoptosis, and cytoskeleton/cell motility markers, and molecules involved in the control of cell differentiation, metabolism, signalling, and transcription shown to be differentially expressed by Noggin overexpression [121]. Furthermore, Noggin was able to restrict the anti-tumour effects of BMP signalling, as shown by a decrease in P-SMAD1/5 expression, resulting in activation of oncogenic pathways WNT and hedgehog signalling, and subsequently an increase in tumour staging. Noggin has also widely been shown to have negative implications in gastric cancer, and overexpression of the gene is associated with poor prognosis. Noggin overexpression in gastric cancer cell lines promoted cell-cycle progression, proliferation, invasion, and colony formation, whereas knockdown produced the reverse effects [122]. Furthermore, Noggin was shown to exert its effects on overexpression of the well-known oncogene EGFR via nuclear localisation of Beta-catenin.

Within colorectal cancer, Noggin was not found to be an independent prognostic factor, however, is concurrently expressed with WNT activation in tumour tissues, and its effect as a BMP inhibitor promotes CRC pathologies. Low-noggin expression in patients has been correlated with increased survival [123]. The BMP target gene, ID1, was downregulated by Noggin overexpression, resulting in down-regulation of tumour-suppressive gene signatures by lack of BMP signalling [123].

BMP inhibition by Noggin in mice transgenic lines restricted to the intestinal tract resulted in the phenotypic similarities to that of Juvenile-polyposis syndrome, which is characterised by BMP dysregulation. Noggin overexpression in these lines resulted in the formation of numerous ectopic intestinal crypts perpendicular to the crypt-villus axis. Severe abnormalities of the intestinal epithelial were observed during early development, with villi showing blunt tips and abnormal invaginations perpendicular to the villus that contained proliferative ki-67 cells and expressed WNT signalling targets shown by increased Beta-catenin staining. Mice then developed similar intestinal epithelial abnormalities to that of patients with Juvenile-polyposis, with polyps developing and foci of dysplastic epithelium and adenomatous change, in which the cells were abnormally arranged with an increased nuclear cytoplasmic ratio and pseudo stratification of the cells [124].

GREM1

GREM1, previously known as Dnm, is a 20 kDa highly conserved glycoprotein that acts as a BMP antagonist. GREM1 is part of the DAN family of proteins, which is a subgroup of the CAN (Cerberus and dan) family. GREM1 is a secreted glycosylated protein and contains a C-terminal cystine knot with an eight-membered ring [125][126]. GREM1 was originally identified as a pathogenic mediator of diabetic nephropathy (DN) [127], and has a variety of functions as a BMP antagonist, including the regulation of limb bud formation [128]. GREM1 has subsequently been implicated in a variety of disorders, including idiopathic pulmonary fibrosis, pulmonary artery hypertension, osteoarthritis, chronic pancreatitis, and cancer [129].

In CRC, RNAseq analysis has revealed that high GREM1 expression leads to poor patient outcome, and that GREM1 neutralising antibodies were able to shrink tumour size and promote LGR5+ intestinal stem cell differentiation, confirming that GREM1 mediates promotion of WNT stem cell signalling [130]. Furthermore, BMP antagonism through GREM1 has been shown to cause regenerative stem-cell activation, suggesting that BMP antagonism is able to functionally switch cells into a de-differentiated state [131]. Overexpression of GREM1 was also shown to enhance the motility and invasion of CRC cells by epithelial-

mesenchymal transition (EMT), as well as upregulate activating transcription factor 6 (ATF6) and downregulated ATF4, modulating the unfolded protein response (UPR) through activation of PI3K/AKT/mTOR and antagonization of BMP2 signalling pathways [132]. The UPR response in CRC has been recently gaining attention, with research showing that activation of the UPR enables cancer cells to evade apoptosis under ER stress, such as hypoxia, ATP shortage, and nutrient deficiency [133]. When using UPR inhibitors, such as 4-PBA (4-Phenylbutyric acid) or TUD-CA (Tauroursodeoxycholic acid), cancer progression and metastasis are significantly decelerated [134]. There is also evidence that downregulation of ATF4 induces apoptosis upon ER stress through activation of C/EBP homologous protein (CHOP), therefore limited cancer cell death under ER stress [135].

One of the most significant roles of GREM1 in disease progression has been shown in hereditary mixed polyposis syndrome (HMPS) in CRC. A 2012 study found that HMPS was caused by a germline gene duplication spanning the 3' end of the SCG5 gene and a region upstream of the GREM1 locus. Furthermore, this mutation caused overexpression of GREM1 mRNA expression and loss of BMP signalling [136]. This was duplicated by a study in 2016 that found a duplication in the GREM1 gene in a family with attenuated/atypical polyposis syndrome [137]. A further study in 2014 [138] found that HMPS polyps were characterised by ectopic crypt formation, that developed orthogonally to the crypt axis and contained actively proliferating cells. This same laboratory then used a transgenic mouse line to create a VIL1/GREM1 mouse line that develop ectopic crypts in a similar fashion to HMPS patients. These Ectopic crypts budded off to become actively proliferating intra-villus lesions, which quickly developed into pan-intestinal polyposis, with mixed serrated, adenomatous, and cystic phenotypes such as those seen in HMPS patients. In this Vil1/GREM1 mouse model BMP signalling was lost, evidenced through reduction in p-Smad1/5/8 staining in the vertical axis on the intestinal villi, suggesting that GREM1 overexpression abrogated the BMP signalling gradient which is characteristic of healthy luminal villi. Furthermore, intestines of transgenic mice were found to be 28% longer than wild-type counterparts, owing to increased villi formation, and enhanced stem-cell progenitor pools.

Chordin Like 2

Chordin-like-2 (CHRDL2) is a BMP ligand antagonist and binds to BMPs to prevent BMP from interacting with its cognate cell surface receptors [96][141]. The CHRDL2 gene is 34kbp and encodes a 47kDa protein with a repeated cysteine-rich motif, known as Von Willebrand factor C (VWC) and a signal peptide. VWC domains are a signature of several extracellular matrix proteins within the TGF β family, such as thrombospondin-1, thrombospondin-2, Von Willebrand Factor-1, and pro-collagen type 1 [142]. Further studies found that the VWC domain mediates binding to BMP2, 4 and 5, as well as the Tsg (twisted gastrulation) protein, allowing the formation of a tertiary complex consisting of BMPs, CHRDL2, and Tsg [143]. This tertiary complex allows increased binding affinity of CHRDL2 to BMPs, and therefore may increase WNT signalling through BMP inhibition.

CHRDL2 has been shown to be upregulated in CRC tumour tissues, and overexpression has been shown to predict poor prognosis in CRC patients [144]. Furthermore, high CHRDL2 levels correlated with clinical features of CRC patients, including tumour size, TNM staging, and tumour differentiation. Genetically predicted CHRDL2 protein levels were also shown to correlate with increased CRC risk [145]

DNA polymerases and Colorectal cancer

Colorectal cancer may also arise in response to defects in the DNA replication and repair systems. DNA polymerases are essential to living organisms; their function as enzymes allows replication of genomic and mitochondrial DNA for the purpose of cell division. Furthermore, DNA polymerases are also involved in the proofreading, maintenance and repair of DNA, as well as the maintenance of telomeres at chromatin ends. DNA polymerases work with remarkably high fidelity with few errors; the replication machinery providing a robust system where errors are promptly identified and resolved. However, mutations or epigenetic modifications in the DNA replication, surveillance, and repair machinery can cause failure of these systems, resulting in the manifestation of oncogenic mutations and DNA damage [149].

DNA polymerases (Pol) consist of five main distinct types, α , β , γ , δ , and ϵ . α , δ , and ϵ belong to the B family of polymerases, whereas β are part of the X family, and γ the A family. A smaller set of lesser known polymerases belong to the Y family. Pol α is responsible for initiation of replication at DNA replication origin sites, and synthesis of Okazaki fragments of the lagging strand (Diagram 5)[149]. DNA Pol β is found in non-dividing cells, suggesting its function is primarily for the repair of DNA damage. Pol γ is found in mitochondria, for the replication of mitochondrial DNA. The bulk of genomic DNA replication therefore lands on Polymerases δ and ϵ , who replicate DNA for lagging and leading strand synthesis respectively. Both Pol δ and Pol ϵ also function as repair enzymes, with Pol δ functioning in mis-match repair (MMR), base-excision repair (BER), nucleotide-excision repair (NER), and double stranded break repair (DSBs). Pol ϵ also functions in the repair of BER, NER, and DSBs, as well as it's additional role in the s-phase checkpoint [149]. It is therefore logical that errors in DNA replication will arise from malfunction of Pol δ or Pol ϵ , as they carry the majority of DNA replication and repair in the genome. This is done with remarkably high fidelity, with only one incorrect base for every 10^9 to 10^{10} nucleotides replicated [150].

This high fidelity of DNA polymerases relies on the proofreading of the 3' exonuclease domain, as well as post-replication surveillance by MMR pathways. Deficiency of the MMR pathway has long been shown to increase the risk for the development of CRC, through high frequency of microsatellite instability (MSI), and is present in approximately of 15% of all sporadic cases. Germline mutations of the MMR pathways also account for hereditary forms of CRC (5% of all cases), such as Lynch syndrome (also known as HNPCC), which have deficiency in the MMR genes (*MLH1*, *MSH2*, *MSH6*, *PMS2*)[151][152]. For example, hypermethylation of the promoter region of the *MLH1* gene results in *MLH1* silencing, causing microsatellite instability through MMR deficiency, which is one of the most common occurrences in MMR disruption in CRC[151][153] [154].

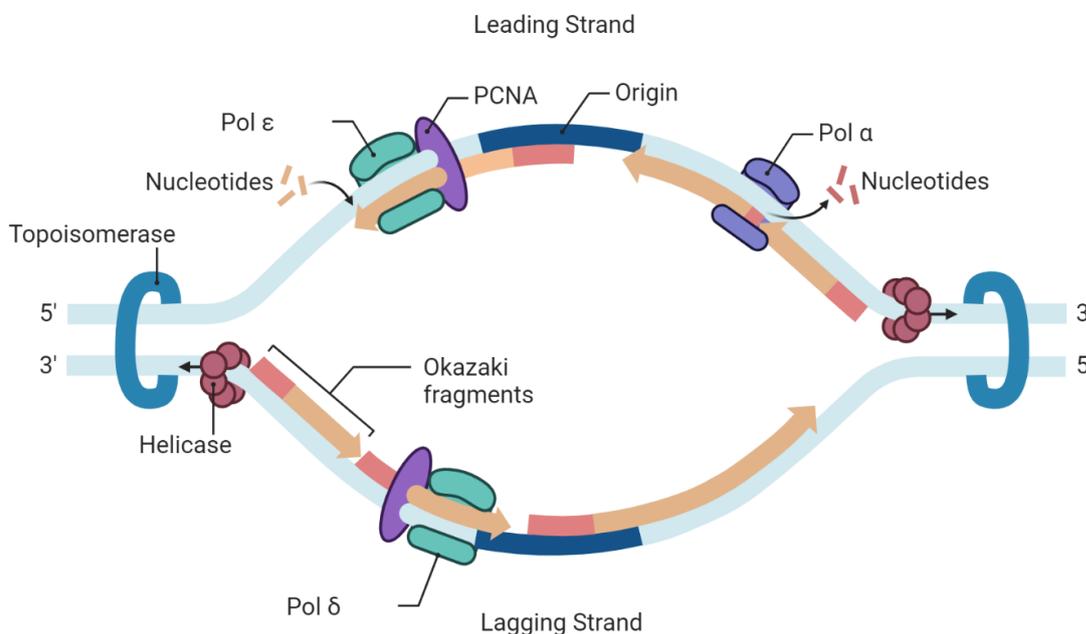


Diagram 5: Eukaryotic DNA replication by DNA polymerases α , δ , and ϵ . The replication fork progresses with the assistance of various enzymes and proteins. Helicase unwinds the parental DNA strands, while topoisomerase relieves the torsional strain ahead of the replication fork. On the leading strand (synthesized continuously in the 5' to 3' direction), DNA polymerase ϵ (Pol ϵ) extends DNA synthesis from an RNA-DNA primer laid down by DNA polymerase α (Pol α). Proliferating cell nuclear antigen (PCNA) acts as a sliding clamp to enhance processivity. On the lagging strand, replication occurs discontinuously, forming Okazaki fragments. These fragments are initiated by Pol α and extended by DNA polymerase δ (Pol δ) with the aid of PCNA. The direction of strand synthesis is indicated by orange arrows, and the origin of replication is marked on the DNA. Nucleotides are incorporated into the growing strands during synthesis.

Within mutations of the polymerase family, deficient proof-reading by exonuclease activity is the most well characterised. Pol δ and Pol ϵ are both capable of synthesising DNA as well as the degradation through exonuclease activity when an error is observed. Germline and somatic mutations within the exonuclease domain have recently been discovered which confer an increased risk of cancer due to high probability of replicative errors. Recent attention has been drawn to germline mutations in the catalytic subunits of Pol δ and Pol ϵ , known as POLD1 and POLE, which disrupt the proof-reading domains and predispose to CRC and other malignancies [150] [155] [156][157] [158]. Pol δ and Pol ϵ are part of the B family of polymerases and have 3'-5' exonuclease activity which proofreads newly synthesised DNA. Somatic mutations of POLE are found in 1-2% of CRCs, where they produce ultra mutated tumours. POLE mutations are also found in 7-12% of endometrial cancers [159], as well as breast, stomach, pancreas and brain [160][161][162].

Pol δ and Pol ϵ proof-reading is carried out by several highly conserved exo motifs in the exonuclease domains, within which lie the catalytic site residues that are essential for exonuclease activity (D316 and E318 in Pol δ , and D275 and E277 in Pol ϵ). Misincorporation of a base into the DNA strand causes the polymerase to pause and switch from the catalytic to the exonuclease domain. The incorrect base is then excised, and the correct base inserted [150]. In studies of mice that harbour alanine substitutions of the exonuclease site, there is deficient exonuclease activity, resulting in increased mutation rate and the development of tumours [163]. Notably, this is only found in homozygous animals for proof-reading deficiency of POLD1 and POLE. *POLD1* mutant mice develop lymphomas and carcinomas of the skin and lung, whereas *POLE*-mutants produce intestinal tumours and histiocytic sarcomas[150].

Mutations in the proof-reading domains of POLD1 and POLE are found in human carcinomas, with mutations in the P286R and V411L residues found in ~7% of sporadic endometrial cancers, associated with ultra-mutation and microsatellite instability [159]. In familial studies of colorectal cancers, linkage analysis and whole genome sequencing showed heterozygous germline mutations of *POLD1*^{S478N} and *POLE*^{L424V} [155]. Current data suggest that germline *POLE* and *POLD1* mutations are present in 0.5–2% of patients in intestinal polyposis and CRC cohorts enriched for familial disease [150].

Mutations in the exonuclease domain of POLE and POLD1 do not appear to present any abnormal clinicopathological features in comparison to other causes of CRC. However, germline POLE and POLD1 mutations seem to occur in patients of young age at onset (typically <50 years) compared to other CRC patients [155]. The majority of germline and sporadic cases do not appear to be microsatellite instable (MSI), however they display ultra mutation, and are predominantly base substitute mutations, with a unique signature of 100-

fold increase in C→A transversion. This results in a strong bias for substitution of serine to tyrosine or leucine, arginine to isoleucine or glutamine, and an increase in glutamic acid to stop codon mutations [164]. This often results in missense and truncation mutations in oncogenes and tumour suppressors, such as that of PI3K, PTEN, APC, MSH6 and P53 [150].

The other components of the Pol δ and Pol ϵ complexes have also been highlighted as potential causes of disease. POLD2,3 and 4 serve several important roles, such as the stabilisation of the holoenzyme complex, and stimulation of polymerase activity via interaction with the proliferating cell nuclear antigen (PCNA) clamp to Pol δ [165]. Within Pol ϵ , the second subunit (POLE2) mediates the interaction with GINS complex, helping to target the Pol ϵ holoenzyme to the leading strand during the initiation of DNA replication. The third and fourth subunits (POLE3, 4) are crucial for binding double-stranded DNA and for processive DNA synthesis and processive 3'–5' exonuclease degradation.

Interestingly, the *POLD3* gene, the third subunit of Pol δ , has been highlighted as harbouring risk SNPs for the development of CRC, and depletion of POLD3 manifests into chromosomal instability and high DNA damage [166]. POLD3 has been implicated in breast cancer formation, with high levels of the protein in tumours resulting in poor clinical outcomes for patients. Furthermore, it has been shown to be stimulated by E2F1, to promote carcinogenesis. POLD4, the final subunit of the Pol δ complex, has also been shown to be highly expressed in many cancer types, and promotes cellular proliferation in glioma [167]

Other, lesser studied polymerases have also been implicated in the initiation and progression of cancer. In fact, with each discovered eukaryotic polymerase, there has been a cancer-related function identified and characterised, giving status to the role of DNA polymerases in cancer progression. For example, Pol β , encoded by the *POLB* gene, has gained attention due to its roles on ovarian, bladder and breast cancer [168–170]. Pol β is the main polymerase involved in BER, providing gap-filling synthesis at apurinic/apyrimidinic sites of damaged DNA. Several risk SNPs have been identified in the *POLB* gene as producing significant increases in the development of cancer. Furthermore, somatic mutation of *POLB* has been associated with a “mutator phenotype” in colon cancer, with the mutant K289M shown to synthesise DNA at a lower fidelity than wild-type counterparts, resulting in the manifestation of genomic instability [171]. Another mutant of *POLB* has also been identified in colon cancer as interfering with BER, again increasing genomic instability. On the other hand, overexpression of *POLB* resulted in abnormal telomere production, and destabilisation of other cellular processes which eventually lead to tumour progression. For example, higher expression on *POLB* resulted in increased genetic instability due to its lower fidelity in DNA replication compared with other, higher fidelity polymerases. Furthermore, enhanced Pol β expression may lead to

enhanced chemotherapy resistance, such as to that of drug Oxaliplatin. These cells were able to bypass genomic crosslinks, enabling cells to replicate mutated DNA in the presence of chemotherapy [171].

Other polymerases, such as those belonging to the Y family, are specialised in bypassing DNA damage during replication through a mechanism known as 'translesion synthesis' (TLS). TLS involves the switching to Y family proteins to facilitate the insertion of bases opposite damaged nucleotides. However, these translesion polymerases have low fidelity and are often prone to error. Their ability to promote replication past DNA lesions prevents the collapse of stalling replicative forks, and maintains genome stability, but at the risk of increased mutagenic rates. This process is often used in response to DNA damage, such as the use of Pol η , in the response of UV radiation. However, this ability to bypass DNA damage can also confer resistance to chemotherapy. For example, Pol η , Θ , and ν , can confer resistance to platinum-based drugs (cisplatin, oxaliplatin and carboplatin) as well as to replicative stress-inducing agents, including AraC and gemcitabine [172][173].

It is clear the DNA polymerases play a substantial role in cancer formation. Given their intrinsic role in DNA replication and repair, it is unsurprising that germline and somatic mutations in the various Pol genes will increase the incidence and development of CRC, through manifestations of chromosomal and genetic instability, as well as high mutagenic load. However, high expression of polymerases can also confer greater proliferative capacity, and the ability to evade standard chemotherapies through the repair of DNA lesions. Therefore, Polymerases can be seen as an important biomarker in cancer progression and response to therapy.

Metastasis

Metastatic CRC occurs when Cancer cells from primary tumours infiltrate the vascular system and colonise distal sites. Metastatic CRC occurs in stage 4, often after previous treatments at stages 1-3 have failed, or may present de novo at stage 4. The most common sites of metastatic CRC are the lymph nodes, liver, lung, and peritoneum [17]. Cancers originating from the rectum are more frequently metastasized into thoracic organs, such as the liver and the nervous system, compared to colonic. On the other hand, mucinous and signet ring adenocarcinomas are more frequently metastasized within the peritoneum compared with generic adenocarcinoma, and less frequently into the liver. Metastasis to the lung is often associated with nervous system metastases, whereas peritoneal metastases were often associated with ovarian and pleural metastases [17].

At stage 1 of CRC, patients are less likely to develop metastasis, at 10% likelihood. This is increased at stage 2 to 10% - 20% likelihood of metastasis, and 25% to 50% for stage 3, which is characterised by lymph node-positive CRC [18]. However, complete surgical resection and adjuvant chemotherapy for stage 3 CRC decreases recurrence and risk of metastases to 20% to 30%.

Metastatic competency is governed by several factors, including cell-intrinsic causes, such as genetic abnormalities and activation of the epithelial-mesenchymal-transition pathway (EMT), as well as cell external factors such as the tumour microenvironment and immune infiltration.

The EMT pathway occurs when cells lose their epithelial characteristics and gain those of mesenchymal cells, which allows for increased migratory ability, invasiveness, and the ability to resist apoptosis. [19]. More than 90% of human CRC cell lines exhibit at least partial EMT, making it a clear target for treating metastatic CRC [20]. The EMT pathway occurs when polarised epithelial cells undergo biochemical changes in which they adopt a mesenchymal phenotype.

Epithelial cells normally interact with the basement membrane through its basal surface, but under EMT activation, degrade the basement membrane through secretion of ECM components. This allows the invasion of the underlying basement layer, and detachment of the original epithelial layer. There are multiple biological processes that occur for an epithelial cell to go through the EMT pathway. Epithelial cells will lose epithelial markers, such as E-cadherin, cytokeratin, MUC1 and laminin-1, and begin to show mesenchymal markers such as N-cadherin, betaBeta-catenin, and vimentin [21]. Within cancer cells, EMT is activated by the induction of several key pathways emanating from the tumour-associated stroma, including TGF- β , ERK, MAPK, PI3K, Akt, Smads, RhoB, and β -catenin [22].

The tumour microenvironment also plays a role in the risk of metastasis of a tumour. The tumour microenvironment comprises the extracellular matrix (ECM) and basement membrane (BM), endothelial cells, adipose cells, tumour-infiltrating immune cells, cancer-associated fibroblasts (CAFs), neuroendocrine cells, pericytes, as well as a multitude of signalling molecules that influence tumour progression [23]. Cancer cells secrete growth factors and cytokines such as IL-6, IL-1 β , TGF- β 1, TGF- β 2, FGF-2, and PDGF. The growth factors recruit and reprogram stromal cells, such as immune cells and fibroblasts, as well as enzymes that degrade and remodel the surrounding ECM and BM, such as matrix metalloproteinases (MMPs).

Metastatic competency also relies on the vasculature surrounding the tumour. In order for epithelial cells to invade distal sites, they need to enter the blood stream to travel to other areas. Therefore, well vascularised tumours facilitate increased migration of cancer cells into the blood stream, allowing for metastasis to occur. Cancer cells may also secrete factors that will promote the growth of blood vessels, known as angiogenesis, which allows delivery of oxygen and other nutrients to the tumour [23]. This is a response to the increased hypoxic and acidic environment of tumour tissue which can arise due to rapid cell growth. Hypoxic inducible factors (HIFs) can stimulate endothelial cells such as vascular endothelial growth factor (VEGF), which begins the process of angiogenesis. In turn, this also promotes the possibility of metastasis through increased opportunity of cancer cells to enter the blood stream[23].

Tumour microenvironment and immune landscape

As discussed, the tumour microenvironment can play influential roles on the metastatic competency of colorectal cancer. Growing evidence indicates that non-cell-autonomous pathways (pathways that externally affect cell behaviour) significantly affect CRC progression. The tumour microenvironment, or TME, is the biological environment formed by malignant and non-malignant cells and their components. The main components of the TME are the extracellular matrix, which contains immune and endothelial cells, and fibroblasts (Diagram 6). Furthermore, due to the close proximity of the intestinal lumen, gut microbiota may also play significant roles in the TME.

TUMOR MICROENVIRONMENT

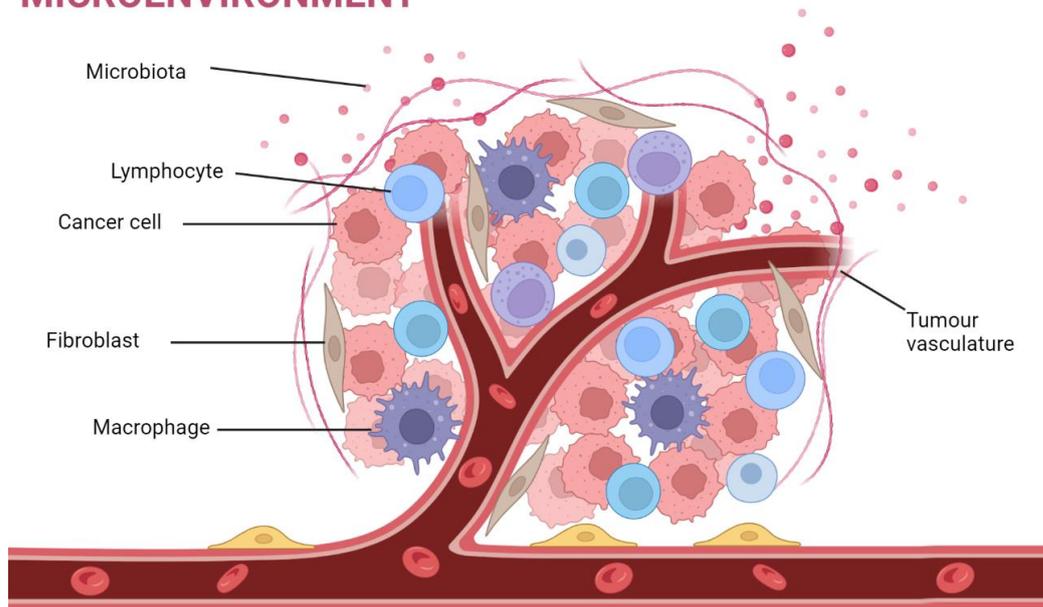


Diagram 6: Schematic representation of the tumour microenvironment. The tumour mass is comprised of cancer cells, fibroblasts, macrophages, lymphocytes and microbiota. The tumour mass is supported by the angiogenesis of the tumour vasculature which enables delivery of oxygen to tumour cells and facilitates metastasis of tumour cells. Infiltrating lymphocytes and macrophages are depicted within the tumour tissue, indicating immune cell involvement. Cancer cells comprise the core of the tumour, supported by fibroblasts that contribute to the stromal matrix. This complex microenvironment plays a critical role in tumour progression and response to therapeutic interventions.

For example, *H. pylori*, *B. fragilis*, and *F. nucleatum* infection in gastric and colorectal cancer has been well recognized, and can stimulate the WNT signalling pathway, which is hyperactive in nearly all CRC cases. In fact, many microbiota have found to increase colorectal carcinogenesis, such as *Enterococcus faecalis* (*E. faecalis*), *Streptococcus gallolyticus* subsp. *gallolyticus* (*Sgg*), *H. pylori*, *Bacteroides fragilis* (*B. fragilis*), *Clostridium septicum* (*C. septicum*), *Escherichia coli* (*E. coli*), and *Fusobacterium nucleatum* (*F. nucleatum*) [24]. *E. faecalis* produces an extracellular superoxide which induces DNA damage and genomic instability in colonic epithelial cells [25]. *Streptococcus gallolyticus* subsp. *gallolyticus* (*Sgg*), is also well characterized in enhancing CRC progression, with *Sgg* activating oncogenic pathways such as WNT/ β -catenin, c-MYC, and PCNA, and thereby promoting CRC [26].

Immune infiltration has also been shown to be crucially important in CRC development. Tumour infiltrating lymphocytes, such as CD4⁺T cells, CD8⁺T cells, B cells, and NK cells, are involved in tumour recognition, destruction, and elimination [27]. Macrophages, a myeloid lineage cell, have been shown to inhibit the proliferation, migration and invasion of CRC cells, therefore also contributing to the reduction in CRC formation. Macrophages can be classified as classical (M1) or alternative activated (M2) subtypes. During normal immune responses, most macrophages differentiate to the M1 phenotype, which inhibits CRC and is involved in Th1 cytokine responses upon pathogen challenging [28]. However, M2 macrophages may promote tumour progression through production of epidermal growth factor and fibroblast growth factor-1 which promotes cancer cell growth, as well as vascular endothelial growth factor A, which promotes angiogenesis. Furthermore, they may release matrix metalloproteinases which can degrade the ECM to promote invasion. In addition, M2 macrophages inhibit immune responses by producing immunomodulators: IL-10, IL-6, and TGF- β 1[29].

Neutrophils may also play both anti-tumorigenic and tumour-promoting roles in CRC. Recent studies have indicated that neutrophils can restrict the microbiota in tumours to reduce CRC progression and metastasis in mouse models [30]. However, evidence suggests that neutrophils may also stimulate CRC progression and metastasis through the CXCL1/CXCR2 chemokine axis, and degrade the ECM microenvironment through production of MMP9, therefore promoting cellular invasion [31].

Cancer-associated fibroblasts (CAFs) make up a large proportion of the tumour microenvironment, and promote cellular proliferation, migration and metastasis, as well as tumour angiogenesis[32]. CAFs secrete various cytokines, chemokines and miRNAs that interact with cancer cells to promote CRC progression. CAFs stimulate pathways such as KRAS, MYC, and TGF- β . For example, CAFs secrete miRNA-17-5p, which targets RUNX family transcription factor 3 (RUNX3). RUNX3 interacts with the proto-oncogene MYC and binds to the promoter of TGF- β 1, thereby activating the TGF- β signalling pathway [33]. The RUNX3/MYC/TGF- β 1 pathway promotes CRC proliferation, chemoresistance, and metastasis. CAFs may also help to shape the immune landscape, as they are negatively correlated with tumour-infiltrating lymphocytes, but were found to recruit monocytes and promote M2 polarization of macrophages [34]. Furthermore, CAFs may promote endothelial cells to release VEGF, leading to angiogenesis [35].

Endothelial cells also make up a large component of the TME. As discussed, endothelial cells are required for the composition of vascular formation during angiogenesis and produce VEGF and other growth factor receptors to enhance blood vessel growth. In addition to this,

endothelial cells may produce a soluble form of Jagged-1, which in turn activates Notch signalling to promote the cancer stem cell phenotype in CRC cells [36]. Endothelial cells may also express adhesion molecules, such as E-selectin, to facilitate CRC invasion and metastasis [32]. As with CAFs, endothelial cells modulate the immune landscape of the TME, expressing FasL to eliminate CD8⁺ T cells and enhance immune evasion of cancer cells [37]. Other proteins produced by endothelial cells include SPARC, COL1A1, COL1A2 and IGFBP3, which produce immune-inhibitory responses [32]. E-selectin, which is also produced by endothelial cells in the TMA, may attract neutrophils to further establish an immune-suppressive environment [38].

GWAS studies in CRC patients

Genome wide association studies have revealed over 100 common genetic variant SNPs that can increase CRC risk [43]. A 2008 study on HMPS-affected individuals identified a disease locus, known as *CRAC1*, to chromosome 15q13.3–q14 in three HMPS- families with multiple affected individuals [178]. Fine-mapping of the *CRAC1* locus constricted the risk area to a region containing 3 genes; *SCG5* a chaperone protein (chr15:30,721,252–30,776,590), *FMN1*, which is involved in adherens junction formation (chr15:30,846,102–31,147,525), and *GREM1*, a BMP antagonist (chr15:30,797,497–30,814,158). This region was associated with low-penetrance predisposition to CRC. An additional 2008 study found several loci associated with CRC risk [179]. The first strongest loci was identified by SNPs rs961253 and rs355527, which mapped 342 kb telomeric to the *BMP2* gene. The next strongest loci mapped 9.4 kb from the transcription start site of *BMP4*. The third strongest loci found was that of *CDH1* (E-cadherin), which is downregulated upon aberrant WNT activation. Another meta-analysis found a strong association of CRC risk by SNPs located in intron 3 of *SMAD* (rs4939827, rs12953717 and rs4464148) [180]. A 2011 study confirmed common susceptibility variants of the BMP pathway in genes *GREM1*, *BMP4* and *BMP2* [181]. *BMP5* was also identified in a separate 2018 study [182]. It is clear that SNPs targeting the BMP pathway are therefore often associated with increased CRC risk.

Several genes directly affecting the WNT pathways have also been identified. In 2015, 6 more SNPs were identified for increasing CRC risk [183]. Among them SNPs in the region of *CTNNB1*, which encodes beta-catenin, was found to be most strongly associated with CRC. A second locus identified located at 3p14.1 (rs812481) is intronic of *LRIG1*, a gene encoding a transmembrane protein that interacts with epidermal growth factor receptor-family.

LRIG1 has been described as a marker of quiescent colon crypt stem cells activated to proliferate following injury.

Another study conducted in 2012 found common variants near the locus for *CDKN1A*, *POLD3*, and *SHROOM2* to be influential in colorectal cancer risk [184]. *CDKN1A* encodes p21^{WAF1/Cip1}, which mediates p53-dependent G1 growth arrest, and as a master effector of multiple tumour suppressor pathways that function independently of classical p53 tumour suppression. *SHROOM2* is known to have broad roles in cell morphogenesis during endothelial and epithelial tissue development, and harbour recurrent mutations in cancer cell lines. *POLD3*, which is part of the polymerase delta complex, was also identified as containing risk SNPs in the study. Several other genes encoding DNA polymerase have been identified as containing risk SNPs for CRC, such as *POLE* and *POLD1* [185], making DNA polymerases another clear functional target for CRC risk.

Other genes identified by GWAS as coding risk SNPs include *APC*, *MLH1*, *MSH2*, *MSH6*, *PMS2*, , *SMAD4* (another BMP pathway protein), *NTHL1*, *LKB1/STK11*, *MUTYH*, *TFEB*, *SPSB2*, *TCF7I2*, *PNKD*, *TMBIM1*, and *TOX2*, among others [186–188]

A recent paper by Law et al. looked to confirm and expand on this list by a GWAS analysis which analysed 4,627 CRC cases and 71,379 controls of European ancestry that identified 53 new risk SNPs across 31 loci. This research filtered out SNPs with an allele frequency of <0.5% and imputation quality score <0.8. They then assessed associations between CRC and SNP genotypes in each study using logistic regression. This analysis found an over-representation of binding for MYC, which is often overrepresented in CRC and is a downstream target of WNT, as well as *POLD3*, *CHRD2*, *ETS2*, *RAD21*, *SMC1A* and *SMC3*. TCF4 is a key transcription factor of WNT, and was also shown to have a risk SNP in its active promoter sequence [43].

Project Background and rationale

Genome wide association studies

In this project we will investigate two neighbouring risk genes that have been highlighted within a recent GWAS; *POLD3* and *CHRD2* [43].

Interestingly, these genes are located together on the long arm of chromosome 11 (11q13) and have been chosen as they lie within target loci that represent physiological significance within the context of CRC, which is supported by previous studies [189][190][177]. Furthermore, unlike the other risk genes discussed above, very little is known about *CHRD2* and *POLD3* in the setting of CRC. Therefore, studying how these genes may affect the incidence and development of CRC is important to uncovering future screening methods and patient therapies. Both *CHRD2* and *POLD3* genes have been shown to be overexpressed in CRC, and elucidating the risk these genes may prove to CRC development, we may reveal novel therapeutic targets and enhance screening for CRC risk.

A meta-analysis of 8,682 cases and 9,649 controls has revealed 3 new SNPs within this region that may contribute to CRC risk. One of these SNPs lies at 11q13.4 (rs3824999) which lies within intron 9 of *POLD3* and is associated with increased CRC risk at genome wide significance of $P = 3.65 \times 10^{-10}$ [148]. However, at the 11q13.4 locus, SNP rs72977282, which maps 3,188 bp 5' to *POLD3*, was more strongly associated with CRC than rs3824999 [148]. It should be noted that there are multiple risk SNPs within the *POLD3* gene, and that most of these are common variants. Furthermore, these SNPs lie within intron sequences so are likely to affect gene transcription levels rather than protein sequences. However, this study did not significantly associate rs3824999 with a change in *POLD3* expression, but these analyses could only detect >5% differences in RNA expression by genotype with 80% power at a single time point, so does not rule out more subtle differences to gene expression control, or other SNPs lying within the same region [148]. Another paper has also identified a new low-frequency risk SNP At 11q13.4, near *POLD3* and *CHRD2*, (lead SNP rs61389091, MAF 3.94%) separated by a recombination hotspot [191]. Therefore there are at least 2 independent risk associations at 11q13.4. One within, and likely to be associated with the *POLD3* locus, and another within, and likely to be associated with the *CHRD2* locus.

Despite efforts through GWAS to elucidate risk SNPs, many of the genes associated with CRC are not yet fully understood. SNPs may lie outside of protein coding regions and so do not affect protein sequence, but expression changes to these genes. Furthermore, the loci at the *POLD3/CHRD2* region displays linkage disequilibrium and is therefore impossible to narrow

down the causal SNP. Therefore, it is important to discover the effects of both POLD3 and CHRDL2, and how they may affect the incidence, development, and prognosis of CRCs.

CHRDL2

Chordin-like 2 (CHRDL2) was first discovered in 2003 through differential analysis of cDNA fragments in vascularized breast tumours from cancer patients [139]. First termed BNF-1, it was found through qPCR to be expressed in liver, heart, prostate, testis, and ovary tissues, and further qPCR of matched paired tumour tissues found CHRDL2 to be overexpressed in breast, lung, and colon tumours [139][140].

CHRDL2 has been shown to undergo extensive and complex alternative splicing in different tissue types, with alternative isoforms possibly affecting BMP binding affinity [142]. CHRDL2 is also known to be secreted into the extracellular matrix, and staining patterns have shown CHRDL2 localisation in intestinal crypts, further supporting its role as a BMP antagonist.

. Co-immunoprecipitation assays have shown that CHRDL2 binds to BMP2, 4, and 6. CHRDL2 antagonism of BMPs resulted in blocked SMAD1/5 phosphorylation, thereby promoting CRC cell proliferation and inhibiting apoptosis. Furthermore, CHRDL2 prevented BMP mediated cell cycle arrest, thereby promoting cell cycle progression and increased CRC cell proliferation in vitro and in vivo, through up-regulation of Cyclin D and downregulation of P21 [144]. *In vivo* xenograft tumours injected with CHRDL2 overexpressing cells in mice displayed increased weight and volume of tumours, with increased proliferation shown by increased Ki67 staining.

Interestingly, investigations into sporadic CRC using genome wide association studies (GWAS) highlighted several single nucleotide polymorphisms (SNPs) in the CHRDL2 promoter sequence and throughout the gene [148]. This study found that the area surrounding the CHRDL2 gene was significantly highlighted as containing several SNPs that may lead to an increased risk for developing CRC. Whether these risk SNPs enhance CRC risk by affecting alternative splicing of CHRDL2 or by modifying CHRDL2 expression levels is yet to be elucidated. It is clear that as a BMP antagonist, CHRDL2 has a definitive role in the inhibition of BMP and promotes tumorigenicity. However, the full effects that CHRDL2 may elicit functionally in the development of CRC is unknown. So far, there has been no evidence into the pathways that CHRDL2 may upregulate in response to BMP inhibition in colorectal cancer. Furthermore, although there has been speculation that CHRDL2 may increase resistance to chemotherapy in other cancer types, there has been no research into this role in the context of CRC.

CHRD2 has also been shown to play a role in other cancer types. For instance, CHRD2 overexpression was found in osteosarcoma tissue compared to healthy adjacent tissue. Furthermore, CHRD2 knockdown in osteosarcoma cells lines decreased proliferation, migration, and clonogenicity, which are seen as hallmark characteristics of cancer cells. In contrast to CRC cell lines, CHRD2 was only able to bind BMP9 in osteosarcoma. Inhibition of CHRD2 increased the combination of BMP-9 to its receptor ALK1, which prevented ALK1 inhibition of the PI3K/AKT pathway. CHRD2 overexpression therefore could mediate its effect through regulating PI3K/AKT, which is commonly aberrantly activated in tumour tissues, resulting in with tumour progression and chemotherapy resistance [146]. In gastric cancer, CHRD2 overexpression was found in gastric cancer cells, and was found to strongly correlate with later cancer stage and poor prognosis. CHRD2 was also found to increase cellular proliferation and clonogenicity, and xenograft tumours with CHRD2 were found to be larger than those of a control. Interestingly, CHRD2 overexpression in gastric cancer cell lines highlighted deregulation of the YAP/TAZ pathway. The YAP/TAZ pathway consists of transcriptional activators, which are found to correlate with cancer stem cells and chemotherapy resistance [147].

POLD3

POLD3 was first identified using proliferating cell nuclear antigen affinity chromatography (PCNA column) and glycerol gradient centrifugation from mouse and calf thymus, which revealed a distinct subunit that reacted strongly with both Pol δ complex and the PCNA binding domain [174] [175]. POLD3 has dual roles in Pol δ complex, first as a stabilizer of the POLD1-POLD2 interaction, and second as a facilitator to the binding of the POLD complex to PCNA through a C-terminal PIP box. POLD3 also exhibits 3' to 5' exonuclease activity, which increases the processivity of DNA synthesis during replication as seen in diagram 6. POLD3 mediates the binding of PCNA to Pol δ through a canonical PCNA-binding sequence located in its C terminus (83). Removal or malfunction of the POLD3 subunit prevents PCNA binding to Pol δ , and therefore ineffective DNA replication.

In addition, Pol δ has roles in DNA double strand break repair via homologous recombination, an important aspect of DNA repair. The Pol δ complex also participates in DNA mismatch and base-excision repair, which is a key process shown to be defective in many hereditary CRC susceptibility disorders, such as Lynch syndrome [176]. Phosphorylation of POLD3 by cyclin-

dependent kinases shows that POLD3 activity may be dependent on cell cycle regulation, an important factor when considering cancer cell cycle progression [177].

Ineffective Pol δ can lead to frequent mutations and genome instability. In cells depleted of POLD1 (the first subunit of Pol δ) and POLD3; research has shown a general increase in genome instability manifested through DNA breaks, S phase progression impairment, and chromosomal abnormalities [177]. It was shown that both proteins were required to maintain sites of open replication, and that POLD3 depletion caused anaphase bridge accumulations. It has been theorised those mutations effecting POLD3 expression increases genome instability and mutation rates, leading to the development of CRC [166]. In mice deficient of *Pold3*^{-/-}, there was replicative stress, micronucleation and aneuploidy . Furthermore, haplo-insufficient (*Pold3*^{+/-}) mice displayed impaired double-strand break repair, telomere shortening and loss, and chromosome breaks [166].

Due to the recent attention to the loci surrounding the *POLD3* gene as containing areas that are susceptible to mutations that increase CRC risk, we sought to analyse the functional effects of POLD3 in the context of CRC.

Project aims.

The aim of this project is to uncover the functional effects of common DNA risk variants; *POLD3* and *CHRD2*, within CRC.

We hypothesise that increasing *CHRD2* levels will lead to a repression of BMP signalling which in turn will increase WNT signalling causing downstream effects such as increased cell proliferation, migration, stem-like qualities.

We also hypothesise that knockdown of *POLD3* will increase the levels of genomic instability, mutation rates, and may activate downstream oncogenic pathways.

Objectives:

1. Generate model systems to investigate *CHRD2* and *POLD3*
 - Generate cell lines to overexpress *CHRD2*.
 - Generate cell lines to knockdown *POLD3*.
 - Test organoid models in response to *CHRD2*
2. Test the hypothesis that modifying *CHRD2* will affect CRC cancer pathways, including cellular proliferation, migration, differentiation, and chemotherapy resistance.
 - Determine the effect of *CHRD2* on gene transcription and the activity of known cancer pathways.
 - Investigate if cell lines overexpressing *CHRD2* have increased proliferative/migratory capabilities.
 - Investigate the effects of *CHRD2* expression on chemo-therapy drug resistance.
3. Test the hypothesis that knockdown of *POLD3* expression will affect genome stability in cancer cells.
 - Determine if cell lines with knocked down *POLD3* have increased proliferative/migratory capabilities.
 - Investigate the effects of *POLD3* expression on chemo-therapy drug resistance.
 - Determine if *POLD3* increases/decreases genetic instability or telomere maintenance.

Chapter 2: Methods

Cell culture and maintenance

Immortalised human colorectal adenocarcinoma cell lines CACO2 (HTB-37), COLO320 (CCL-220), LS180 (CL-187) and RKO (CRL-2577) were purchased from (Sigma-Aldrich, UK) and were maintained in T25 flasks with High Glucose Gibco Dulbecco's Modified Eagle Medium (DMEM) (Sigma-Aldrich) supplemented with 10% fetal bovine serum (FBS) (Sigma-Aldrich), and 1% penicillin streptomycin (Sigma-Aldrich). Cells were grown in an humidified atmosphere at 37°C with 5% CO₂.

Subculturing was performed every 72 hours to maintain a cell confluency of less than >80%. This was achieved through washing cells with phosphate buffered saline (PBS) (Sigma, UK), and then detaching using trypsin-EDTA (Life Technologies, UK). Medium was added to neutralise the trypsin, and cells were pelleted by centrifugation at 400 x g for 5 minutes and resuspended in fresh medium.

Cell lines with overexpression of the CHRDL2 gene and cells with POLD3 knockdown shRNA were routinely given puromycin to ensure removal of non-plasmid containing cells. Details of puromycin concentrations can be seen below:

Cell Line	Puromycin (mg / 10ml media)
CACO2	0.025
COLO320	0.005
LS180	0.012
RKO	0.005

Overexpression of CHRDL2

The generation of CHRDL2 overexpressing inserts was carried out by J Sandhu and A Lewis (unpublished). Overexpression of CHRDL2 was achieved using the Tet-on system. Firstly, the CHRDL2 full length cDNA (COSMIC) was cloned into pCW57.1 (Addgene #41393) using Gateway technology (Invitrogen, Thermo Fisher, US), followed by validation by Sanger sequencing and restriction digest. The vector was then transfected in HEK293 cells along with viral packaging vectors (2nd generation system – pCMV-dR8.2 and pCMV-VSV-G) using Lipofectamine 2000 (Invitrogen, US). Virus containing media was collected, sterilized and titre measured (Go-Stix, Takara).

Transduction to cell lines

The cell lines CACO2, COLO320 LS180, and RKO were transduced, and integrated with the pCW57.1-CHRD2 overexpression vector. To achieve this, cells were first seeded at a density of 100,000 cells per well in a 6 well plate and left to adhere for 72 hours. Next, the viral vector was administered to the cells in either a low virus concentration of 20 ng/ ml, or a high virus concentration of 50 ng/ml. A solution of 4 ml 2x polybrene (Biotechne), the virus containing media, and IMDM (Sigma) to a final volume of 8 ml was made. 2 ml of virus/polybrene solution was added per well and left for 24 hours. After this, media was refreshed with standard IMDM, and puromycin was added to select for virus-containing cells as described previously. Subculturing was performed every 2-3 days. We observed a greater number of cells in the high virus treated group, so this was selected for future experiments.

Knockdown of POLD3

Lentiviral shRNA of POLD3 was used to knockdown gene expression. Target sequences were implemented to create shRNA oligos, which were then inserted into the EZ-Tet-pLKO-Puro vector (Figure 1), which allowed tet-on induction of POLD3 shRNA production.

Target sequences were identified through the use of the Broad institute GPP web portal.

Two shRNAs with different target sequences were created to control for off target effects.

Target sequences are as follows:

shRNA CTAGCACTGACTATGACATCCTTAAATACTAGTTTTAAGGATGTCATAGTCAGTTTTTTTG
233258

shRNA AATTCAAAAACTGACTATGACATCCTTAAATGATCATTAAAGGATGTCATAGTCAGTG
233261

and then centrifuged again for 15 minutes at 30000g. The supernatant was removed, and the pellet was left to air dry at room temperature.

The pellet was then resuspended in 500 μ l nuclease free water. To phosphorylate the annealed oligos, 85 μ l of oligos in water were given 10 μ l ligase buffer (NEB), and 5ul of PNK (NEB). This was incubated for 1 hr at 37 °C , before heat inactivation at 65 °C for 20 minutes. Oligos were then diluted to a working concentration of 100 ng/ μ l

Plasmid Ligation

Ligation of the EZ-Tet-pLKO-Puro empty vector and shRNA target oligos was performed using the Promega T4 DNA ligation kit (PROMEGA, UK).

In a microcentrifuge tube the following was added:

Linearized vector DNA (100 ng)	2ul
Insert DNA	11ul
10X Rapid Ligation Buffer	2 μ l
T4 DNA Ligase	1 μ l
Nuclease free water	70 μ l

Ligation mixture was vortexed and then incubated overnight at 4 °C .

Transformation

JM109 competent *E.coli* cells (PROMEGA) were given 2 μ l of ligated vector and incubated on ice for 1 hour. 50 μ l of competent cells were then heat shocked at 42 °C for 40 seconds, before returning to ice for 2 minutes. Cells were then given 500 μ l of LB medium and allowed to recover at 37 °C for 4 hours with slight agitation. 200 μ l of cells were then plated on LB-agar plates containing ampicillin (100 μ g/ml) and incubated overnight at 37 °C . Single colonies were grown in 1 ml LB medium at 37 °C for 8 hours and then expanded in 100 ml LB medium overnight. Bacterial cultures were then pelleted by centrifugation at 4000g.

Mini prep and restriction digest of ligated clones

DNA was extracted from samples using DNA Maxi kit (QIAGEN). Sample concentrations were measured via Nanodrop.

Colony PCR

Plasmid DNA extracted from bacterial cultures was used in Polymerase chain reaction (PCR) to validate the inclusion of the target sequence in the plasmids.

A master mix was prepared as below with quantities multiplied by number of samples required:

- 1 µl Forward primer (Sigma-Aldrich)
- 1 µl Reverse primer (Sigma-Aldrich)
- 5 µl 10x Buffer (Invitrogen)
- 4 µl dNTPS (Applied Biosystems, US)
- 1.5 µl Mg (Bio Line, US)
- 0.5 µl Taq polymerase (Bio Line)
- 36 µl RNase-free water

49 µl of the master mix was added to each well. Plasmid DNA was diluted 1:10 and 1 µl of plasmid DNA was added to the desired well. The plate was then incubated at 94 °C for 5 minutes. The plate was then cycled at 95 °C for 45 seconds, 55 °C for 1 minute, and 72 °C for 1 minute. This was repeated 34 times.

1 µl of the resulting DNA was added to 9 µl Orange G (Sigma-Aldrich) and samples were run on a 2% Agarose gel supplemented with 0.01% SYBR safe (Invitrogen) at 120 V for 20 minutes.

Plasmid sequencing

Samples were sent to the company GENEWIZ (Oxford, UK) for sequencing.

Viral packaging

The vector containing POLD shRNA was then transfected in HEK293 cells along with viral packaging vectors (2nd generation system – pCMV-dR8.2 and pCMV-VSV-G) using

Lipofectamine 2000 (Invitrogen, US). After 24hrs, virus containing media was collected, sterilized and titre measured (Go-Stix, Takara).

Transduction to cell lines

The cell lines CACO2, COLO320 LS180, AND RKO were transduced, and integrated with the pCW57.1-CHRD2 overexpression vector. To achieve this, cells were first seeded at a density of 100,000 cells per well in a 6 well plate and left to adhere for 72 hours. Next, the viral vector was administered to the cells in either a low virus concentration of 20 ng/ ml, or a high virus concentration of 50 ng/ml. A solution of 4 ml 2x polybrene (Biotechne), the virus containing media, and IMDM (Sigma) to a final volume of 8 ml was made. 2 ml of virus/polybrene solution was added per well, and left for 24 hours. After this, media was refreshed with standard IMDM, and puromycin was added to select for virus-containing cells as described previously. Subculturing was performed every 2-3 days. We observed a greater number of cells in the high virus treated group, so this was selected for future experiments.

Doxycycline treatment

To induce expression of the target gene, doxycycline was administered. For CHRD2 cells, doxycycline was given at 10 µg/ml, 1 µg/ml, and 0.1 µg/ml, termed CHRD2, CHRD2+ ad CHRD2 ++. For POLD3 cells, doxycycline was given at 10 µg/ml, 1 µg/ml, and 0.1 µg/ml, termed POLD3, POLD3-, and POLD3--.

Wound healing assay

To assess the migratory ability of cells, a wound healing assay was performed. Cells from culture were pelleted and resuspended in fresh media and seeded at a density of 1×10^5 /ml in a 6 well plate. To induce expression, cells were treated with 2 µl of doxycycline suspended in dimethyl sulfoxide (DMSO) (Sigma-Aldrich) at a concentration of 10 mg/ml to give a final concentration of 10 µg/ml, or 2 µl DMSO in the control. Cells were then incubated at 37 °C with 4% CO₂ for 72 hours. The medium was then removed, and using a sterilised 200 µl pipette tip, a scratch was made down the centre of each well. The cells were then washed with PBS to remove dead cells and replenished with fresh media. A second doxycycline treatment was then added at the same concentrations.

Wells were imaged at 0 hours and 24 hours post scratch. Wound size was measured via ImageJ software using the wound healing plugin. Percentage of wound healing was calculated by:

$$\frac{\text{Initial wound size}}{\text{Final wound size}}$$

Migration assay

To further assess migratory capabilities of cells through a barrier, a migration assay was performed. This assay involved plating cells within inserts that possessed a mesh bottom within a 24 well plate. The membrane of these inserts had a pore size of 8.0 μm .

From culture, cells were pelleted and resuspended in fresh serum-free media at a density of 1×10^5 /ml. 300 μl of resuspended cells was given to each insert, to give a total of 30,000 cells/well. 300 μl of complete medium was given to the wells beneath the insert. Doxycycline-inducible expression was given in the same concentration as the wound healing assay (see above). Cells were then incubated at 37°C with 4% CO₂ for 96 hours.

Inserts and media were then removed, and cells from the upper surface of the insert membrane were swabbed away. The membrane was cut and mounted on a coverslip, and cells on the underside of the membrane were counted.

Clonogenic assay

To assess the ability of single cells to generate colonies and cell survival ability, a Clonogenic assay was performed. Cells from culture were pelleted, and resuspended to a dilution of 100 cells/ ml.

Using a 6 well plate, 100 cells were plated per well and supplemented with 1 ml fresh media giving a final volume of 2 ml per well. Cells were treated with 2 μl doxycycline treatment as before in concentrations of 10 mg/ml and 1 mg/ml. The control was given 2 μl DMSO.

Plates were incubated for 2 weeks at 37 °C with 4% CO₂ until visible colonies were formed. Every 72 hours doxycycline and DMSO treatments were refreshed.

After 2 weeks, the media was aspirated, and the cells washed with 2ml PBS which was then aspirated. 1 ml of >98% ethanol which has been stored at -20 °C was given to each well of the 6 well plate to fix the cells. This was incubated for 10 minutes at room temperature, and then aspirated. To stain colonies, 1 ml of crystal violet stain (0.5% in 20% methanol) was added to each well and incubated at room temperature for 10 minutes, before aspiration.

Stained plates were submerged in baths containing clean water to remove excess stain. Wells were imaged, and colonies were counted by ImageJ.

Colony numbers were normalised to 1 for each plate, and then data compiled.

Proliferation assay / MTS assay

To assess the proliferation of cell lines, an MTS assay was performed. Cells from culture were seeded at a density of 1×10^5 in 70 μl standard media in a 96 well plate. Expression was induced by doxycycline at concentrations 10 $\mu\text{g/ml}$ and 1 $\mu\text{g/ml}$, with a ratio of 1 μl doxycycline treatment to 1 ml media. Control samples were given DMSO in the same proportions. 9 replicates were used for each concentration. Cells were incubated for 72 hours at 37 °C with 4% CO_2 . For low glucose proliferation assays, cells were resuspended in standard media conditions with Gibco Dulbecco's Modified Eagle Medium low glucose (DMEM) (Sigma-Aldrich). Cells were incubated for 120 hours, and MTS assay was performed.

To perform the MTS assay, each well was given 10 μL MTS assay solution (CellTiter 96® AQueous One Solution) (PROMEGA, UK) and plates were gently shaken and incubated at 37°C for 4 hours. Absorbance was read at 490 nm using CLARIOstar microplate reader (BMG Labtech, UK) .

Drug dose response curve

To assess the ability of cell lines to withstand treatment from commercial chemotherapy drugs, a drug dose response curve was performed. The chemotherapy drugs 5FU, Oxaliplatin, and Irinotecan are the most commonly used drugs to treat CRC within the UK, so these were selected for this experiment. All chemotherapy drugs were obtained in powdered form and solubilised in DMSO as per manufacturers guidance to give a concentration of 0.5 M. 5FU (F6627-1G, Merck), Oxaliplatin (O9512-5MG, Merck), Irinotecan (I1406-50MG, Merck).

A serial dilution of chemotherapy drugs in standard media was performed to give the following concentrations:

5FU (μM)	OXALIPLTIN (μM)	IRINOTECAN (μM)
0	0	0
20	5	5
100	25	10
500	50	30
1000	100	60
5000	500	100
10000	4000	300

Cells from culture were seeded on a 96 well plate at a density of 1×10^4 in 100 μl standard media with 10 $\mu\text{g/ml}$ doxycycline and incubated for 24hrs. PBS was added in the surrounding wells to prevent evaporation of media. The media was then aspirated, and 100 μl of the diluted drug with 10 $\mu\text{g/ml}$ doxycycline was added to the corresponding well and incubated for 72 hours.

For secreted drug dose response curves, cells were seeded in the same manner but as 0.5×10^5 cell density, and in half conditioned media, half standard media. Conditioned media was generated through addition of 10 $\mu\text{g/ml}$ doxycycline to CHRDL2 or parental cell lines at 70% confluency. After 48hrs media was harvested and filtered. Serial dilution of chemotherapy drugs was performed in 50:50 conditioned media and standard media.

An MTS assay was then performed as above to measure the numbers of surviving cells present. IC50 values were found using GraphPad PRISM, and two-way ANOVA was used to find differences between the curves. T-tests were used to find significant differences between IC50 values.

GraphPad Prism was used to draw drug dose-response curves after treatments. To find the IC50 values (half-maximal inhibitory concentration) a curve of best fit was generated using non-linear regression. Absolute IC50 values were then determined using baseline of 0.

Organoid preparation and maintenance

Organoids were maintained in a humidified atmosphere at 37°C with 4% CO₂. Organoids were grown in ADF media as described: ADVANCED DMEM/F12, 2mM GLUTAMAX, 1mM N-ACETYLCYSTEINE, 10mM HEPES. Supplemented with 1% PS, 10% B27, 5% N2 Growth factors were also given to media surrounding Matrigel: 1% Mouse recombinant Noggin (Peprotech, UK), 1% mouse recombinant EGF (Invitrogen), 5% Recombinant human R-spondin (Peprotech).

Wild-type mice were culled, and the small intestine was removed and washed with PBS. Villi and differentiated cells were scraped off the intestinal membrane using a glass microscope slide. Sections of intestine were then cut into 2mm segments and transferred to ice-cold PBS. Pipettes were coated in FBS, and intestinal segments were washed through pipetting up and down to dislodge single cells and debris. PBS was then removed, and washes were repeated 5 times.

Segments were then resuspended in 2.5mM EDTA/PBS to loosen crypts and rotated at 4 °C for 30 minutes. The supernatant was then removed, and segments were resuspended in ADF media. The entire volume was pipetted up and down several times, and then the supernatant removed and centrifuged for 5 minutes at 1200 rpm at 4°C. The supernatant was then removed, and the resulting pellet was resuspended in 10 ml ADF media and passed through a 70 µm cell strainer into a clean 15 ml falcon tube. The tube was then centrifuged for 2 minutes at 600 rpm at 4°C so that single cells will not be included in the pellet, and the supernatant was removed. This was repeated 3 times.

Finally, the pellet was resuspended in 50ul ADF media and 100ul Matrigel, and pipetted 40ul/well in a 24 well plate, with 500 µL of ADF media supplemented with growth factors. Passaging of organoids was repeated every 48 hours and consisted of transferring organoids to a 15 ml conical tube and pipetting up and down to break up organoids. Organoids were then centrifuged for 2 minutes at 600 - 800 rpm at 4°C and resuspended in ADF + Matrigel and plated as described previously.

For organoid counting, independent researchers were blinded to sample type and counted the number of organoids, and following chemotherapy, the number of live and dead organoids.

Organoid immunofluorescence staining

Organoid samples were prepared for staining by removal of growth media and pelleted through centrifugation at 600g. Organoids were then fixed through resuspension in 500 μ l formalin (Sigma-Aldrich) for 10 minutes, before pelleting at 400g and resuspension in 70% ethanol for 1 minute. Organoids were then pelleted at 400g and resuspended in 50 μ l of low gelling agarose (Sigma-Aldrich) and incubated on ice for 30 minutes, before embedding in paraffin blocks. Sectioning of organoids was performed at 5 μ M through standard microtome sectioning and left to dry on slides.

Slides containing organoid sections were dewaxed through xylene (Fisher Bioreagents) submersion for 5 minutes and rehydrated through submersion in ethanol at 100% 90% and 70% for 5 minutes. Antigen retrieval was performed by submerging slides in boiling 10 mM sodium citrate buffer (Sigma-Aldrich), before washing with PBS. Samples were then blocked through addition of Goat serum (Zytochem Plus, 2bscientific, UK) for 1 hour. Primary antibodies diluted in PBS were added for 1 hour, and secondary antibodies were incubated for 1hr in the dark. Coverslips were mounted using VECTASHIELD Vibrance [TM] Antifade Mounting Medium with DAPI (2bscientific) for imaging.

Organoid Antibodies:

Antibody	Catalogue number	Concentration used
OLFM4	39141, Cell Signalling, US	1:200
Alexaflour Goat anti rabbit	A-11011, Abcam	1:50

Flow cytometry

Cells were plated in six-well plates at a density of 1×10^5 cells/well under standard media conditions, supplemented with DMSO, or doxycycline treatments. For chemotherapy flow cytometry analysis, cell was treated with 25 μ M Oxaliplatin. Cells were grown for 48 hours before harvesting by trypsinization and washed once with cold PBS. To investigate cell-cycle progression, cells were resuspended in Hoeschst 33342 (62249, Thermo Scientific) and incubated at 37°C for 2 hours with slight agitation. Finally, samples were resuspended in PBS, and flow cytometry analysis was performed using ACEA Novocyte system (Agilent, US). The percentages of cells in various phases of the cell cycle were analysed through Novocyte software. To investigate apoptosis, cells were stained with ZombieAqua (423101, Biolegend, US) and annexin-V antibody (V13242, Thermo scientific) for 30 minutes each at room temperature. Subsequently, cells were washed once with PBS and analysed by flow cytometry using manufacturers software. For Ki-67 analysis, cells were incubated with Ki-67 BV711 (407-5698-82, Thermo-fisher) in a dilution of 1:1000 in standard media conditions for 30 minutes, before progression onto cell-cycle analysis staining.

Western blot

Sample preparation

For intracellular protein detection, cells were seeded in a 6 well plate at a density of 1×10^5 cells/ml. Cells were treated with doxycycline to induce CHRDL2 expression in concentrations of 10 μ g/ml, 1 μ g/ml, and 0.1 μ g/ml as stated previously. Cells were incubated at 37°C for 72 hours for CHRDL2 samples and 120 hours for POLD3- samples. Cell culture medium was aspirated, and the cells washed with PBS X2. 300 μ l of lysis RIPA buffer (Thermo Scientific, UK) was given per well, and cells were removed from the plate by cell scraper. Lysed cells were incubated on ice for 30 minutes.

For secreted protein expression, cells given doxycycline expression at 10 μ g/ml were grown in a T75 flask at a confluency of 70-90% and incubated for 72 hours. The media was collected and centrifuged at 400 x g through Amicon® Ultra centrifugal filters (Merck, UK) with pore size of 30 kDa to concentrate the samples and remove proteins <30 kDa. Media was replaced with PBS through centrifugation, and concentrated protein samples were resuspended with RIPA buffer.

Samples were then sonicated for 10 seconds at 100% power. Samples were stored at -80°C.

Bicinchoninic acid (BCA) assay

5 µl of protein samples were diluted with 45 µl RIPA buffer. Protein standards were created using BSA at concentrations of 0 – 1600 µg/ml. BCA assay was performed with using the BCE assay kit (Thermo-Scientific) and read at 592 nm. Protein samples were then diluted with RIPA buffer to achieve 30 µl per 20 µl for gel loading.

Protein separation and membrane transfer

30 µl protein samples (30 µg) were boiled at 95°C 5 minutes with 9 µl loading dye (NUPAGE, Invitrogen) and 1 µl sample reducing buffer (Invitrogen). Protein samples were separated via 4-12% sodium dodecyl sulphate polyacrylamide gel electrophoresis under 125 V, against Spectra™ Multicolor Broad Range Protein Ladder (ThermoFisher, UK), and then transferred onto the polyvinylidene difluoride membrane (Millipore, UK) under 20 V.

Blocking and antibody staining

Membranes were blocked with 10 ml semi-skimmed dried milk powder made in 1X Tris-buffered saline PH.8 (TBS) for 1 hr on a rocking platform. Membranes were then incubated with primary antibody in 10 ml TPBS-milk overnight at 4°C on a rocking platform.

Membranes were then washed with TBS with 1% tween (Sigma-Aldrich) for 10 minutes on a rocking platform which was repeated three times. A secondary was added in 10 ml TBS-milk and incubated on a rocking platform for 1 hour at room temperature. The membrane was then washed as before.

Western blot antibodies:

Antibody	Catalogue number	Concentration used
CHRD2 primary	AF2448, Abcam, US	0.1 µg/mL (1:2000)
POLD3 primary	ab182564, Abcam	1:250
Goat anti Mouse	ab205719, Abcam	1:2000
Donkey anti Goat	ab6885, Abcam	1:2000
B-actin	Ab8226, Abcam	1:2000

Imaging

Membranes were imaged through incubation with Enhanced chemiluminescence (ECL). The ratio of optical density of the bands was measured by a gel image analysis system (Bio-Rad) and normalized to β -actin.

Loading control

Membranes were washed as before and then incubated with loading control membrane at 4°C on a rocking platform overnight. Membrane was then washed and imaged as before.

Protein concentration was determined through ImageJ analysis. To do this, mean grey value was taken of each band. Loading control values were taken away from protein band values to determine normalised protein expression.

Immunofluorescence

Coverslip preparation

Coverslips were soaked in 1M HCl (150 ml) at 55 C overnight at 150 – 170 rpm. HCL was aspirated at coverslips were washed dH₂O 10-20 times. PH was checked to ensure in range 6-7 ph. 100 ml of poly-lysine (0.1 mg/ml in 100ml water) was poured over coverslips and left to shake at room temperature for 1 hour. Coverslips were washed with dH₂O and dried on filter paper overnight, before sterilization by UV light.

Cell preparation

Cells were grown on coverslips and treated with doxycycline at 10 μ g/ml or DMSO until 90% confluent. For oxaliplatin treatment, cells were given 5 μ M oxaliplatin.

Immunofluorescence staining

Coverslips were washed with ice cold PBS X2 and fixed with ice cold methanol for 10 minutes. Coverslips were then washed again with PBS X2.

Samples were permeabilised with 0.5 ml 0.2% Triton X for 2 minutes at room temperature and washed with PBS X2. Samples were then blocked in 1% BSA for 2 hours at 37 °C . Samples were then incubated with 100 μ l primary antibody H2AX (ab195188) for 30 minutes at 37 °C . Coverslips were then washed with PBS X2. 100 μ l of DAPI was added to coverslips for 2

minutes before coverslips were washed with PBS. Coverslips were placed on microscope slides and secured with DPX.

Slides were imaged with Leica DM4000 at 20x and 40x magnification using 346, 496 and 650nm wavelengths. During acquisition, all imaging parameters were kept the same during each experiment; exposure, gain, and pinhole size, to enable comparison of cellular fluorescence between samples.

Immunofluorescence antibodies:

Antibody	Catalogue number	Concentration used
H2AX	ab195188, Abcam	1:50
ATM	ab2354, Abcam	1:50
P53	25275, Cell signalling, US	1:800
Beta-catenin	610154, BD Biosciences, US	1:100
Ki67	D3B5, Cell signalling	1:100
RAD21	Ab992, Abcam	1:100
Ku70	Ab2171, Abcam	1:100
PCNA	Ab29, Abcam	1:100
ARTEMIS	MA5-46962, Thermo Scientific	1:100
BRCA1	MA1-23164, Thermo Scientific	1:100
ERCC1	Ab129267, Abcam	1:100
IQGAP1	33-8900, Thermo Scientific	1:100
EPCAM	ABC-304, Merck-Millipore	1:100
OLFM4	39141, Cell Signalling	1:200
Alexaflour Goat anti mouse	A-11001, Abcam	1:50
Alexaflour Goat anti rabbit	A-11011, Abcam	1:50

For Corrected total cell fluorescence, Image J was used. A background measurement was taken for each image, and a fluorescent measurement from 50 cells per replicate was taken. CTCF was then calculated using the following:

$$\text{CTCF} = \text{Cell fluorescence} - (\text{Area of selected cell} \times \text{Mean fluorescence of background})$$

Comet assay

Microscope slides were dipped in 1% agarose melted in dH₂O. Bottom of the slides were wiped on one side, and slides left to dry at 37°C.

Cells from culture were treated with 10 µg/ml doxycycline or DMSO, and incubated for 120hrs. Cells were then trypsinised, and pelleted. Pellet was then resuspended in PBS.

65 µl of 1% of low-temperature gelling agarose was resuspended with 10 µl of suspended cells (10³ cells) and placed on the agarose-coated slides. The agarose drops containing the cells was covered with a coverslip and placed at 4°C to set. The coverslips were then removed, and 75 µl of low-temperature gelling agarose was placed on top of the previous one, covered with a coverslip, and placed 4°C to set. After agarose slides were set, slides were placed in a Coplin jar containing lysis buffer and placed at 4°C for 1 hour. Lysis buffer was then removed, and slides were immersed in AES (alkaline electrophoresis solution) in the dark at 4°C to allow DNA unwinding for 1 hour. Slides were then placed into the electrophoresis tank and submerged in new AES buffer. The level of AES buffer did not exceed 2 mm from the top surface of the slide. The tank was placed at 4°C and electrophoresis conducted for 40 minutes at 20V. The slides were removed from the electrophoresis tank and rinsed with dH₂O. 100 µl of green fluorescence nucleic acid staining solution was placed on each slide, covered with parafilm, and incubated at room temperature for 20 minutes. The slides were then rinsed with dH₂O. Finally, the slides were submerged sequentially in to 70%, 90% and 100% ethanol, for 10 minutes each.

The slides were visualised with a fluorescent microscope Leica DM4000 and images analysed with ImageJ and violin plots were created with RStudio.

qPCR

Sample preparation

Cells were seeded a 6 well plate at a density of 1×10^5 cells/ml. Doxycycline treatment given at concentrations 10 $\mu\text{g/ml}$, 1 $\mu\text{g/ml}$, and 0.1 $\mu\text{g/ml}$. Cells were incubated at 37 °C for 72 hours and were harvested at >70% confluency. Cells were pelleted and resuspended in lysis RLT buffer (QIAGEN).

RNA extraction

RNA was extracted from samples using RNeasy® Maxi kit (QIAGEN) and eluted in RNase free water. Sample concentrations were measured via Nanodrop.

cDNA synthesis

RNA samples were diluted to 1 $\mu\text{g/}$ 10 μl in a 96 well plate.

To each well 10 μl of diluted RNA sample was added, with 1 μl DNase Enzyme (Thermo-Scientific), and 1 μl DNase Buffer (Invitrogen). The plate was then incubated at 37°C for 30 minutes.

1 μl of Ethylenediaminetetraacetic acid (EDTA) (Thermo-Scientific) was then added to each well. The plate was incubated at 65°C for 10 minutes and then at 4°C for 5 minutes.

For reverse transcription (RT) a RT master mix was made:

- 2 μl RT buffer (Applied-Biosystems)
- 0.8 μl dNTP (Applied-Biosystems)
- 2 μl Random primers (Applied-Biosystems)
- 1 μl Reverse Transcriptase (Applied-Biosystems)
- 4.2 μl RNase free H₂O

10 μl of the RT master mix was added to each well and the plate was put on a heating program consisting of:

- 25°C for 10 minutes
- 37 °C for 2 hours
- 85 °C for 5 minutes
- 4 °C for 5 minutes

Real-time polymerase chain reaction (qPCR)

Gene expression was quantified by real-time reverse transcriptase polymerase chain reaction (qPCR) using TaqMan technology (Applied biosystems). For each target gene 3 technical replicates were used in a 96 well plate (Applied-Biosystems).

A master mix was made where quantities reflect amount per well (96 well plate):

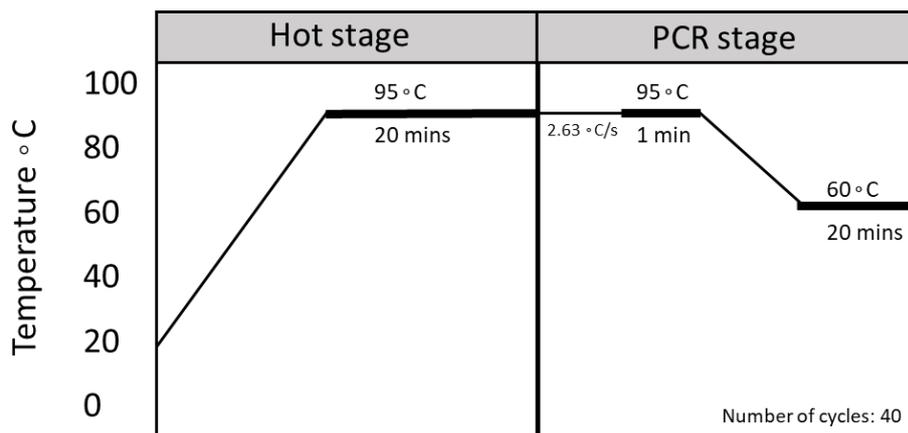
- 5 μ l TaqMan™ Fast Universal PCR Master Mix (Applied Biosystems™, US)
- 3.5 μ l RNAase-free water
- 0.5 μ l probe

To each well 9 μ l of the master mix was added, with 1 μ l of the desired probe.

TaqMan probes used:

- POLD3
- CHRDL2

The reaction was run on Quant Studio 7 Flex Real-Time PCR machine (Thermo Fisher Scientific) with the cycling program shown below:



Quant studio 7 Flex Real-time PCR cycling program used for qPCR

Telomere Length analysis

To assess telomere length, real-time PCR was used. Relative telomere length was compared to that of a single copy gene (36B4) to control for amplification of each sample, and to determine genome copies per sample. Genomic DNA was extracted from our POLD3 knockdown cell lines (DNeasy, QIAGEN) treated with doxycycline at 10 µg/ml, 1 µg/ml, and 0.1 µg/ml or DMSO. Telomere length analysis was performed as described [67]. Briefly, two q-PCR master mixes were prepared with either the copy gene primer (36B4) or the telomere primer. Master mixes prepared as followed: 10 µl 2X Power SYBR Green PCR Master Mix (Applied Biosystems), 2 µM forward primer (Telomere-F or 36B4-f), 2 µM reverse primer (Telomere-R or B6B4-R), 4 µl of 5ng/µl DNA, and nuclease free water to 20 µl. Each sample was prepared in triplicate. Primers were used as described here [68].

A standard curve was generated by S. Adulami (unpublished) using serial dilutions of 36B4 standard (6125000 kb to 6.125 kb). Plasmid DNA (pBR322) was also added to each standard to maintain a constant 20 ng of total DNA per reaction tube. A telomere standard curve was established by serial dilution of the telomere standard (1018400 kb to 10184 kb) and was used to measure content of telomeric sequence per sample. Real-time PCR runs were performed in triplicate for each of the DNA pools.

C-circle analysis

To assess the level of c-circle present, real time PCR was used. C-circle levels were determined as a % of total c-circles present in cell line U2OS. The method for C-circle amplification was a previously described [192].

Briefly, genomic DNA was extracted from our POLD3 knockdown cell lines (DNeasy, QIAGEN) treated with doxycycline at 10 µg/ml, or DMSO. Rolling cycle of C-circle amplification was performed using 40 ng total genomic DNA and made up to 10 µl with 10 mM Tris. To diluted DNA 10 µl of master mix was added: 0.8 µl 10 M DTT, 2 µl Phi 29 DNA pol buffer (NEB, UK), 0.4 µl 10 mg/ml BSA, 0.2 µl 10% tween, 0.8 µl 1 mM dTTP, 1.5 µl phi 29 DNA pol (NEB), 4.3 µl water. Separate samples were also made without phi 29 DNA pol as a control. Samples were run for 8 hours at 30 °C and deactivated at 65 °C for 20 minutes. qPCR on amplified C-circle DNA was then performed as described before in telomere length analysis.

Primers:

Oligomer name	Sequence
Telomere standard (human)	(TTAGGG) ₁₄
36B4 standard (human)	CAGCAAGTGGGAAGGTGTAATCCGTCTCCACAGACAAGGCCA GGACTCGTTTGTACCCGTTGATGATAGAATGGG
Telomere-F (human)	CGGTTTGTGGTTTGGGTTTGGGTTTGGGTTTGGGTTTGGGTTT
Telomere-R (human)	GGCTTGCCTTACCCTTACCCTTACCCTTACCCTTACCCTTACCCT
36B4-F (human)	CAGCAAGTGGGAAGGTGTAATCC
36B4-R (human)	CCCATTCTATCATCAACGGGTACAA

Statistical analysis

Statistical analysis was carried out using GraphPad Prism. Shapiro-Wilk test was used to determine whether data sets followed a uniform distribution. Student's t-test was used to compare two means from normal distributed data. TWO-way ANOVA (Analysis of Variance) was used to compare the means of two or more unmatched groups that were normally distributed.

Bioinformatics

Analysis of RNA-seq data

Samples for RNA-seq analysis were prepared by growing cells from culture in standard media conditions with overexpression of the *CHRD2* gene through doxycycline-inducible expression. Doxycycline was given in concentrations 10 µg/ml, 1 µg/ml, and 0.1 µg/ml.

RNA-seq was performed by the Oxford Genomic centre. Data for bioinformatics analysis was given in the format of fastq raw reads. Data was analysed using the open-source software package Tuxedo Suite. Tophat2 and Bowtie2 were used to map paired end reads to the reference *Homo sapiens* genome build GRCh38. GENCODE38 was used as the reference human genome annotation. Use of Tophat suite source paper can be found here: Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 2013 Apr 25;14(4):R36. doi: 10.1186/gb-2013-14-4-r36. PMID: 23618408; PMCID: PMC4053844.

Aligned reads were filtered for quality using Samtools with the minimum selection threshold of 30. Transcripts assembly and quantification was done through Cufflinks, and differential expression analysis was achieved using Cuffdiff software. Differential expression was expressed in the form of \log_2 fold change between sample and control.

Data visualisation and R

Data was cleaned and significant data was extracted using R software. Graphs we generated using R studio 4.1.0 using libraries ggplot2 and heatmap2

GSEA

Gene-set-enrichment analysis was performed using the GSEA software 4.2.3. The Chip annotation platform used was Human_Ensembl_Transcript_ID_MSigDB.v7.5.1.chip.

Gene sets used:

- c6.all.v7.5.1.symbols.gmt
- h.all.v7.5.1.symbols.gmt
- GOBP_REGULATION_OF_BMP_SIGNALING_PATHWAY
- enplot_REACTOME_PI3K_AKT_SIGNALING_IN_CANCER_13
- enplot_GOMF_BMP_RECEPTOR_BINDING_58
- WP_NRF2_PATHWAY.v2023.1.Hs.gm

Code for RNAseq analysis can be found in Appendix 1.

Chapter 3: Effects of CHRDL2 overexpression on 2D CRC cell lines and 3D models

Introduction

To investigate the role of CHRDL2 in CRC, we aimed to study the functional effects of modifying CHRDL2 expression in 2D cell lines and 3D organoid models. We sought to achieve this through a variety of assays to measure proliferation, migration, clonogenic potential, and differentiation.

CHRDL2 is a BMP antagonist, which, through binding to BMP receptors, prevents BMP signalling. Canonically, BMP signalling within the intestine promotes differentiation, and a reduction in proliferation [193]. WNT signalling promotes stem-cell renewal and proliferation within the stem-cell compartment of the crypt and works in a counter gradient to BMP signalling [194]. Therefore, we hypothesised that through BMP signalling inhibition, CHRDL2 would promote a more stem-cell phenotype in CRC cells, thus causing an increase in cellular proliferation, migration, clonogenicity and stem-cell pathways through increased WNT signalling.

Previous studies have supported the hypothesis that CHRDL2 would increase proliferation, with Sun. et al, [189] showing that inhibition of CHRDL2 through an shRNA method reduced proliferation by approximately 50% compared to that of an empty vector in HCT116 cells, and that stable overexpression of CHRDL2 in HCT8 cells caused an increase in proliferation. This was also repeated using a colony formation assay in soft agar, which showed CHRDL2 overexpression increased cellular proliferation [189]. A study by Chen (2021) has shown similar results in osteosarcoma cells, where CHRDL2 repression through stable transfection of shRNA showed a decrease in proliferation, colony formation, and migration [190]. This role of CHRDL2 was also later supported by Wang (2022) who silenced CHRDL2 expression in gastric cell lines, leading to decreased proliferation, with equivalent increases in proliferation and colony formation following CHRDL2 overexpression [147].

To achieve our own investigation of the effects of CHRDL2 on CRC, we developed an inducible tet-on system which allows us to overexpress CHRDL2 in epithelial cancer cells through the addition of doxycycline. A tet-on system was utilised in order to create cell lines in which CHRDL2 could be overexpressed in varying levels. Furthermore, tet-on systems significantly improve the safety of gene therapy approaches. The tet-on system was developed using endogenous mechanisms found in bacteria, which produce antibiotic resistance. Tetracycline, or orthologues such as doxycycline, binds to the rtTA transcription factor and allows binding

to the promoter of the controlled gene [195]. This also allows temporal control of *CHRD2* overexpression that can be turned on and off. Moreover, our Tet-on system allows us to make comparisons on the effects of *CHRD2* overexpression within isogenic pairs of cell lines, which has not been previously attempted.

Doxycycline was given to cells to give final concentrations of 0.1 ug/ml, 1 ug/ml and 10 ug/ml to induce *CHRD2* expression, which were termed *CHRD2*, *CHRD2+*, and *CHRD2 ++* respectively. For this system we used immortalised epithelial cell lines originally derived from patients with colorectal adenocarcinoma. Epithelial cells were chosen for our model as the majority of CRC arises from mutations in the colonic epithelium [196]. Additionally, epithelial cells have little to no intrinsic *CHRD2* expression. In the intestinal tract BMP repressors are predominantly expressed by fibroblasts in the mesenchyme, which signal in a paracrine fashion to epithelial cells, making our addition of *CHRD2* to epithelial cells a more faithful model to the *in vivo* system. Furthermore, epithelial cells, despite being differentiated cells, retain the ability to gain stem-like characteristics. This enables us to test whether *CHRD2* increases stem-ness through hyperactivity of the WNT pathway when BMP signalling is inhibited. .

However, 2D cell line systems have their limitations, so we also developed 3D organoid models. Derived from crypts of mouse small intestine, our organoids were grown in a 3D scaffold. This allowed organoids to differentiate into different cell-types and offered an insight into the role of *CHRD2* during this differentiation process. It is well documented that BMP expression in intestinal epithelial cells inhibits stem-cell self-renewal and restricts stem-cell phenotypes evidenced by a reduction stem cell marker's such as LGR5 [197][198]. Therefore, we sought to treat a 3D organoid model with extrinsic *CHRD2* to repress BMP signalling, in hope we would see an increase in stem cell markers, and reduction in differentiation.

With these models we aimed to elucidate the functional effects of *CHRD2* overexpression in CRC.

Establishing CHRDL2 cell lines.

Previous studies have highlighted CHRDL2 to be over-expressed in many CRC cases and have shown that overexpression can lead to increased migration and clonogenicity in cancer cell lines [189][190]. To further study these effects of overexpression of CHRDL2 in CRC, we established CHRDL2 overexpression in 2D cancer cell lines. Firstly, the mRNA expression of *CHRDL2* in common CRC cell lines was analysed. There was found to be moderate expression of CHRDL2 in CACO2 and LS180 cells lines, and little to none in COLO320 and RKO cell lines (Figure 1). BMP levels in our experimental cell lines were also analysed, to gain insight to effect CHRDL2 would inflict as a BMP antagonist. Moderate levels of BMP1 and 4 were found in all tested cell lines. BMP2 was found highly expressed in CACO2, with low expression in COLO320 and LS180, and little to no expression in RKO. BMP5 and 6 had low expression in our cell lines, with little to none in RKO cells.

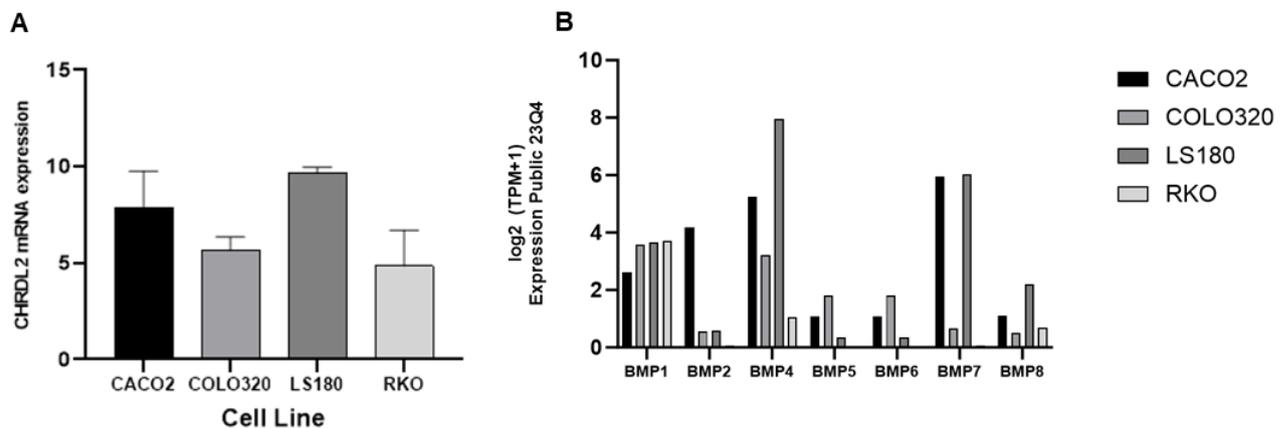


Figure 1: A) qPCR of relative CHRDL2 mRNA expression in CRC cell lines. N=3. B) Levels of endogenous BMP proteins in experimental cell lines given as log₂ (TPM + 1). Data obtained from the Cancer cell line encyclopaedia (CCLE). Error bars given as \pm SEM.

Colorectal adenocarcinoma cell lines CACO2, COLO320, LS180 and RKO cell lines were selected for lentiviral transduction of *CHRD2* overexpression. These cell lines were selected due to variation of levels of wild-type *CHRD2* mRNA expression (Figure 1), and the various mutations within the BMP pathway they exhibited (Figure 2). All cell lines had hyperactivity of the WNT signalling pathway as evidenced by mutations in the APC genes, which are found in ~80% of CRC tumours [199]. Furthermore, CACO2 and LS180 cell lines had mutations in the *CTNNB1* genes which encodes for β -catenin, a major component of the WNT signalling pathway. The LS180 and RKO cell lines have damaging mutations in the *BMP2* gene, a crucial receptor for canonical BMP signalling. Through these combinations of mutations in our pathway of interest we were able to elucidate any differences in the observational effects of *CHRD2* overexpression in these cell lines.

Cell Line	APC mutations	BMP receptor mutations	β -catenin mutations	WNT mutations	Other relevant mutations
CACO2	APC- damaging		CTNNB1- damaging		TP53- damaging SMAD4- damaging
COLO320	APC- damaging				TP53- damaging
LS180	APC/APC2 – non-conserving	BMP2- damaging	CTNNB1- damaging		PTEN- non- conserving
RKO	APC2- non- conserving	BMP2- damaging BMP1A- non-		WNT10a- non- conserving	

Figure 2: Table of cancer cell lines used and mutations. Acquired from the broad institute DepMap portal.

Full length cDNA of *CHRD2* was cloned into the tet-responsive expression vector pCW57.1, packaged into lentivirus and transduced into our experimental cell lines. To ensure overexpression of *CHRD2*, qPCR and Western blotting was performed. As seen in Figure 3, A, each cell line transduced with lentiviral overexpression of *CHRD2* showed an increase in *CHRD2* mRNA levels with addition of doxycycline. Doxycycline was given in 3 concentrations to cell lines: 0.1 μ g/ml (*CHRD2*), 1 μ g/ml (*CHRD2+*) or 10 μ g/ml (*CHRD2++*) (Figure 3, B). The highest levels of doxycycline treated cell lines (*CHRD2++*) showed significantly increased *CHRD2* mRNA expression ($p < 0.05$) in all cell lines. Next, western blotting was performed to validate protein expression (Figure 3,C). In each of our treated cell lines *CHRD2* protein expression increased with addition of doxycycline (Figure 3, C, D). COLO320 and LS180 cell lines showed highest *CHRD2* protein levels. Faint bands of *CHRD2* in the control lane for CACO2 can be observed, which appears to reflect the qPCR data (Figure 1).

We also validated the assumption of CHRDL2 acting as a secretory protein through western blotting of conditioned media of CHRDL2-expressing cells. Previous studies have reported that CHRDL2 is a secreted protein which signals through paracrine methods to neighbouring cells. To performed western blotting of secreted CHRDL2 protein, RKO cells were grown in standard media conditions treated with a final concentration of 10µg/ml doxycycline. The conditioned media was then harvested and diluted in lysis buffer to give dilution ratios of 1:25, 1:50 or 1:100. As seen in Figure 3 C, CHRDL2 protein was observed in the media of treated cells which increased with concentration and was un-detectable in our untreated control cell line. Secreted protein levels were controlled for by ponceau stain (Figure 3 E). Quantification can be observed in Figure 3 F.

To ensure BMP antagonism was present due to CHRDL2 overexpression, western blotting was performed for SMAD 1/5 phosphorylation, a direct downstream transducer of BMP signalling. Total SMAD1 was used to control for varying expression levels. SMAD 1/5 phosphorylation clearly showed a reduction in LS180, RKO, CACO2 and COLO320 cell lines (Figure 3, G), with quantification supporting this observation (Figure 3, H). RKO cells showed the highest reduction in SMAD 1/5 phosphorylation at nearly 50% reduction in CHRDL2 + treated cells. This data confirmed CHRDL2's role as a BMP antagonist in our CRC cell lines.

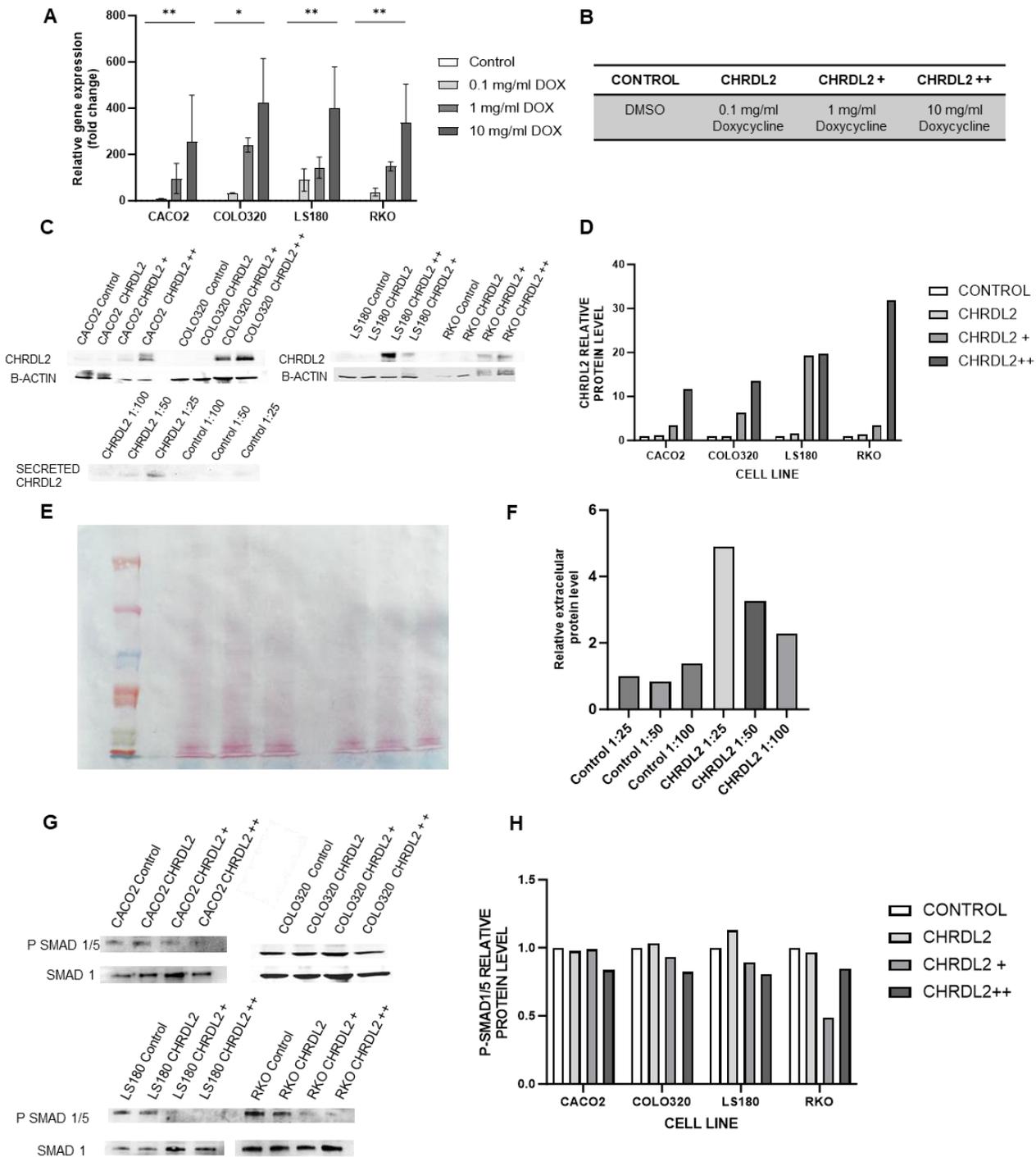


Figure 3: A) qPCR of mRNA levels of CHRDL2 expressed as fold change in 4 experimental cell lines. Cell lines were grown with doxycycline at: 0.1 µg/ml, 1 µg/ml or 10 µg/ml to induce expression. RKO DMSO-10µg/ml $p < 0.01$, COLO320 DMSO-10µg/ml $p < 0.05$, CACO2 DMSO-10µg/ml $p < 0.01$, LS180 DMSO-10µg/ml $p < 0.01$. N=3. B) Table of doxycycline treatment used in figure B. C) Western blotting of corresponding protein levels of CHRDL2 in cell lines with lentiviral overexpression, and secreted CHRDL2 present in cell culture media. Shown as ratios of dilution with loading buffer. D) Quantification of CHRDL2 protein levels as measure by western blot using Image J software. E) Ponceau stain of media from CHRDL2 cell lines for secreted CHRDL2 protein analysis. F) quantification of secreted CHRDL2 by western blotting. G) Western blotting of SMAD1/5 phosphorylation in cell lines overexpressing CHRDL2. H) Quantification of P-SMAD1/5 protein levels in CRC cells with CHRDL2 overexpression. Error bars given as \pm SEM.

CHRD2 overexpression increases WNT signalling through β -catenin nuclear localisation

Next, we investigated whether CHRD2 inhibition of BMP signalling would cause the upregulation of WNT signalling. As described in diagram 2, WNT activation of G-coupled frizzled receptors leads to the sequestering of the β -catenin destruction complex, allowing β -catenin to enter the nucleus. Therefore, increased WNT signalling in our cell lines was shown by enhancement of Beta-catenin staining in the nuclei over cytoplasmic staining. The ratio on Beta-catenin nuclear localisation was found by measuring the fluorescence of the nuclei of each cells, divided by the average fluorescence of the cytoplasm. As seen in figure 4 A, B, CHRD2 overexpression increased Beta-catenin nuclear-localisation, a hallmark of WNT signalling ($P < 0.005$, Figure 4 B), confirming the hypothesis that CHRD2 increases WNT signalling through BMP inhibition.

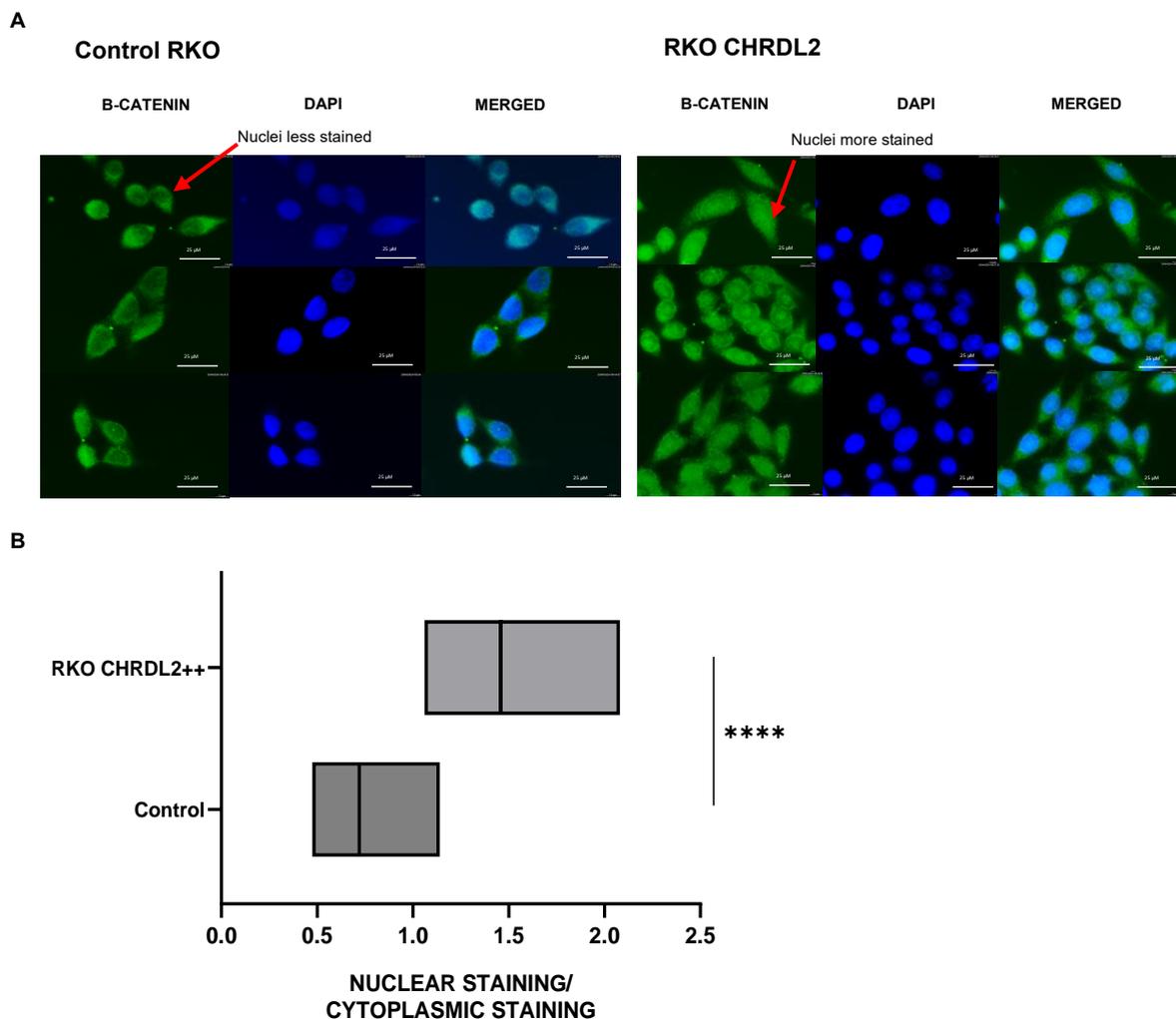


Figure 4: A) Immunofluorescence staining Beta-catenin on RKO cells with 10 $\mu\text{g/ml}$ doxycycline to induce CHRD2 expression. N=3 B) Quantification of Beta-catenin nuclear staining over cytoplasmic staining on RKO cells with 10 $\mu\text{g/ml}$ doxycycline to induce CHRD2 expression N=3. Quantification found using CTFC of nucleus only. Error bars given as \pm

Overexpression of CHRDL2 decreases cellular proliferation and slows cell cycle progression.

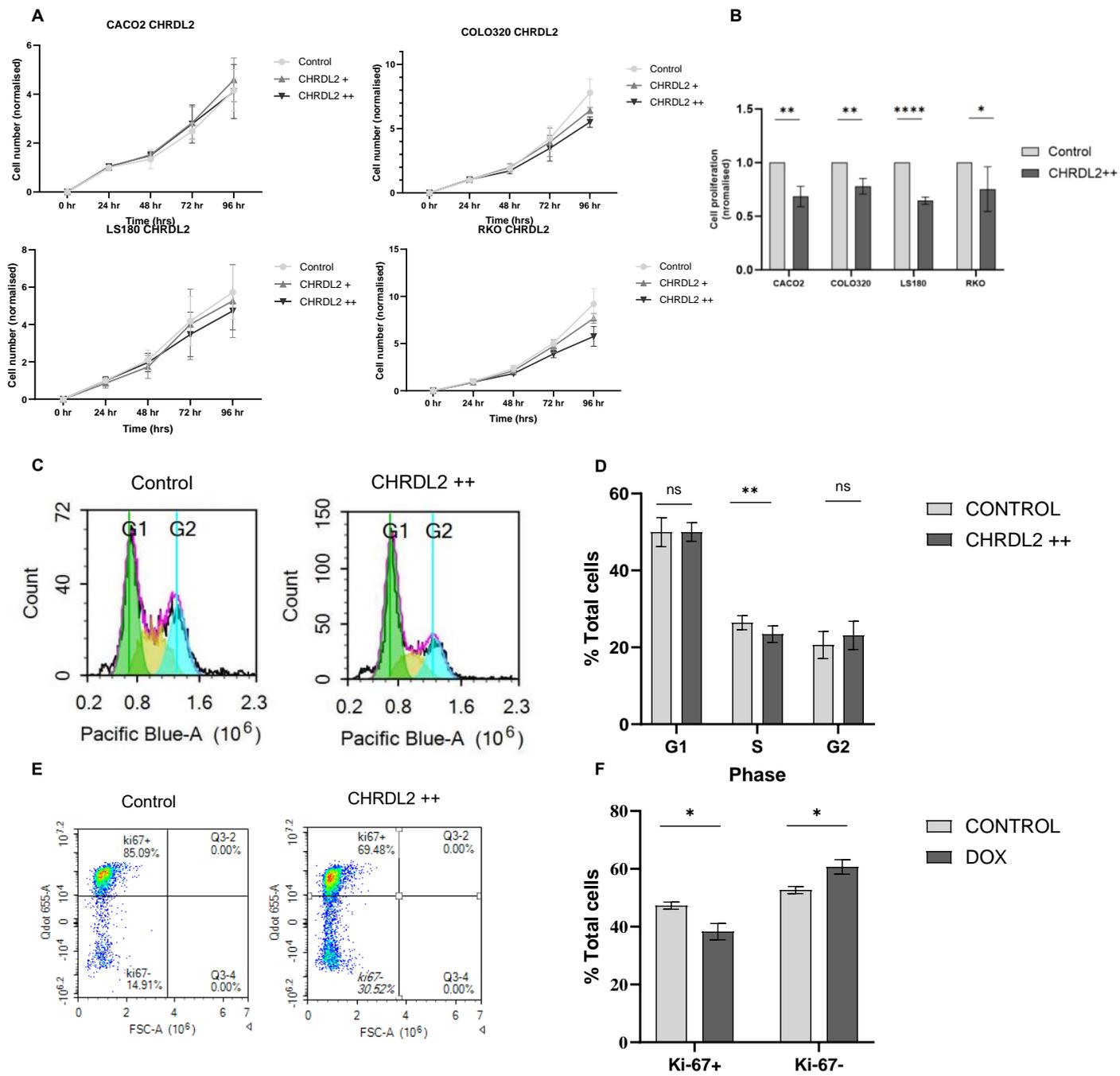
Aberrant proliferation is a hallmark of cancer cells and is often also linked to stem cells with increased cellular proliferation causing higher tumour growth. Furthermore, cancer stem cells are often linked to rapid proliferation through unlimited self-renewal. Therefore, we wished to measure the effects of CHRDL2 overexpression on cellular proliferation in our experimental cell lines. To do this, cells were seeded at a density of 100,000 cells/ml in a 96 well plate and treated with doxycycline at final concentrations of 0.1 ug/ml (CHRDL2), 1 ug/ml (CHRDL2+), and 10 ug/ml (CHRDL2++) to induce CHRDL2 expression. Cell growth was measured at 24, 48, 72, and 96hrs by colorimetric metabolic assay (MTS assay). As seen in figure 5 A, cell growth was slightly reduced during overexpression of CHRDL2 in COLO320 and RKO cell lines ($P < 0.01$, $P < 0.036$, t-test).

Intriguingly, this unexpected effect was enhanced when cells were cultured under low glucose conditions. Low glucose DMEM at 1g/ml was given compared to standard 4.5g/l. Final cell numbers after 72hrs of CHRDL2++ treatment was assessed by MTS assay under standard and low glucose conditions (Figure 5 B). High levels of CHRDL2 (++) expression reduced proliferation in all our cell lines by approximately 25%: CACO2 ($p < 0.01$), LS180 ($p < 0.001$), RKO ($p < 0.05$), and COLO320 ($p < 0.01$) (Figure 5 B).

Many studies have shown that glucose levels are altered within the tumour microenvironment. Cancer stem-cells show a preference for metabolism via aerobic glycolysis, also termed the Warburg effect, and require increased glucose and glutamine levels compared to normal cells in order to sustain energy requirements for the formation of tumours and long-term cell survival [200]. It is due to this increased glucose demand that there is often a deficiency in glucose within the tumour microenvironment [201]. Our observed sensitivity to glucose reduction could be a marker of enhanced stem-cell properties of cancer cells under CHRDL2 overexpression. It should be noted that glucose-depletion dramatically alters Cancer stem cells (CSC) gene expression and cellular behaviour, compared to non-stem cancer cells, which allows them to reduce reactive oxygen species and survive for longer periods of time [202][203][204].

To determine the cause of slower growth of CHRDL2 expressing cells, we performed a series of assays to measure cell cycle status. As seen in figure 5, C, Flow cytometry analysis showed that CHRDL2 overexpression decreased the proportion of cells in S phase and increased that in G2 phase, shown by reduction of the S phase bridge between G1 and G2 peaks. Quantification of this can be observed in figure 5D, which clearly shows an increase in the portion of cells in S phase, again suggesting a decreased rate of proliferation. Ki-67, a marker of proliferating cells, was also used to determine the number of cells in S, G2 and M phases

of the cell cycle, in which Ki-67 is accumulated before degradation in G1 and G0. As seen in figure 5 E, CHRDL2⁺⁺ overexpression decreased the number of Ki-67 positive cells and increased the proportion of Ki-67 negative cells. This suggests that CHRDL2⁺⁺ decreases the number of actively proliferating cells, enhancing the number of cells present in either G0 or G1 of the cell cycle, which would indicate an increase in quiescence or slow cycle cells respectively. These findings were confirmed by immunofluorescent staining of Ki67 in COLO320 cells with CHRDL2 overexpression, which clearly shows a reduction in Ki67 compared to a control ($P < 0.001$, Figure 5 G, H). Overall, contrary to our hypothesis, CHRDL2 overexpression appears to reduce proliferation, however our observation that CHRDL2 cells are more dependent on high levels of glucose, suggests a more complex phenotype. The observation of fewer proliferating cells, marked by Ki-67, would suggest CHRDL2 overexpression either increases quiescence of cancer cells, or a switch to a slower-cycling cell type, which is potentially a marker of increased stem-ness.



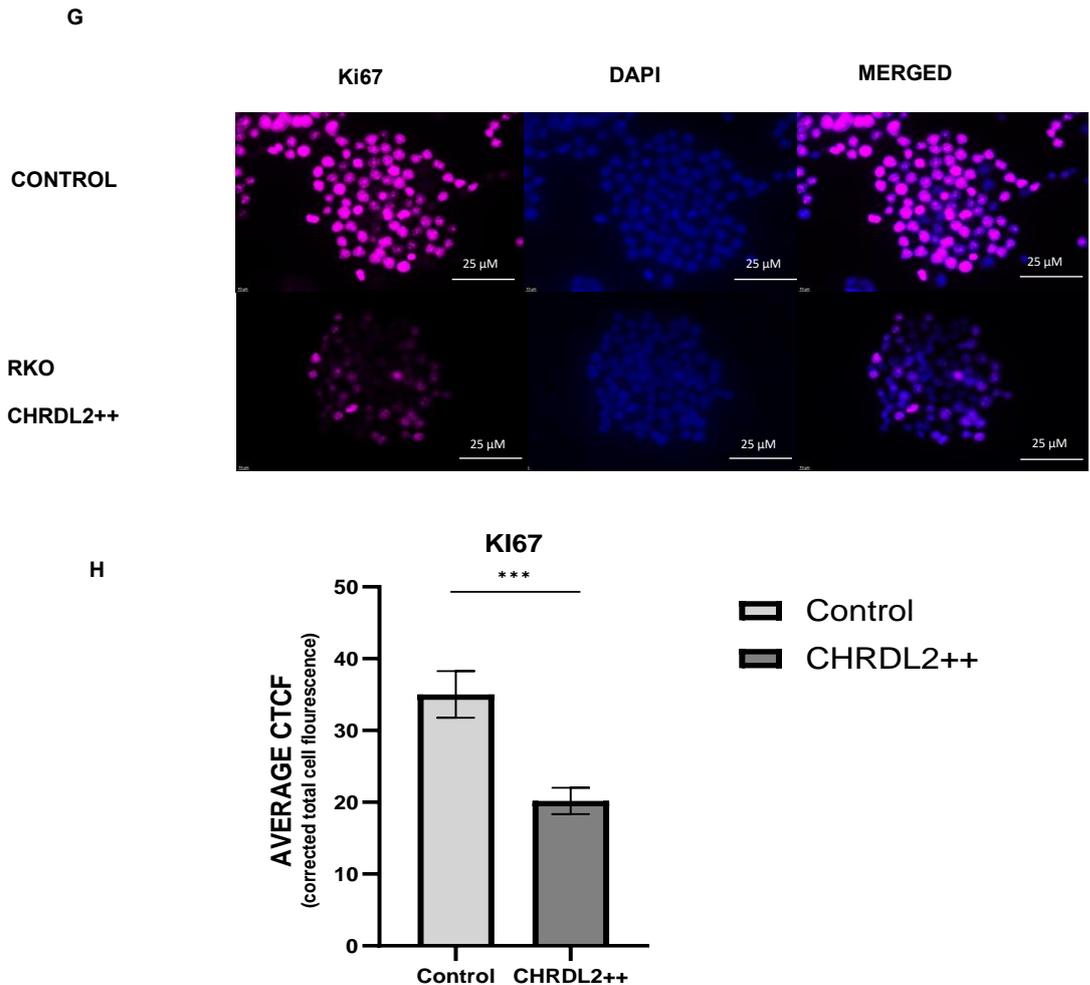


Figure 5: A) MTT assay of cellular proliferation of CHRDL2 cell lines. In COLO320 cells, two-way ANOVA between Control and CHRDL2++ was $P < 0.0118$. LS180 cells two-way ANOVA between Control and CHRDL2= $P < 0.0107$. Control and CHRDL2++ $P < 0.0114$. RKO cells Control and CHRDL2++ $P < 0.0181$. One-way ANOVA at 96hr was also performed: CACO2 $P < 0.44$, COLO320 $P < 0.0411$, LS180 $P < 0.121$, RKO $P < 0.0476$. $N = 3$. B) Cellular proliferation analysis on cells grown in low glucose conditions (1g/l), given 10 $\mu\text{g/ml}$ doxycycline overexpression of CHRDL2. In CACO2, COLO320, LS180 and RKO cells lines, proliferation was significantly decreased with CHRDL2 expression. T-test; $P < 0.01$, $P < 0.01$, $P < 0.001$ and $P < 0.01$ respectively. C) Cell-cycle progression analysis of CHRDL2++ COLO320 cells compared to a control, by flow-cytometry. Histograms display cells stained with DNA stain Hoescht33342, with level of staining correlating to cellular DNA quantify. Green: G1, Yellow: S, Blue: G2. D) Quantification of Cell-cycle progression analysis of CHRDL2++ cells compared to a control, by flow-cytometry. S phase T-test $P < 0.01$. $N = 3$. E) Flow cytometry analysis of COLO320 cells treated CHRDL2 Overexpression. CHRDL2 decreased the number of Ki-67 + cells and increased the number of Ki-67 – cells. $N = 3$. F) Quantification of Ki67 status of COLO320 CHRDL2++ cells. T-test $P < 0.05$, $P < 0.05$. G) Immunofluorescence staining of proliferation marker Ki67 on COLO320 cells with CHRDL2++ overexpression. H) Quantification of Ki67 immunofluorescence staining on COLO320 cells with CHRDL2++ overexpression. $P < 0.001$. Error bars given as \pm SEM.

CHRDL2 decreases colony size and increases migration.

Metastasis is the spread of cancer from its initial origin to other organs around the body. The ability of cancer cells to gain metastatic competency relies on several factors, with cancer cells required to migrate to nearby tissues, invade and survive circulation, and form new colonies in foreign organs to form tumours [205].

The ability of a cell to clone itself to form new colonies is measured by its clonogenic potential. Within the tumour microenvironment, cells with high clonogenic potential usually have higher cellular proliferation and can reflect the potential of cancer to metastasize [206]. We therefore sought to measure the clonogenic potential of cells with CHRDL2 overexpression.

To achieve this, clonogenic assays were performed. Cells were diluted to 200 cells/ml, plated in a 6 well plate and allowed to grow for 1-2 weeks until visible colonies had formed. Cells were treated with DMSO, CHRDL2+ or CHRDL2++ doxycycline to induce CHRDL2 expression. Colonies were then fixed and stained with crystal violet before quantification (Figure 6, A). The results showed that CHRDL2 overexpression dramatically reduced colony formation in all 4 of our CRC cell lines (Figure 6, B). However, smaller colonies excluded from quantification due to sizing thresholds, were still visible in the CHRDL2 ++ groups, (Figure 6, A). This would suggest that CHRDL2 overexpression results in smaller colony sizes, which is linked to slower proliferative ability as demonstrated in figure 6.

Before cells can colonise new tissues, they must migrate to nearby tissues and invade the circulatory system. It is for this reason migration is a hallmark of cancer cells and is an important factor in allowing metastasis through mobilizing across the tissue microenvironment into the blood stream [206]. To assess the migratory potential of cells overexpressing CHRDL2, a scratch assay was performed. 100,000 cells were seeded in 6 well plates and grown until confluent, then a scratch was made down the middle of the well. Images were taken at 0hr and 24hr post scratch, and the difference in scratch wound size was measured to assess the migration rate (Figure 6, C). As you can see in figure 6D, migration was slightly reduced in CACO2 and RKO cells during CHRDL2 overexpression. No significant difference in migratory rate was found in COLO320 and LS180 cell lines although a trend for reduced rate was seen in both. However, as with the clonogenic assay, scratch assays are also reflection of cellular proliferation, as the confluence of cells that fill the scratch void may be due to proliferation of cells rather than migration of cells. So, it can be theorised that this reduction in migration is intrinsically linked to lower proliferation capabilities of cells with CHRDL2 overexpression.

Therefore, we needed a more robust method to analyse migratory potential of CHRDL2 overexpressing cells. To do this, we seeded 50,000 cells/ml in 500 μ l transwell inserts. Transwell inserts were then placed inside standard 12-well plates. Inserts contained a mesh bottom, with a pore size of 0.8 μ M, to allow migration of cells through the mesh membrane to the well below. Cells within the transwell inserts were suspended in media without serum, and the inserts were then placed in media with serum, to attract cells towards the bottom well. Cells that harboured migratory potential would pass through the membrane and colonise the bottom of the well, as seen in Figure 4, E. Interestingly, there was a dramatic increase in migratory potential of CHRDL2⁺⁺ cells compared to the control (Figure 2, F) ($P < 0.05$).

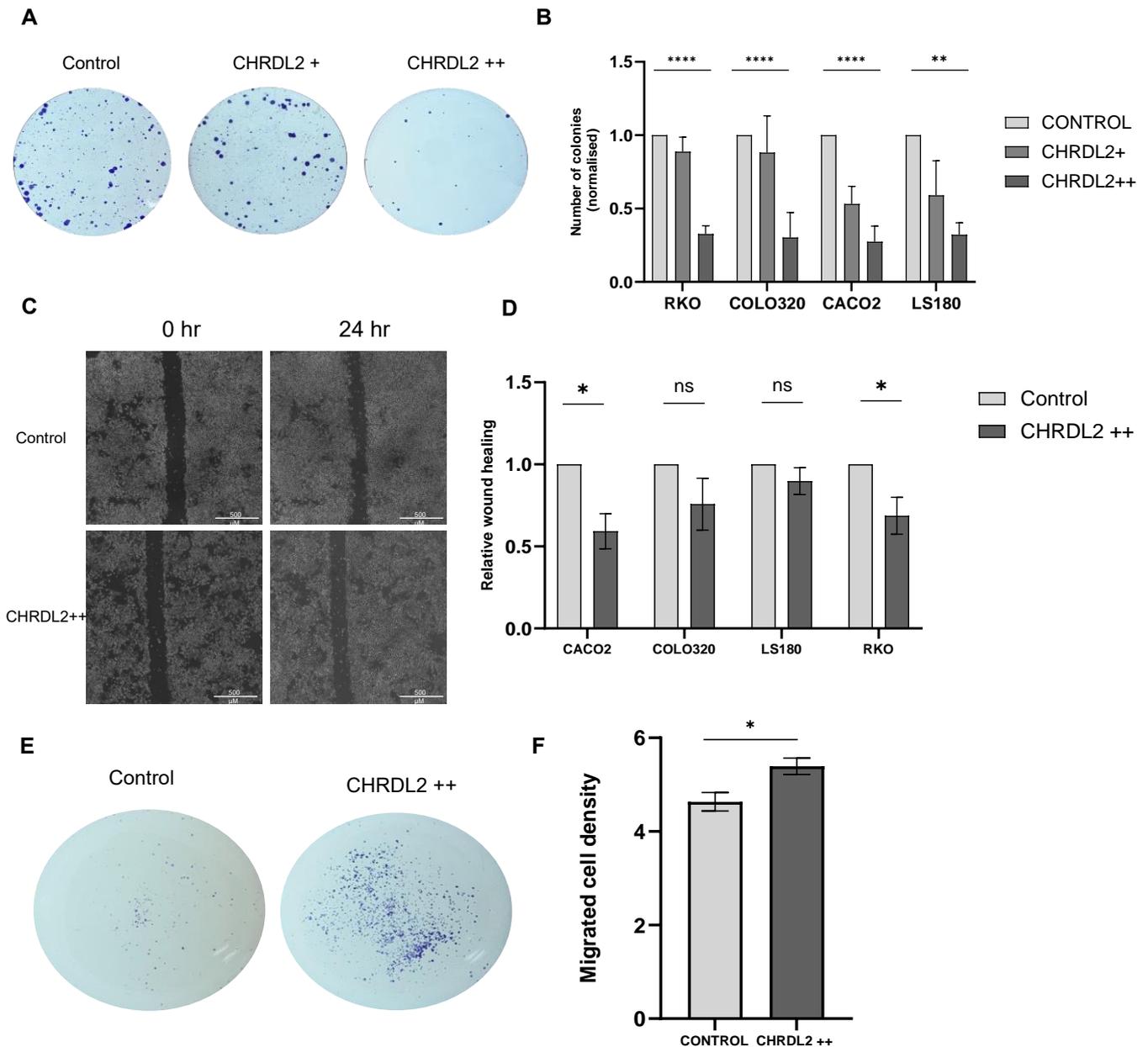


Figure 6: A) Crystal violet staining of colonies of RKO cells treated with CHRDL2+ or CHRDL2++ doxycycline to induce CHRDL2 expression B) Quantification by ImageJ of colonies established in our 4 experimental cell lines with CHRDL2 overexpression. CACO2 and RKO cell lines both showed reduced colony formation in the low and high CHRDL2 treated groups, $p < 0.0001$, $p < 0.0001$, $p < 0.0001$, $P < 0.01$. $N = 3$. C) Images of scratch assay performed on LS180 cell line with CHRDL2++ overexpression compared to a control. D) Quantification of scratch assay performed on CACO2, COLO320, LS180, and RKO cells lines. CACO2 and RKO showed significant reduction in migration rate ($p < 0.05$). $N = 3$. Error bars given as \pm SEM. E) Migrated COLO320 cells colonising bottom of transwell inserts. F) Quantification of COLO320 cells migratory potential. T test, $P < 0.05$. Error bars given as \pm SEM.

In order to confirm differences in proliferation and colony growth were due to CHRDL2 overexpression and not the effects of doxycycline to induce expression, parental cell lines without CHRDL2 lentiviral transduction were assessed. Cells were treated in the same manner as CHRDL2 overexpressing cells, treated with a DMSO control, or 0.1, 1, or 10 $\mu\text{g/ml}$ doxycycline. Proliferation assays were performed after 96hrs treatment. As seen in figure 7 A, no differences in proliferation were seen across all treatments, unlike what is seen in CHRDL2 overexpressing cells, where doxycycline induction of CHRDL2 overexpression significantly reduced proliferation (Figure 5 B). Similarly, there was no trend in reduction of colony growth (Figure 7 B) after doxycycline treatments in parental cells, unlike our CHRDL2 overexpressing cell lines (Figure 6 A). Therefore, we can confirm the results seen are due to the effects of CHRDL2, and not doxycycline.

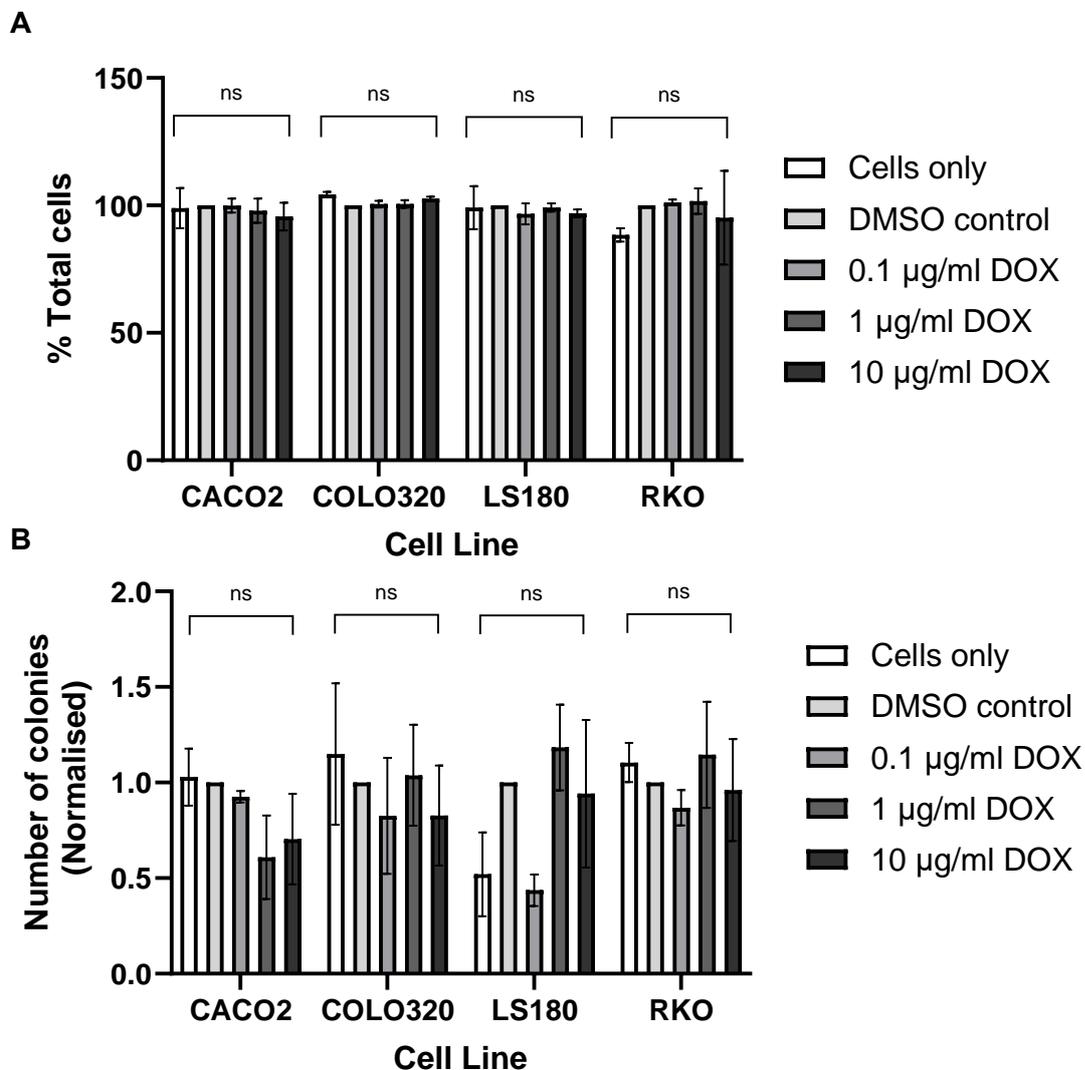


Figure 7: A) Cellular proliferation analysis on parental cell lines after 72hrs given doxycycline at 0.1 $\mu\text{g/ml}$, $\mu\text{g/ml}$, and 10 $\mu\text{g/ml}$. B) Clonogenic analysis on parental cell lines after 2 weeks given doxycycline at 0.1 $\mu\text{g/ml}$, $\mu\text{g/ml}$, and 10 $\mu\text{g/ml}$.

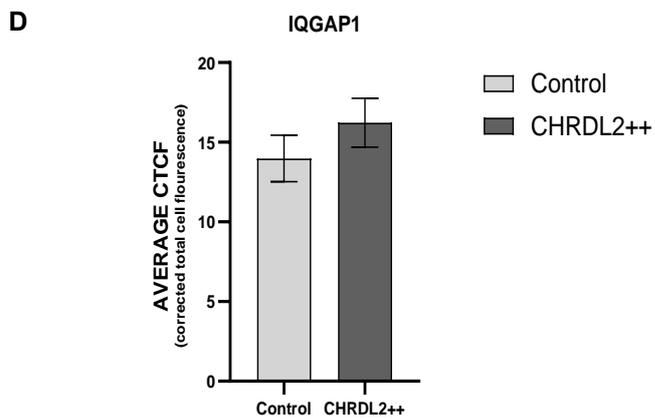
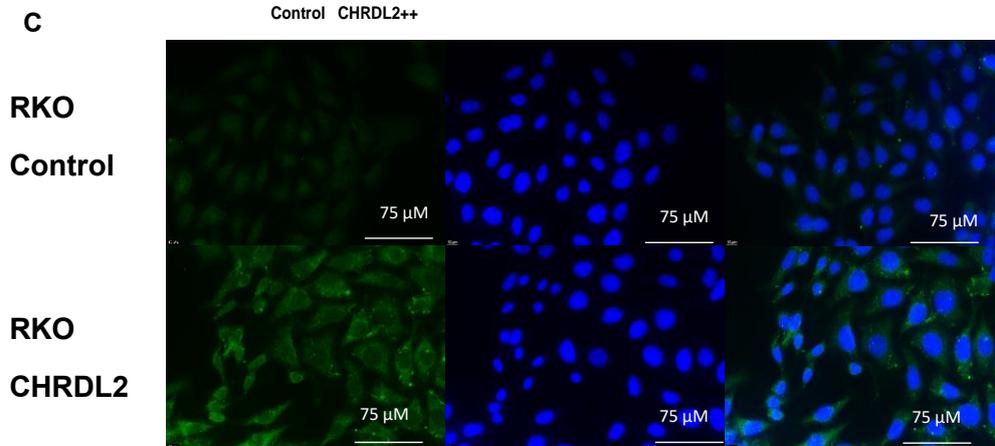
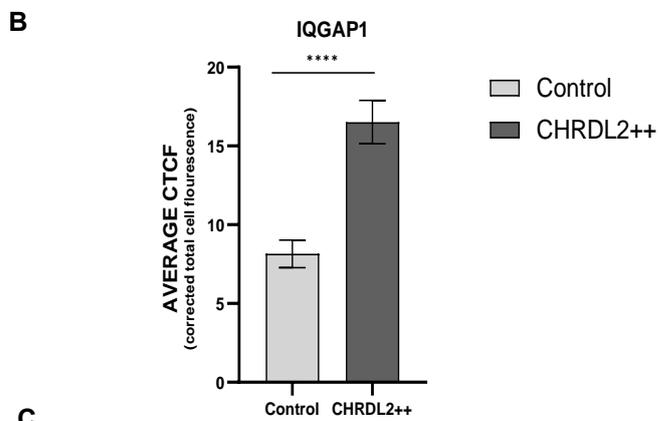
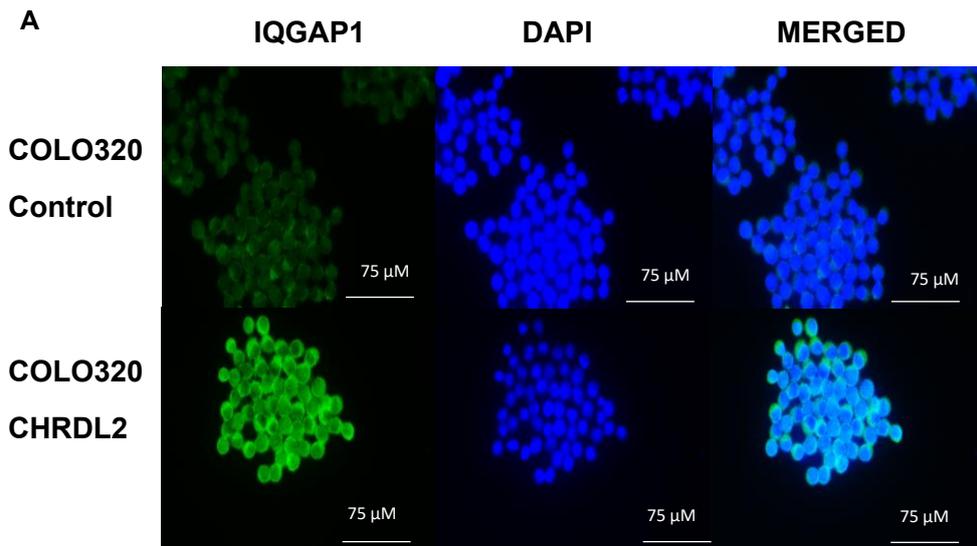
CHRDL2 increases migration and adherence markers.

Next, we looked at migration markers in COLO320 and RKO cells treated with CHRDL2++ compared to a control. First, we looked at the presence of IQGAP1, a marker of migration, invasion, and adhesion. IQGAP1 is oncogenic in many cancers, including CRC, and is a scaffold protein that helps in organisation of the actin cytoskeleton [207]. IQGAP1 has previously been shown to be expressed at the invasive front of CRC tumours[208], so proved a good marker to initially look at migratory ability.

As seen in Figure 8, A, IQGAP1 cytoplasmic expression was significantly higher in COLO320 cells overexpressing CHRDL2, which is supported by quantification in panel B, $P < 0.0001$. In Panel C, there also appeared to be an increase in IQGAP1 expression in RKO cells treated with CHRDL2, however this was not a large enough increase for statistical significance (Figure 8, D). This may be due to non-specific background staining in RKO cells or high endogenous IQGAP1 Expression in RKO cells. However, IQGAP1 foci are clearly visible, and in higher density in RKO CHRDL2 overexpressing cells. This could also be explained by the damaging mutations in BMP receptors harboured by RKO cells, making BMP inhibition through CHRDL2 overexpression less effective in these cell lines (Figure 2).

Next, we looked at expression of well-known epithelial cell adhesion molecule, EpCAM. EpCAM is a glycoprotein involved in cell cycle progression, proliferation, and migration [209]. Increased expression of EpCAM has been shown to correlate with tumour progression, and is present in more aggressive CRC patient tissue [210]. There was an obvious increase in EpCAM expression in COLO320 cells treated with CHRDL2++ compared to a control, as seen in figure 8 E, which is supported by quantification (Figure 8, F)($P < 0.01$).

Together, this data supports our previous findings in figure 6, of increased migratory potential. Cancer stem cells have long been shown to harbour increased migratory potential, so these findings seem to support our increased cancer stem-cell hypothesis during CHRDL2 overexpression, and may explain the increase for CHRDL2 overexpression in more aggressive CRCs that harbour this more migratory phenotype



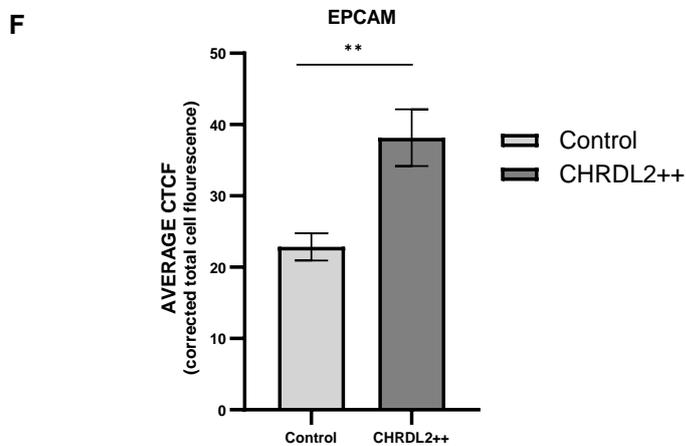
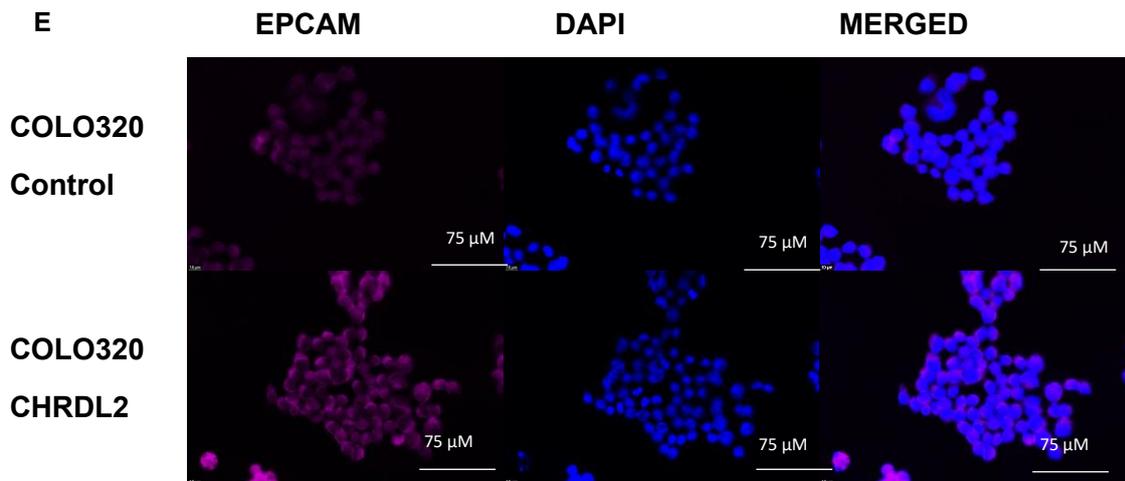


Figure 8: A) Immunofluorescence staining of IQGAP1 on COLO320 cells treated with CHRDL2++ compared to a DMSO control. B) Quantification of IQGAP1 in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). $P < 0.0001$ $N = 3$ T-test. C) Immunofluorescence staining of IQGAP1 on RKO cells treated with CHRDL2++ compared to a DMSO control. D) Quantification of IQGAP1 in RKO cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). $N = 3$ T-test. E) Immunofluorescence staining of EpCAM on COLO320 cells treated with CHRDL2++ compared to a DMSO control. F) Quantification of EpCAM in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). $P < 0.001$ $N = 3$ T-test. Error bars given as \pm SEM.

CHRD2 increases stem cell markers in 2D cell lines.

Next, to confirm increased stemness in our 2D cell lines, we performed qPCR analysis on unknown stem cell markers on our 2D CHRD2 overexpressing cell lines. Well known stem cell markers *LGR5* and *LGR6*, as well as B lymphoma Mo-MLV insertion region 1 homolog (*BMI1*) were chosen for this analysis. We found that *BMI1* was significantly upregulated in CACO2 CHRD2 overexpressing cells, and appeared, but not significantly, upregulated in COLO320 cells ($P < 0.05$) (Figure 9). We also found that *LGR6*, was also upregulated in CACO2 cells ($P < 0.05$). *LGR5*, the most commonly used marker for intestinal stem cells, was increased 2-fold in expression of COLO320 cells ($P < 0.01$).

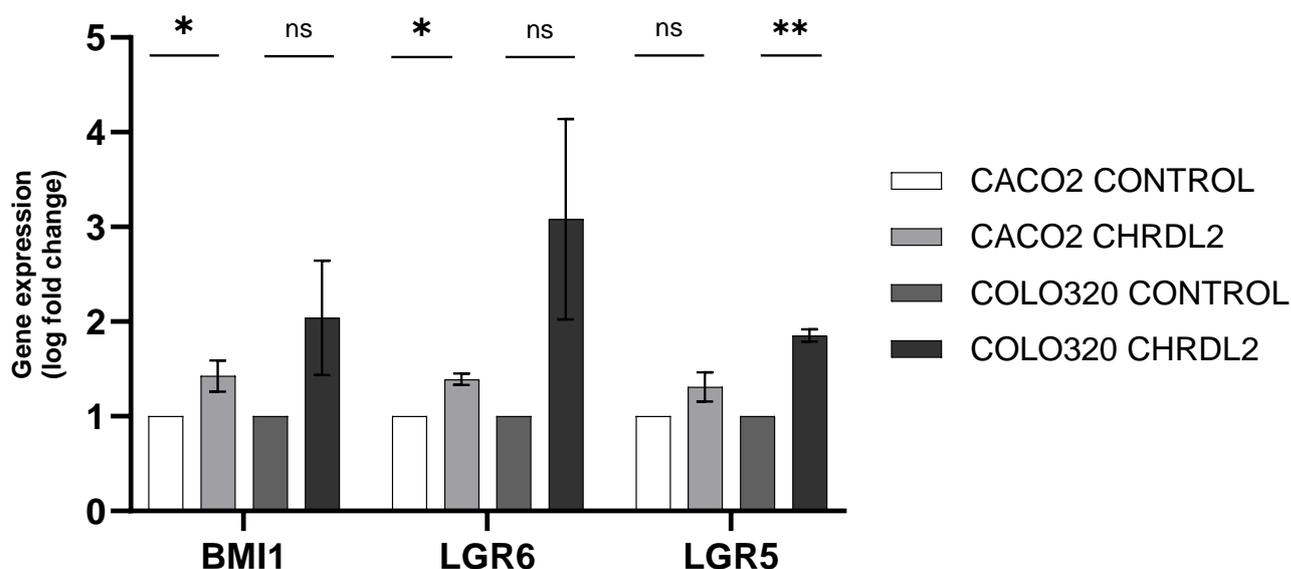


Figure 9: A) qPCR analysis of mRNA gene expression of known stem cell markers BMI1, LGR6, and LGR5 in cOLO320 and CACO2 cell lines with CHRD2⁺⁺ overexpression. CACO2 BMI1 $P < 0.05$. CACO2 LGR6 $P < 0.05$, COLO320 LGR5 $P < 0.01$. N=3 T-test. Error bars given as \pm SEM.

The effects of CHRDL2 on 3D organoid models

Traditionally *in vitro* studies on CRC have relied heavily on the use of 2-D cell culture systems, as we have used in our previous assays. Unfortunately, there are many limitations to 2D models, which lack the extracellular matrix and forces which effect cellular growth and characteristics. Therefore, we decided to study the effects of CHRDL2 in a 3D model.

There are many drawbacks to using 2D cell culture models to study CRC, for example, 2D models do not reflect the structure of intestinal shape nor that of a tumour, and therefore lack the biophysical cues seen *in vivo* [211]. Furthermore, cell morphology is altered during the transfer from tissue to 2D petri dish, and there is a loss of heterogeneity after multiple passages [212]. These factors can alter the structure inside the cell, as well as secretion and cell signalling [213]. Adherently grown cells have also been shown to lose polarity when grown in 2D [214]. Furthermore, commercially available immortalised cell lines that have been used in our previous assays are extensively mutated, so do not offer the best model for studying the initiation of cancer formation. These cell lines have often been in culture for decades and will therefore have gained many extra mutations since they were originally isolated, making them further distinct from *in vivo* conditions. Consequently, there has been a growing pressure to develop alternative 3D models to better understand cancer development *in vitro*.

First developed by the Hans Clever lab in 2009, organoid models had increased in popularity in recent years [215]. These organoids can be derived from mice or humans and are grown in a synthetic extracellular matrix scaffold to mimic *in vivo* growth. We utilised this approach and developed murine intestinal organoids to model the effects of CHRDL2 in a 3D environment. Organoids were given secreted forms of CHRDL2 that were harvested from conditioned media from our 2D cell lines, to mimic paracrine CHRDL2 signalling.

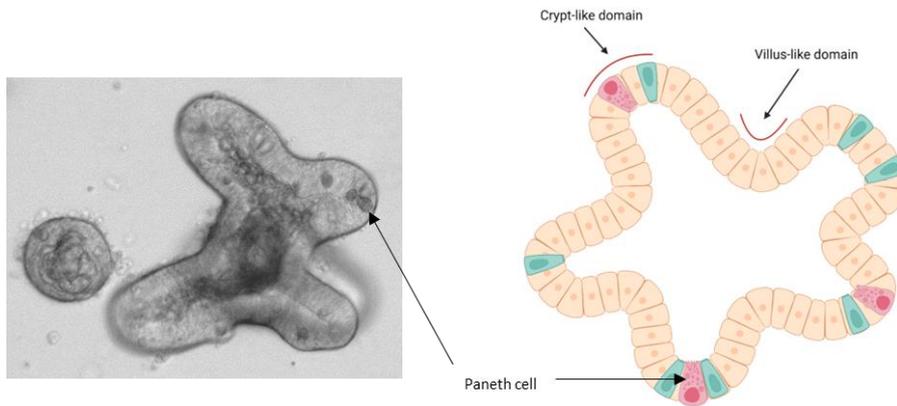


Diagram 7: Image of intestinal organoid derived from mouse small intestine and diagram depicting organoid structure. Invagination's in organoid display villus-like domains, whereas organoid "buds" display crypt-like domains with active stem cells [215].

Diagram 7, showcases the distinctive morphological features of intestinal organoids, characterized by a villi-bud-like configuration, with the outer epithelial layer forming distinct protrusions and invaginations. Prior investigations have extensively looked at gene expression of intestinal organoids and have shown epithelial cells within the "buds" of the organoids show crypt-like expression, and the evaginations villus-like expression. Furthermore, organoids show visible Paneth and goblet cells, with Paneth cells shown in the crypt-like domain of the organoids as granular-type cells (Diagram 5)[215].

For creation of CHRDL2 conditioned media, our CHRDL2-overexpressing cell lines were grown in serum free media and given doxycycline to induce expression. Media was collected after 48 hours, and standard organoid media was supplemented with 20% conditioned media, termed "CHRDL2 media". To control for generic secreted protein and doxycycline in our conditioned media, we added doxycycline to our parental cell lines in the same manner as we did to our CHRDL2 cell lines. Media from the treated parental cell lines was then given to organoids as the same with CHRDL2 media, termed "control media". Organoids that were given nonconditioned media were termed "control".

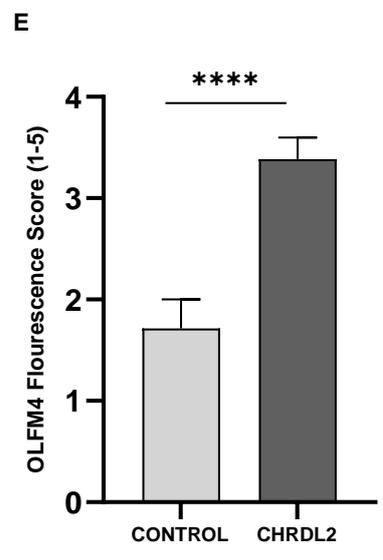
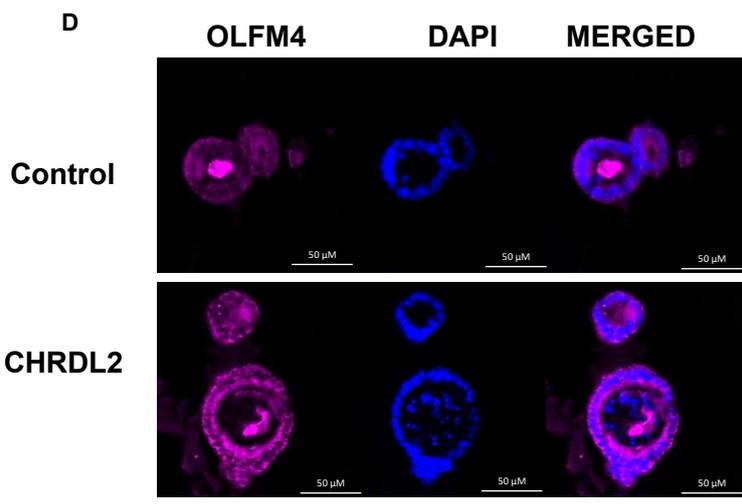
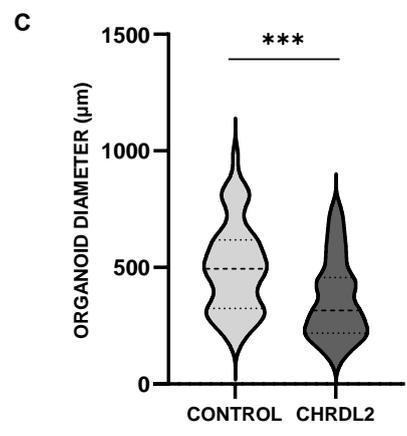
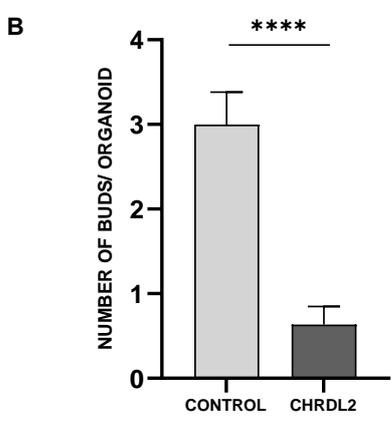
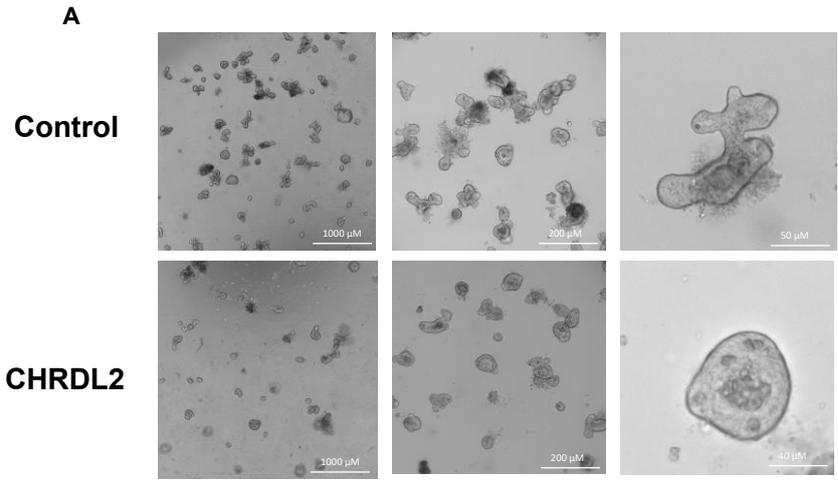
CHRDL2 addition affected the morphology and gene expression of murine intestinal organoids. As seen in figure 10 A, observational differences were noted in the number of villi "buds" on each organoid, as well as reduced overall organoid size. Organoids treated with the CHRDL2 conditioned media appeared to have a reduced number of differentiated buds ($P < 0.001$) (Figure 10 A and B), which may signify an increase in stem-like characteristics. Furthermore, organoids were significantly smaller in size, measured by average organoid

diameter ($P < 0.001$) (Figure 10, C). These effects are similar to that seen in our 2D cell lines, where cells have slower growth and increase stem-cell markers during CHRDL2 overexpression. This was corroborated by immunofluorescence staining of organoids which showed an increase in Olfactome-*din-4* (OLFM4), a marker for LGR5+ stem cells $P < 0.0001$ (Figure 10 E, F) [100]. OLFM4 was used as a substitute marker to LGR5, due to lack of suitable LGR5 antibodies.

Furthermore, we have shown an enhancement of WNT signalling through observed increase in β -catenin nuclear localisation. As seen in Figure 10 F, in control organoids β -catenin is localised to outer cellular membrane, (blue arrow), whereas upon extrinsic CHRDL2 treatment, β -catenin can be observed in the cytoplasm and nucleus of organoid cells (red arrow).

Moreover, as illustrated in Figure 10 G, through qPCR analysis we observed a significant increase compared to the controls ($P < 0.05$) in the expression of stem cell markers LGR5 (indicating crypt CBCs) and BMI1 (slow-cycling crypt stem cells). We also observed a slight increase in stem cell markers SOX9 and MSI1. These findings collectively suggest that exposing intestinal organoids to CHRDL2 diminishes differentiation and enhances stem cell numbers. No difference was found in the expression of vi

These data suggest that CHRDL2 addition in a 3D model increases stem-like characteristics, observed by decreased differentiation, and increased stem cell markers and enhanced WNT signalling. This supports our hypothesis that CHRDL2 increases stem-ness through BMP inhibition.



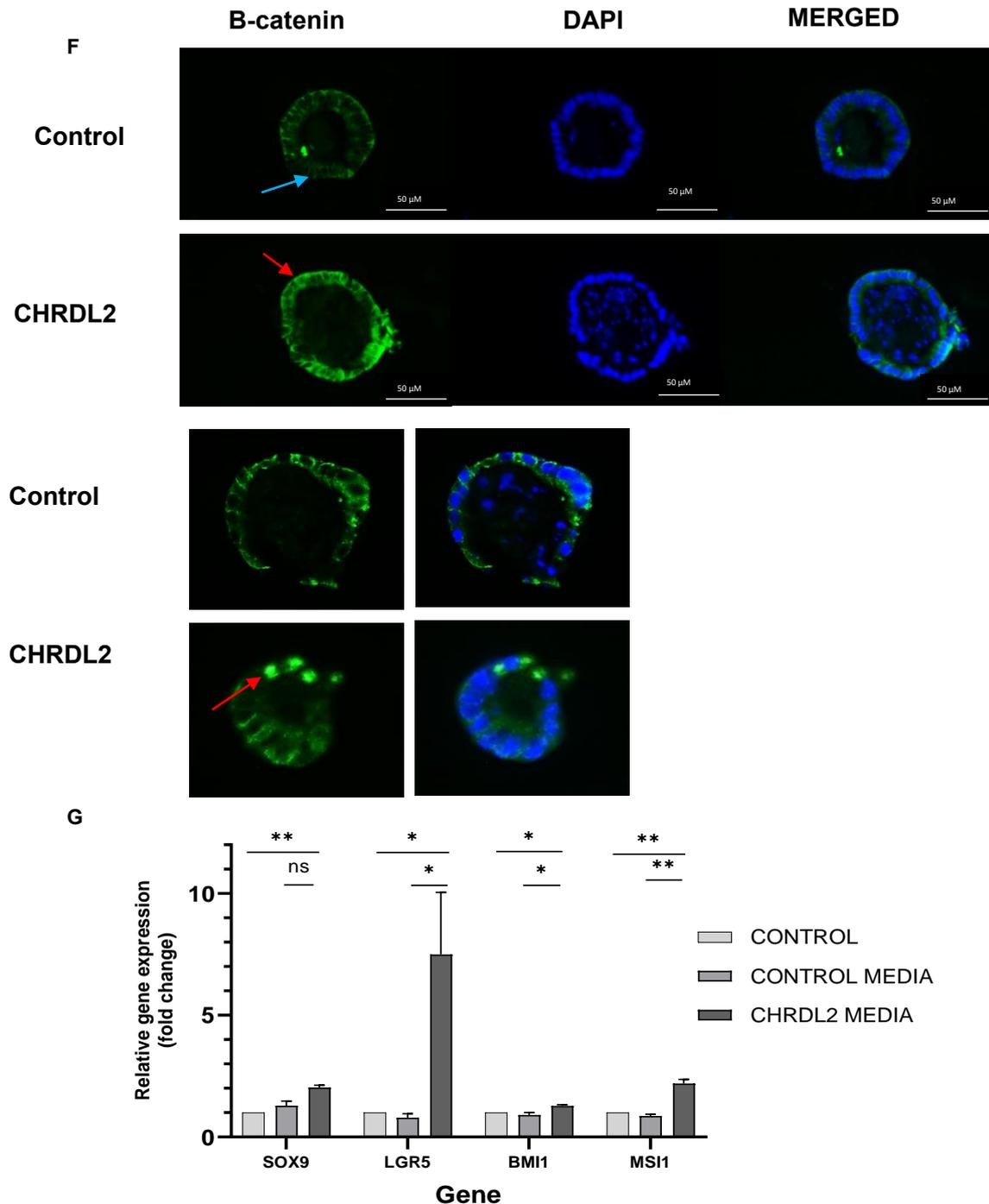


Figure 10: A) Image of murine-derived organoids treated with conditioned media containing secreted forms of CHRDL2 compared to conditioned media from control cells with no CHRDL2 overexpression. B) Quantification of buds per organoid in CHRDL2 treated murine organoids compared to a control. T-test $P < 0.0001$. C) Quantification average organoid diameter in CHRDL2 treated murine organoids compared to a control. T-test $P < 0.001$. D) Immunofluorescence staining of OLFM4 on murine organoids treated with secreted CHRDL2 compared to a control after 1 week. E) Quantification of immunofluorescence scoring of OLFM4 on murine organoids treated with secreted CHRDL2 compared to a control. T-test $P < 0.0001$. F) Immunofluorescent staining of β -catenin localisation in murine small intestinal organoids upon CHRDL2 addition and a control. G) qPCR of stem cell markers from CHRDL2 treated murine organoids compared to a control. T-test SOX9. Students T-test $P < 0.0014$, LGR5 $P < 0.04$, $P < 0.043$, BMI1 $P < 0.0113$, MSI1 $P < 0.0067$, $P < 0.009$. $N = 3$ Error bars given as \pm SEM.

Discussion

In order to investigate the effects of CHRDL2 as a BMP antagonist in colorectal cancer, we transduced CHRDL2 overexpression in CRC cell lines. To ensure BMP antagonism, western blotting was then performed for SMAD1/5 phosphorylation, an intracellular downstream transducer of the BMP signal. CHRDL2 overexpressing cells showed a reduction in P-SMAD1/5, without compromising levels of total SMAD1, confirming CHRDL2 as a BMP antagonist. Next, we tested our hypothesis that CHRDL2 overexpression would increase WNT signalling upon BMP inhibition. Immunofluorescence staining of CHRDL2 cells revealed increased nuclear-localisation of β -catenin, a hallmark of WNT signalling. These data support our hypothesis that CHRDL2 overexpression increases WNT signalling, which is a hallmark of cancer stem-cells.

Traditionally, when considering enhanced oncogenic potential, we think of increased cellular proliferation, migration, and metastasis. This is especially true in the case of WNT driven CRC tumours, in which hyperactivity of the WNT pathway leads to hyperproliferation and increased stem-like phenotypes of cancer cells. Therefore, we wished to test the effects of CHRDL2 overexpression on proliferative abilities of cancer cells.

Therefore we generated growth curves for our 4 experimental cell lines. Surprisingly, CHRDL2 overexpression showed a slight reduction in proliferation in both COLO320, LS180, and RKO cells. This was also shown in low glucose conditions, in order to replicate a low-glucose tumour microenvironment and cause cell stress. Again, CHRDL2 overexpression reduced cell confluency at 96hrs post treatment. To confirm if this was due to lower proliferation or apoptosis, cell cycle analysis was performed. Ki-67 is a marker of proliferating cells, with Ki67 expressed in S, G2, and M phases of the cell cycle, and rapidly degraded during G0 and G1. Initially we gated for Ki-67 + cells and then determined cell cycle status (G1, S, or G2). As seen in Figure 5 C and E, CHRDL2 overexpression decreased the number of cells present in S and G2 phase, and reduced the number of Ki67+ cells. This would suggest that CHRDL2 cells had lower proliferation rate due to less actively proliferating Ki67+ cells and a longer duration in G1. This could indicate a slower growing phenotype, or an entering to a quiescent cell state. Further work would need to be completed to determine if a quiescent phenotype is observed, through staining for quiescent cell markers, such as P27.

This conflicts with previous data by Sun 2017, which showed a significant increase in proliferation with CHRDL2 overexpression in colorectal cancer cell line HCT8 [189]. However, it is important to note that the models for overexpression we used have a doxycycline inducible system, which allows comparison of behaviour within one cell population, and not of separate

cell lines that may have diverged over multiple passages. Previous reporting of CHRDL2 affecting proliferation by Sun 2017 (CRC) and Wang 2022 (Gastric cancer) have used a stably on- shRNA against the CHRDL2 gene, and in the case of Sun2017, a lentiviral overexpression without an inducible system. Furthermore, the cell line used by Sun 2017 to overexpress CHRDL2 was HCT8, which we did not test, so could potentially generate different results due to variation in BMP pathway and BMP receptor expression. These methods, although useful, have their drawbacks, as stably overexpressed or knockdown cell lines may be divergent from their control cell lines and differ in the growth potential over long periods in culture. Furthermore, plating densities from using two different cell populations might differ despite attempts at normalising cell numbers. Therefore, the use of a Tet-on system, which takes a single population of cells and activating CHRDL2 overexpression, is a more robust system.

This may explain some of the reported differences found in CHRDL2 overexpression and proliferation. Furthermore, in our tested cell lines by both qPCR and western blotting, we found minimal to no CHRDL2 expression, which draws into question the validity of shRNA knockdown techniques. In our data we have used 4 different CRC cell lines, which have shown consistently similar results, supporting our findings over previous research. Therefore, we believe our experiments reveal the most accurate portrayal of CRHDL2 overexpression on proliferation.

To further investigate any potential effects of CHRDL2 on proliferation, we analysed proliferation in a low glucose media, to mimic the low-glucose environment present in tumours, and to emphasise any potential differences by causing cell stress. We have shown CHRDL2 overexpression to have a significant reduction in proliferation during low glucose conditions. It has been previously shown that cancer stem-cells prefer to rely on aerobic glycolysis compared to more differentiated cancer cells, known as the Warburg effect. This principle dictates that cancer-stem cells prefer to rely on glucose-dependent growth as glucose uptake by cancer stem cells allows them to fuel long-term survival compared to their non-stem counterparts [200][216]. Furthermore, it has been proposed that the Warburg effect is an early oncogenic transformation and occurs in conjunction with KRAS and BRAF mutations in cancer initiation that is intrinsic to the formation of solid tumours [217]. Therefore, this marked reduction in proliferation in low-glucose environments displayed by our cell lines may indicate that these cells more closely resemble the cancer-stem cell phenotype by their preference for glucose-linked metabolism. This corroborates our theory that CHRDL2 increases stemness through inhibition of BMP signalling and increased WNT signalling.

Clonogenic and migration assays were used to measure the effects of clonogenicity and migratory potential, but also to yet again measure differences in proliferation in our

experimental cell lines. We have shown a reduction in the number of colonies formed by CHRDL2 overexpressing cells compared to the control (Figure 6.A). Stereotypically cancer-stem cells have a higher clonogenic potential, as they retain the intrinsic ability for self-renewal which allows single cells to form new colonies through clonal expansion [218]. These data would suggest that CHRDL2 causes cells to lose this ability for self-renewal, which in turn would be indicative of a less stem-cell phenotype. However, clonogenic assays rely on colonies to grow large enough to be visible to the eye, so are also a measure of proliferative ability as well as clonogenicity. In fact, as seen in figure 6 A, colonies are still present during CHRDL2 overexpression, but these are much smaller in size. Therefore, it is important to take this data into consideration with our proliferation data (Figure 5), which demonstrated a reduction in proliferation during stressed conditions, which would also form under low-density plating which is present in clonogenic assays. Therefore, we cannot draw any conclusions to the self-renewal capacity of CHRDL2 overexpressing cells as this data may be indicative of their reduced proliferative ability.

Scratch assays to measure the proliferation capabilities of CHRDL2 overexpressing cells were also performed. In CACO2 and RKO cell lines there was a decrease in migration/proliferation in CHRDL2 cells as measured by the rate cells migrate into a “scratch” made in confluent adherent cells. This would indicate that in two of our cell’s lines CHRDL2 decreased the migratory ability. This conflicts with our hypothesis that CHRDL2 creates a more stem-cell phenotype as stem-cells are known for their inherent migratory ability [219]. However, as with the clonogenic assay these results could reflect the cells’ proliferative ability rather than migratory. Further studies could be conducted using scratch assays as a model for migration if proliferation blockers, such as mitomycin c, were used, to de-couple proliferation and migration in the quantification for this assay.

Therefore, we utilised a migration assay using transwell inserts where cells were chemo-attracted to a serum-rich environment. This data showed that in fact, CHRDL2 overexpressing cells had an increased migratory potential (Figure 6 E, F), which supports our hypothesis and previous published data [189][190]. Furthermore, this corroborates the conclusion that scratch assays measure proliferative ability, and again showed CHRDL2 overexpression reduced proliferation.

Next, we looked at markers that would signify an increased migratory potential of CHRDL2 overexpressing cells. As seen in figure 8, Immunofluorescence staining of IQGAP1 and EpCAM, known markers of migration, invasion, and adherence, were upregulated in our CHRDL2 overexpressing cells. This supports our findings that CHRDL2 overexpression increase migratory potential.

To further analyse the effect of CHRDL2 on a potential stem-cell phenotype, we analysed known stem-cell markers by qPCR in our 2D cell lines. As seen in figure 9, BMI1, LGR6, and LGR5 were all upregulated by CHRDL2 overexpression. Within the intestinal crypt, normal stem cells are arranged in a hierarchy, with rapidly proliferating stem cells or crypt base columnar cells (CBCs) at the base of the crypt, marked by LGR5+ and LGR6+ expression. A separate population of stem cells lie the +4 position, which appear to cycle more slowly [220,221]. Some studies propose that these slow cycling stem cells at the +4 position are marked by BMI1, which was upregulated in our qPCR data, raising the possibility that CHRDL2 enhances this slow-cycling stem cell phenotype.

Slow cycling CSCs have also been shown to be radiation resistant, like our CHRDL2 overexpressing cells [222,223]. There is some conflicting evidence for the role of these slow-cycling stem cells with several publications proposing that they are key for regeneration of the intestine after injury [221,224]. However more recent studies show that LGR5+ CBCs are also able to fulfil this role or suggest that the two populations support each other to facilitate tissue repair [225]. Nevertheless, recent publications have linked the often over-looked slow-cycling stem cell to worsened patient prognosis, which may explain the effect of CHRDL2 on reduced proliferation but poor patient prognosis and increased stem cell qualities [226]. CHRDL2 may therefore increase the slow-cycling stem-cell phenotype.

We also sought to test CHRDL2 overexpression on a 3D model. As seen in figure 10, murine intestinal organoids were produced in 3D, and treated with secreted human CHRDL2, as used in our secreted drug dose response curves. This allowed us to add measure the effects of CHRDL2 in a 3D environment that more closely mimics *in vivo* conditions, as well as a non-cancer model. CHRDL2 has previously been reported to mainly be expressed in the mesenchyme of the intestinal tract, where it is secreted and acts through paracrine signalling by binding to BMP receptors [189]. Therefore, addition of secreted CHRDL2 in a 3D condition allows us to study the effects of CHRDL2 in an environment that more closely resembles the physiological process. Organoids treated with CHRDL2 displayed a more spherical phenotype, with few differentiated buds, and a smaller diameter (Figure 10 B, C, D). This suggests that CHRDL2 increased a more stem-like phenotype in murine intestinal organoids, with reduced growth which supports our observed reduction in proliferation in 2D cell lines. This is supported by an increase in β -catenin nuclear localisation observed in intestinal organoids, a hallmark of WNT signalling. As expected, we also saw an increase in mRNA levels of stem cell markers LGR5 BMI1, LGR6, and SOX9 with CHRDL2 addition ($P < 0.05$). Furthermore, immunofluorescence staining of OLFM4 marker, which marks LGR5+ stem cells, was also significantly increased (Figure 10, E). This supports our theory that CHRDL2

increases the stem-cell phenotype through blocking the BMP differentiation pathway. However, no other markers were found to be significantly increased or decreased.

It should be noted that human CHRDL2 was used to treat murine organoids, so there will be some inter-species differences in the CHRDL2 protein. As seen in figure 11, human and mouse CHRDL2 do show high homology, with both proteins possessing the same type and number of domains. However, when the sequence was analysed through FASTA, a protein sequence similarity score of 78%.1% was obtained. Therefore, these experiments could be repeated using a mouse CHRDL2 secreted protein, to see the full effects of CHRDL2 in a 3D environment.

A	Identity	302/430 (70.2%)
	Similarity	336/430 (78.1%)
B	Gaps	18/430 (4.2%)

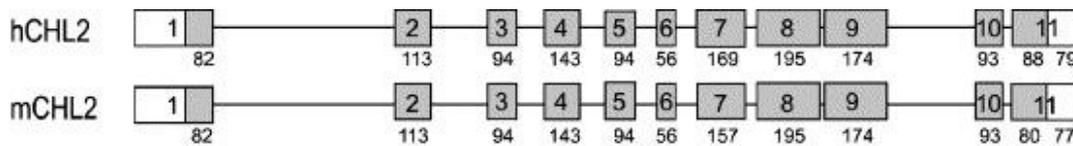


Figure 11: A) Sequence identity and similarity score between human and mouse CHRDL2. B) comparison of human and mouse CHRDL2 transcripts. Diagram represents exons of both human and mouse CHRDL2 protein with number of base pairs in exons. Data obtained by NCBI blast.

Limitations and future work

We have shown that CHRDL2 reduces proliferation through cell confluency assays, and increases migration through a porous membrane, as well as increases the expression of migratory markers. However, our use of scratch assays as a determinant of migration

conflicted with the other results observed. This is potentially due to the reduction in proliferation of CHRDL2 overexpressing cells confounding the migration results within the scratch assay, as previously discussed. However, without the use of proliferation inhibitors, such as mitomycin C, we cannot determine if this observation is a reflection of clonogenicity of migratory potential. Therefore, future work would be needed to uncover this. Furthermore, the clonogenic assay used in this study is tightly linked with proliferation, as larger colonies need to be observed by eye to be counted for analysis, and larger colonies form through cell proliferation. Further work would be needed, such as the use of plating single cells in 96-well plates, to uncover the potential of a single cell to form a colony.

The use of quiescent markers, such as P27, to determine if CHRDL2 overexpressing cells are entering a quiescent stem-like state would also be beneficial to explain the cell-cycle result shown here. Currently, this study cannot determine if CHRDL2 has reduced the proliferation of cells by holding in G1, or entering quiescence at G0.

Another limitation of this study is the use of human CHRDL2 on murine organoids, as previously discussed. Despite high similarity between murine and human CHRDL2, the addition of mouse CHRDL2 on murine organoids would be a more faithful model to uncover the full effects of CHRDL2.

Conclusion

Collectively, this data shows the effects of CHRDL2 overexpression in increasing a slow-cycling stem-like characteristic in colonic epithelial cells. CHRDL2 overexpression caused an increase in migratory potential, and decreased differentiation and increased stem-cell markers in a 2D and organoid model. However, we also observed a decrease in proliferation, and slow cell-cycle progression in CHRDL2 overexpressing cells. This data suggests that CHRDL2 increases the phenotype of a slow-cycling stem cell, rather than the actively proliferating stem cell phenotype.

Chapter 4: CHRDL2 and chemotherapy resistance

Introduction

Standard practice of care for all but the earliest stage of CRC relies on the use chemotherapy and radiotherapy treatments. Chemotherapy and radiotherapy are used in combination with

surgical procedures and aim to prevent the proliferation of tumour cells through disrupting DNA synthesis, resulting in DNA damage, and therefore cell death.

Cancer stem cells have long been known to harbour increased resistance to chemotherapy, through upregulation of DNA repair pathways [227]. As a BMP inhibitor, we have previously shown CHRDL2 upregulation increases stem-cell WNT signalling in CRC cell lines. Therefore, we wished to test whether CHRDL2 upregulation would increase survival during chemotherapy treatment, through increased stem-cell signalling resulting in upregulation of DNA repair pathways.

5-Fluouracil (5FU) is currently the cornerstone of chemotherapy treatment used to treat CRC [228]. 5FU is a fluorinated pyrimidine that acts through inhibition of the thymidylate synthetase, an enzyme involved in pyrimidine nucleotide synthesis [229]. Within the UK, 5FU is used in combination with either Oxaliplatin, a diamino cyclohexane platinum compound that forms DNA adducts (Known as FOLFOX) or Irinotecan, a topoisomerase I inhibitor (known as FOLFIRI)[228].

The addition of Oxaliplatin to 5FU treatment displayed a higher response rate to 5FU alone at improved patient progression-free survival to 9.0 months vs. 6.2 month with 5FU alone, whereas irinotecan also improved patient survival in combination with 5FU to 10.8 vs. 8.5 months [230], 5FU alone has a patient response rate of 22. Oxaliplatin on its own also has a low response rate of only 10% in CRC patients, but when combined with 5FU increases response rate to 53%. Similarly, irinotecan improves response rate to 39% when used with 5FU, highlighting the importance of combinations treatments [230]. In terms of effects, FOLFOX is more likely to cause neuropathy and FOLFIRI is more likely to cause GI toxicities, including diarrhoea, abdominal pain, and nausea/vomiting. Therefore, oxaliplatin based therapy may not be suitable for patients with comorbidities such as diabetes, who might already have neuropathy, so a FOLFOX regimen would be more suitable [231]. On the other hand, women have been shown to experience worse nausea and vomiting compared to men following chemotherapy, in which FOLFOX would be the preferred treatment [231].

Therefore, we tested these 3 chemotherapy agents in our 2D cell lines to study the effects of CHRDL2 overexpression in response to the agents.

Following treatment with chemo and radiation therapy, DNA damage response pathways are activated. Single stranded and double stranded breaks (DSBs) are commonly formed after treatment with chemotherapy agents 5FU, irinotecan, and Oxaliplatin.

DSBs are considered one of the most lethal forms of DNA damage and are the main mechanism of action to cause genome instability and ultimately cell death by chemotherapy

[232]. For example, oxaliplatin causes DNA intra-strand cross-linking which causes DSBs, resulting in cell-cycle arrest and apoptosis [233]. Furthermore, chemotherapy drug 5FU has also been shown to work in collaboration with DSB forming agents such as oxaliplatin and radiotherapy, to increase the instability of DNA damage and ultimately cell death in cancer cells [234]. Together, 5FU and Oxaliplatin are the most common forms of chemotherapy to treat CRC in the UK, and therefore DSBs are an excellent measure of chemotherapy-induced DNA damage. Another chemotherapy agent, irinotecan, causes supercoiling, in which single and double strand breaks occur.

During early DSB response, there is a rapid phosphorylation of the minor histone H2A variant, at mammalian Ser-139 to produce γ H2AX [235]. γ H2AX accumulation at nuclear foci then aids in the recruitment of repair factors to the damaged DNA sites, leading to the activation of repair pathways, in which ATM and ATR serve as central transducers. In recent years immunofluorescence staining of phosphorylated H2AX has become widely popular as a method of detecting DSBs which proves reliable and repeatable. Phosphorylation of H2AX is detected as foci presented in the nucleus of cells which can be quantified as a reflection of DNA damage.

During DNA damage in the form of DSBs, there is a dynamic shift in chromatin structure triggered by HP1-beta release from Histone H3, which in turn causes ATM and ATR to propylated H2AX [236]. Formation of γ H2AX causes recruitment of repair molecules, such as the MRN complex (composed of Mre11, RAD50, and Nbs1), and CIP. These work to stabilise DNA strands and work in conjunction to begin the process of homologous recombination (HR) to repair DSBs. MDC1 is then bound to γ H2AX, and recruits BRCA1, which begins the process of DNA repair. Other proteins, such as RAD51, RAD54, and RAD21 are also recruited. Damaged DNA is then removed through exonucleases EXO1, mediate by RPA (Figure 18). Rad21 and FANCF are then recruited, and correct DNA synthesis is repaired by Pol δ and Pol ϵ , and ligation of newly synthesised DNA by ligase 1.

Alternatively, non-homologous recombination may occur at DSBs at sites in which no overhangs are present (figure 18). In such cases, Ku70 and Ku80 bind to DNA ends, and PCKs are recruited to stabilises the recognition complex. Once DNA-PKCS are recruited, ARTEMIS, and DNA endonuclease, is recruited and dephosphorylated by DNA PCKs, to removed damaged bases. XRCC4 and XL4 are the recruited and bind to the KU70/Ku80 heterodimers, to facilitate the binding of ligases. New DNA synthesis is performed by Pol μ , and ligase IV to synthesis new DNA strands (Figure 18).

Alternatively, in the presence of crosslinks, such as formed by oxaliplatin, that do not result in DSBs, nucleotide excision repair is activated. Recognition of crosslinks is accomplished

through either transcription coupled repair, or global excision repair (figure 18). Unwinding of Crosslinked DNA is performed through proteins TFIH and ERCC5, before excision by ercc1, XPR, and XPA. Gap filling of missing DNA is then performed by PCNA, and Pol δ , before ligation by ligase 1. PCNA is also involved in Base excision repair, where single strand breaks are present (figure 18).

In this chapter we have looked at the response in survival of our 2D cell lines and 3D organoid models following chemotherapy. We have also analysed the repair of damaged DNA following chemotherapy, and that resulting activated DNA damage response pathways.

CHRD2 increases resistance to chemotherapy.

To measure the effects of CHRD2 on chemotherapy resistance, drug dose response assays were performed. This was carried out by plating 100,000 cells/ ml in a 96 well plate and treating with increasing concentrations of the chemotherapy agents before incubation for 72hrs. Cell number was then assessed through MTS assay. As you can see in figure 12 A, a curve was generated in which cell number was plotted against chemotherapy concentration (μM). IC50

values were then calculated from the curve, which represent the dose of chemotherapy at which half the cells survived (50% cell number). As we can see in panel A, cells with CHRDL2 overexpression had a higher survival rate when exposed to chemotherapy, compared to the control. This observation is confirmed by IC50 values, which when compiled showed a 2-fold increase in IC50 values across all cell line and chemotherapy drugs ($P < 0.001$) (Figure 12, B).

To statistically compare these results, IC50 value replicates were compiled and compared between DMSO control, and 10mg/ml doxycycline (CHRDL2++) treated samples. As seen in figure 10 C, in the CACO2 cells line CHRDL2 overexpression significantly increased the IC50 value of oxaliplatin and 5FU treated cells. In COLO320 cells IC50 values were increased by CHRDL2 overexpression in oxaliplatin, and irinotecan treated cells. In LS180 cells, the IC50 value was significantly increased when treated with oxaliplatin. However, in RKO cells, the IC50 values of chemotherapy treatment were not significantly increased by CHRDL2 overexpression. The IC50 values for each chemotherapy drug for our 4 experimental cell lines can be seen in panel D. COLO320 cells treated with Oxaliplatin showed the highest increase in IC50 during CHRDL2 overexpression.

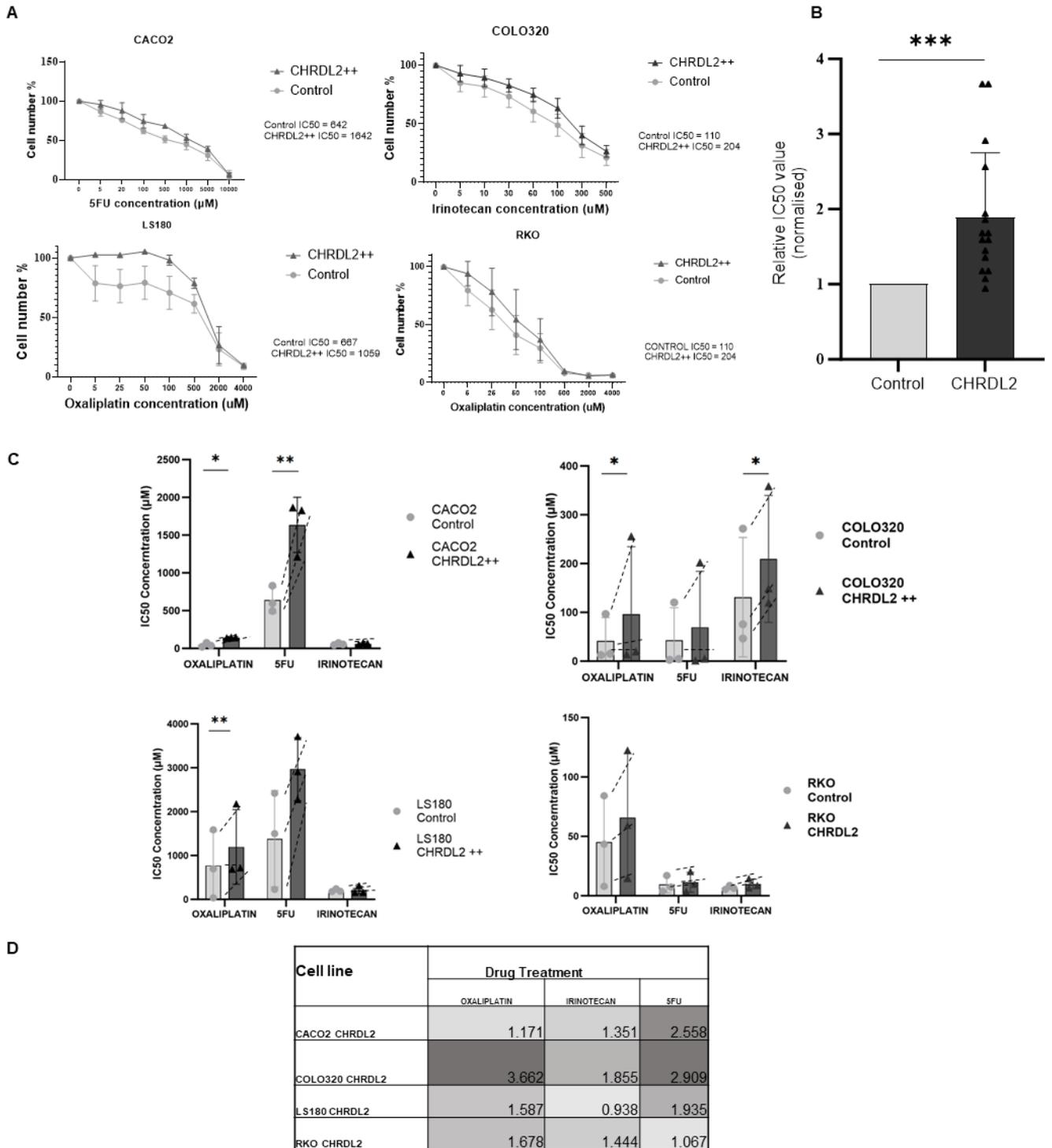


Figure 12: A) Drug dose response curves using CACO2 cells and 5FU, COLO320 cells and Irinotecan, and LS180 and RKO cells and Oxaliplatin N=3. Two-way ANOVA was used to find differences between curves, $P < 0.0068$, $P < 0.0001$, $P < 0.0006$, $P < 0.005$. Linear regression analysis was used to find the IC₅₀ of these cell lines with baseline = 0. B) Average difference in IC₅₀ values across all cell lines and 3 chemotherapy drugs. $P < 0.005$. C) Average IC₅₀ values of CACO2, COLO320, LS180 and RKO cell lines using chemotherapy drugs Oxaliplatin, 5FU, and Irinotecan N=3. D) Table of ratio differences in IC₅₀ values for each chemotherapy drug and cells between CHRDL⁺⁺ cells and control. N=3. Error bars given as \pm SEM.

CHRDL2 overexpression reduces cell cycle stalling during chemotherapy.

We then sought to test how chemotherapy treatment effected the cell-cycle in COLO320 CHRDL2 overexpressing cells. As our chemotherapy agents target DNA synthesis, we can expect to see activation at the S/G2 phase checkpoint, causing cells to gather in S and G2 phases, and a lower proportion of cells in G1. This is observed in figure 13 A, which clearly shows a small proportion of cells in G1 (40% from 50%), and the majority of cells stalled in the S/G2 phase of the cell cycle. This is a large difference between the normal cell cycle observed in figure 5A.

To analyse differences in cell cycle during CHRDL2 overexpression and chemotherapy, we chose COLO320 cells treated with oxaliplatin, as these previously showed the greatest increase in IC50. Interestingly, cells with CHRDL2 overexpression displayed a smaller number of cells stalled in S Phase at 25% compared to 30% of control cells, and therefore showed a greater proportion of present in G1 phase. This is the opposite to that of untreated cells in Figure 5A, where CHRDL2 decreased the number of cells in S Phase possible due to slower cell division.

CHRDL2 also increased immunofluorescence staining of proliferation marker Ki67 in COLO320 cells treated with oxaliplatin (Figure 13, C). This was confirmed by flow cytometry, which showed the number of Ki-67 + cells were increased during chemotherapy treatment ($P < 0.0001$) (Figure 13 D). This would indicate increased growth, and there for survival, of cells treated with CHRDL2 overexpression during chemotherapy treatment.

Further flow cytometry analysis of apoptosis marker Annexin V following chemotherapy treatment was performed. CHRDL2 overexpression decreased the number of cells that had entered early apoptosis ($P < 0.05$) (Figure 13 E and F) demonstrating that CHRDL2 overexpressing cells have the ability to evade apoptosis.

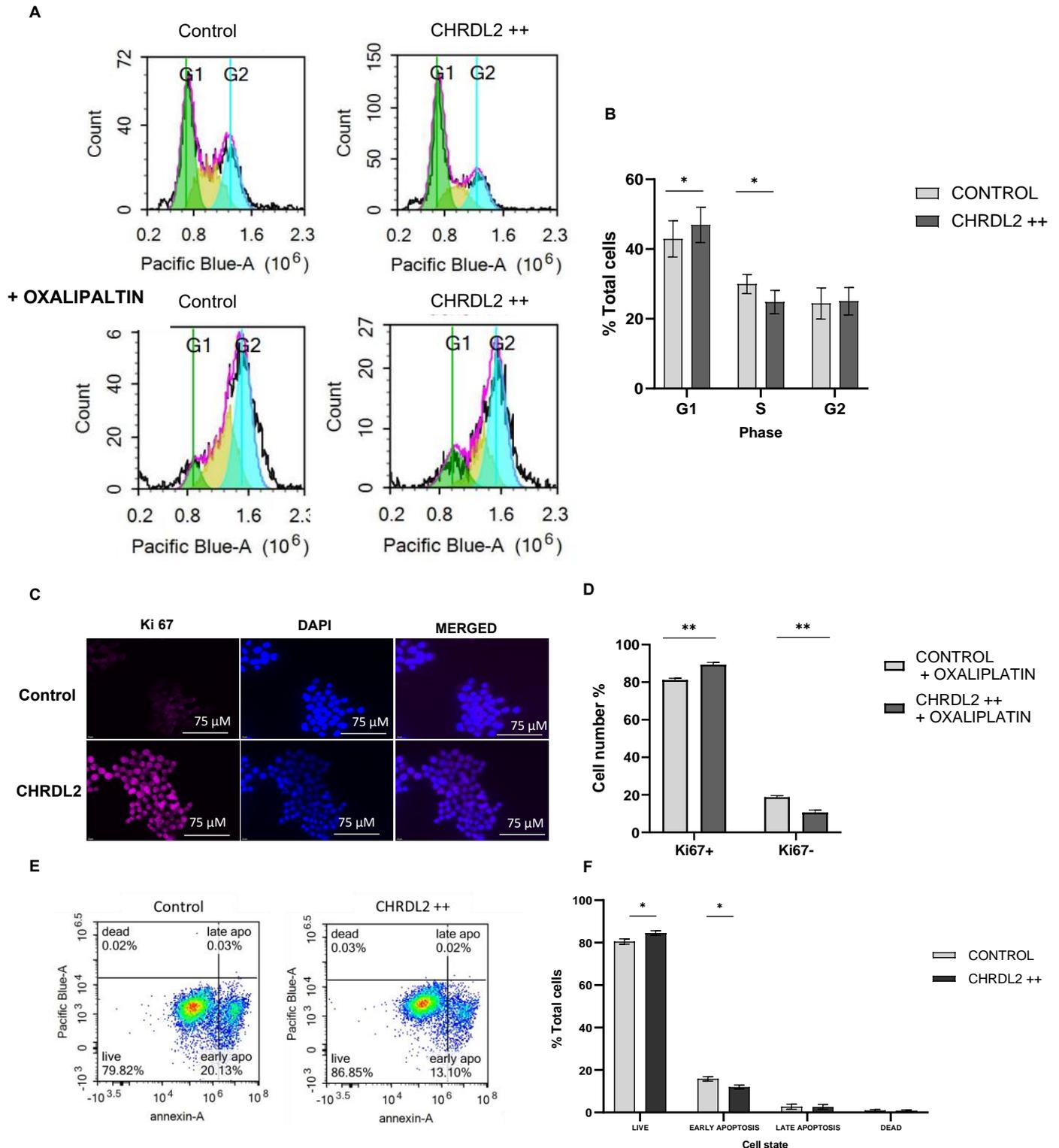


Figure 13: A) Cell cycle expression analysis by flow cytometry of COLO320 cells (top row), and COLO320 cells treated with IC50 Oxaliplatin (bottom row), with CHRDL2++ overexpression or a DMSO control. Green: G1, Yellow: S, Blue: G2. B) Quantification of cell cycle status in COLO320 cells given CHRDL2 overexpression and treated with IC50 oxaliplatin. G1: $P < 0.05$, S: $P < 0.05$. C) Immunofluorescence staining of proliferation marker Ki67 on COLO320 cells with CHRDL2++ overexpression and treated with oxaliplatin. D) Quantification of Ki67 immunofluorescence staining on COLO320 cells with CHRDL2++ overexpression and treated with IC50 oxaliplatin. Ki67+ $P < 0.001$, Ki67- 0.001. E) Flow cytometry analysis of apoptosis markers on COLO320 cells with CHRDL2++ overexpression and treated with IC50 oxaliplatin. F) Quantification of Flow cytometry analysis of apoptosis markers on COLO320 cells with CHRDL2++ overexpression and treated with IC50 oxaliplatin. Live: $P < 0.05$. Early apoptosis: $P < 0.05$. Error bars given as \pm SEM.

Secreted CHRDL2 increases chemotherapy resistance.

In vivo, CHRDL2 is not expressed by non-cancerous epithelial cells, but by fibroblasts in the mesenchyme of the intestinal villi. Secretion of CHRDL2 by these mesenchyme cells then elicits BMP antagonism in epithelial cells through paracrine signalling. Therefore, we sought to repeat our drug dose response curves with parental cell lines that were given secreted forms of CHRDL2 to more closely mimic *in vivo* conditions for CHRDL2 signalling. Drug dose response curves using Oxaliplatin chemotherapy were performed using our parental cell lines CACO2 and RKO which did not contain the CHRDL2 expression vector. These cells were then given conditioned media that was harvested from their CHRDL2-expressing counterparts and a control. Conditioned media was produced by adding standard glucose media to CHRDL2 overexpressing cells or control cells with addition for doxycycline for 48hrs, and then harvested and filtered. Conditioned media was given in a 1:1 ration with standard media.

As seen in figure 14, both cell lines showed an increase in cell survival when given CHRDL2-conditioned media, which mimicked the response we saw in figure 10. When IC50 values were compiled, both cell lines showed an increase in IC50 values when CHRDL2 was overexpressed ($P < 0.05$) (Figure 14 A, B), and when taken together showed a significance in IC50 values of $P < 0.035$ (Figure 14 C).

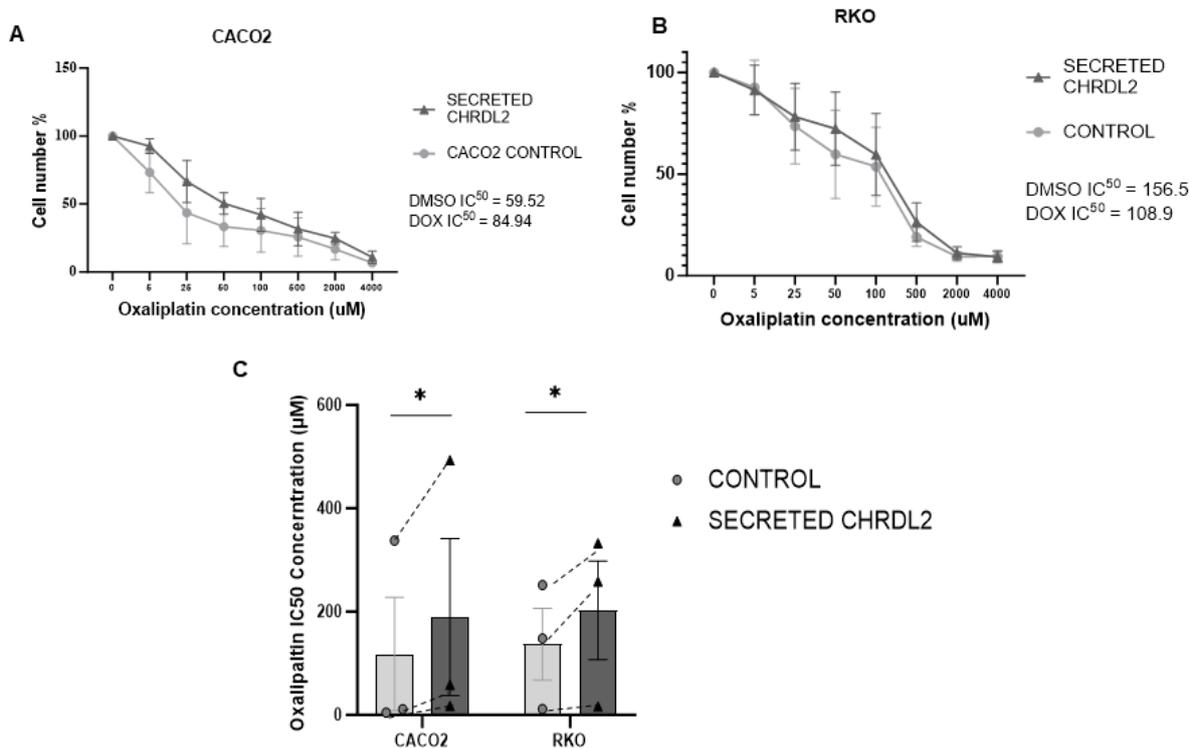


Figure 14: A) Oxaliplatin drug dose response curve of CACO2 cells with CHRDL2 conditioned media. N=3. Two-way ANOVA of curve differences $P < 0.0147$. Linear regression analysis was used to find the IC₅₀ of these cell lines with baseline = 0. B) Oxaliplatin drug dose response curve of RKO cells with CHRDL2 conditioned media. N=3. Two-way ANOVA of curve differences $P < 0.0732$ C) Average IC₅₀ values for chemotherapy drug Oxaliplatin, on cell lines CACO2 and RKO with CHRDL2 conditioned media. CHRDL2 v control $p < 0.0305$. N=3. Error bars given as \pm SEM.

CHRDL2 addition increases chemotherapy resistance in intestinal organoids.

Next, we sought to replicate our findings on chemotherapy resistance in 2D cell lines on a 3D organoid model. Murine intestinal organoids were established as before, with secreted CHRDL2 addition in the media. After 24hrs post derivation, chemotherapy 5FU was added to organoids at 0.5 μ M, 1 μ M, and 5 μ M. As with our 2D cell lines, 72hrs post chemotherapy, images of treated organoids were taken to observed differences between CHRDL2 treated organoids and a control. As seen in images in figure 15 A, organoids exhibited the beginnings of apoptosis (blue arrow) and death (Red arrow), in response to chemotherapy in both control and CHRDL2 treated groups. However, it is clear there is more surviving organoids (green arrow) in CHRDL2 treated groups, especially at the highest dose of chemotherapy. This is confirmed by quantification in figure 15 B, which highlights the increased survival of CHRDL2 treated organoids ($P < 0.0442$). This would suggest, that as with our 2D cells lines, CHRDL2 treatment increased overall survival when subjugated to chemotherapy. Interestingly, CHRDL2 treated organoids appeared to be more differentiated at higher levels of chemotherapy, which could perhaps be an adaptive response.

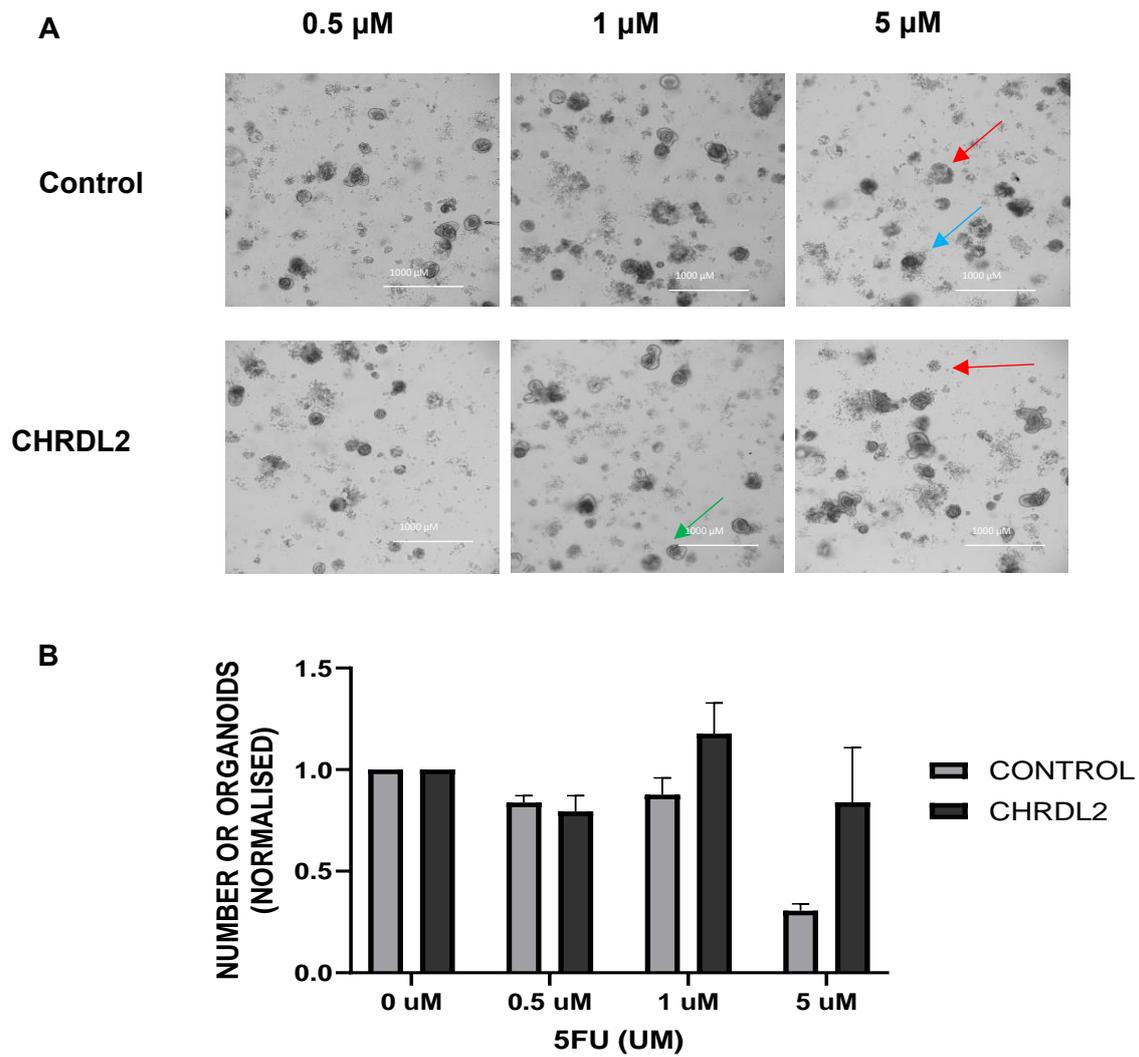


Figure 15: A) Images of murine-derived organoids treated with conditioned media containing secreted forms of CHRDL2 compared to a control and 5FU. Image taken 96hrs post treatment. I) Quantification of number of live organoids in CHRDL2 treated organoids compared to a control. N=3. 2-Way ANOVA $P < 0.0442$. Error bars given as \pm SEM

CHRDL2 increases radiation resistance.

Radiation is the second most common form treatment for CRC and is often used in combination with the use of chemotherapies [237]. Resistance to irradiation, along with chemotherapy resistance, is also attributed to more aggressive cancers that evade treatment [238]. Intestinal stem cells have been shown to have increased resistance to irradiation, therefore we sought to measure the effects of CHRDL2 overexpression on cell survival during X-ray irradiation. Cells were treated with 0 GY, 2 GY, 4 GY or 6 GY X-ray irradiation and cell viability was assessed after 72hrs. As seen in figure 16, CHRDL2 overexpression increased cell survival at 4GY and 6GY radiation ($P < 0.03$, $P < 0.02$, t-test).

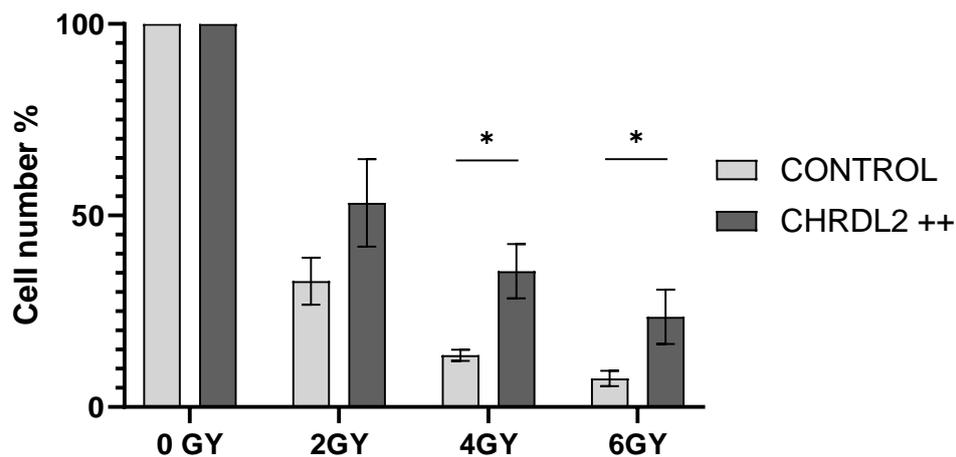


Figure 16: Cell count after irradiation of RKO cells overexpressing CHRDL2 N=3. T-test 4GY: $P < 0.038$, 6GY: $P < 0.0241$. Error bars given as \pm SEM.

CHRDL2 overexpressing cells have high BMP antagonism following chemotherapy treatment.

We then analysed BMP antagonism in CHRDL2 overexpressing cells after surviving chemotherapy treatment. CHRDL2 overexpressing cells and control cells were treated with IC50 5FU for 72 hours in the same manner as our drug dose response assays. Surviving cells were then harvested by trypsinization and centrifugation at 300 x G, and P-SMAD1/5 protein levels were analysed. This was compared the cells that had not undergone chemotherapy treatment. As seen in figure 17 A, B, CHRDL2+ and CHRDL2++ cells had reduced P-SMAD1/5 expression of 75% compared to a control. Compared with previous data shown in figure 3 G, this reduction in P-SMAD1/5 is greater following chemotherapy than in un-treated CHRDL2 overexpressing cells. This would suggest that cells with the highest levels of BMP antagonism were more readily able to survive chemotherapy, resulting in the surviving cells showing a greater reduction of P-SMAD1/5. This would suggest that there is a direct relationship between BMP antagonism and chemotherapy resistance, possible as a result of enhanced WNT signalling causing increased cell survival.

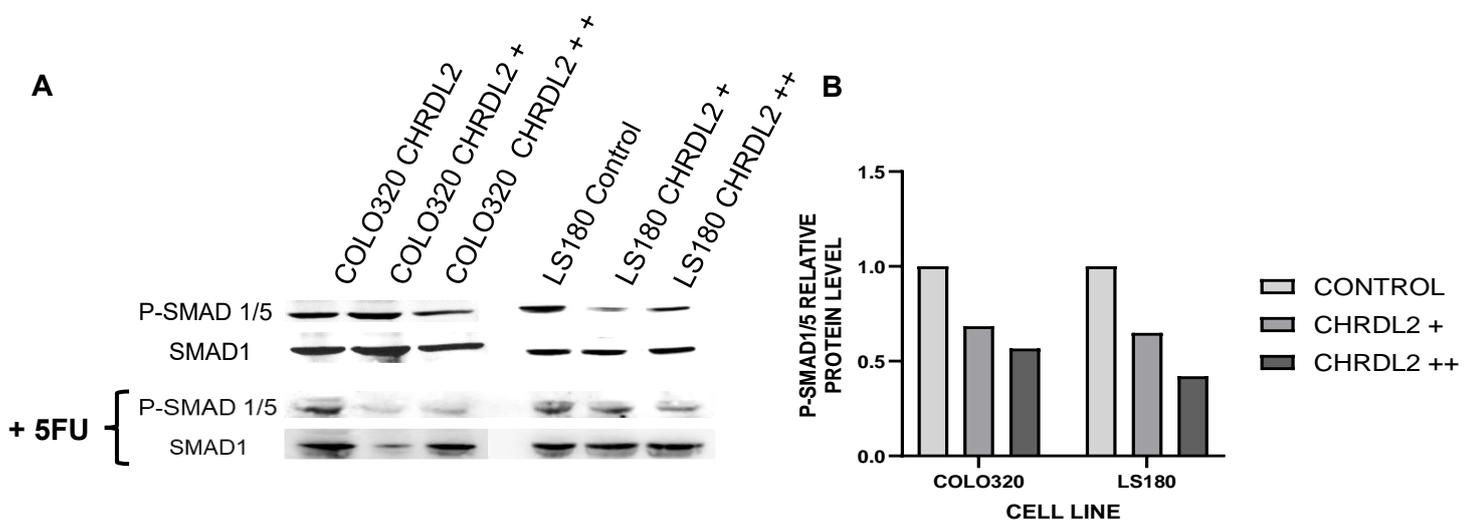


Figure 17: A) Western blotting of SMAD1/5 phosphorylation in cell lines overexpressing CHRDL2 treated with 10 μ M 5FU compared to a control. B) Western blot quantification of P-SMAD1/5 in 5FU treated cells

CHRD2L2 decreases DNA damage through upregulation of DNA repair pathways during chemotherapy.

Chemotherapy exerts its effects through preventing DNA synthesis or damaging DNA, causing activation of cell cycle inhibitors which prevents cell division. This manifests in the production of double-stranded and single-stranded DNA breaks. Following DNA damage, cells can activate DNA repair pathways (Figure 18) to repair damaged DNA in order to survive. This is especially true in the case of cancer stem-cells, which acquire increased capacities for repairing DNA damage, aiding to their resistance to common therapies [238].

Therefore, we tested whether CHRD2L2 overexpression was able to enhance chemotherapy resistance through decreased DNA damage and upregulation of DNA repair pathways. To do this, we treated COLO320 cells with CHRD2L2 upregulation and Oxaliplatin. COLO320 cells treated with Oxaliplatin were chosen as they demonstrated the highest level of chemotherapy resistance. We then analysed several of DNA repair pathways, such as BER, HR, NHEJ, and NER, in our cells.

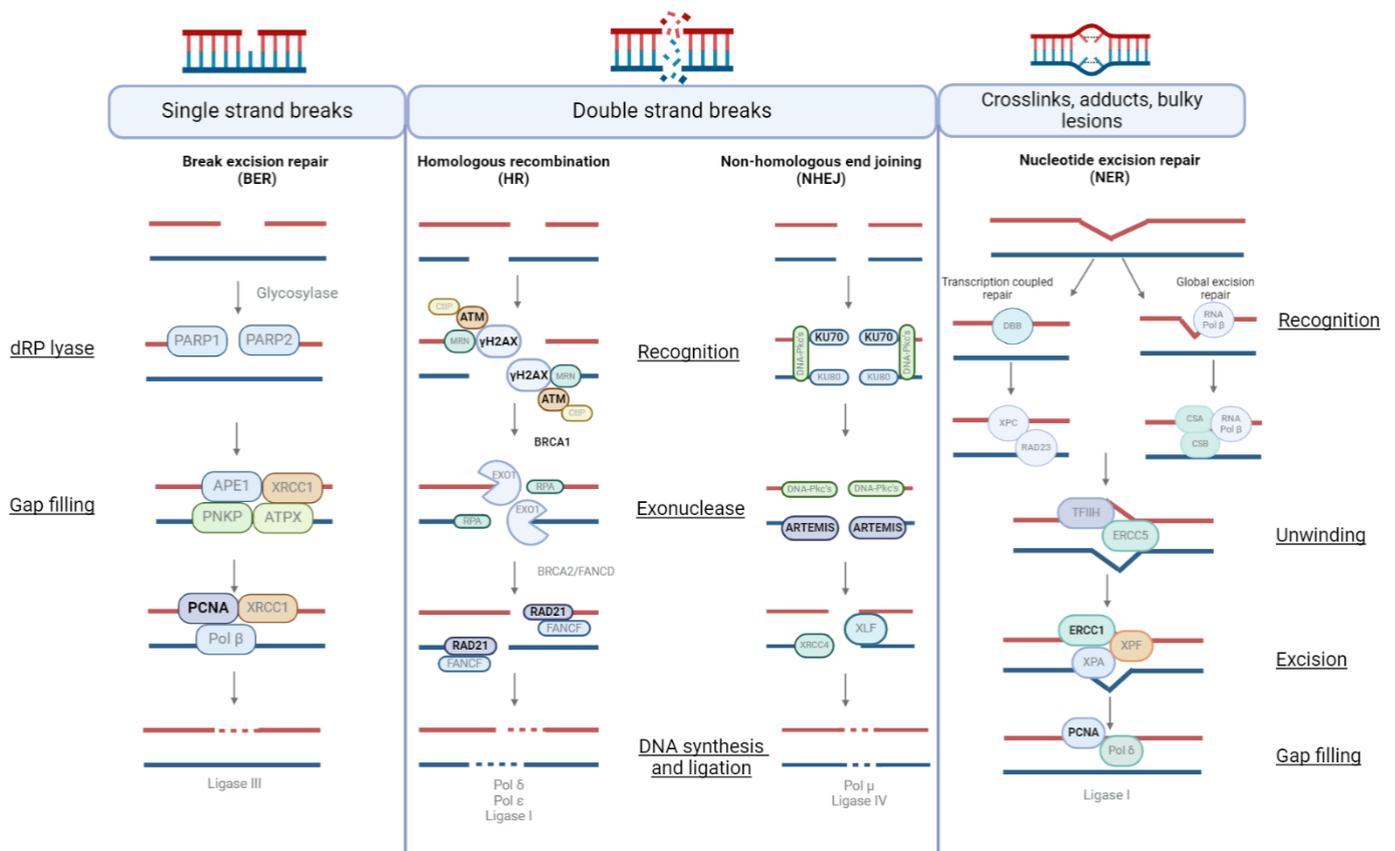


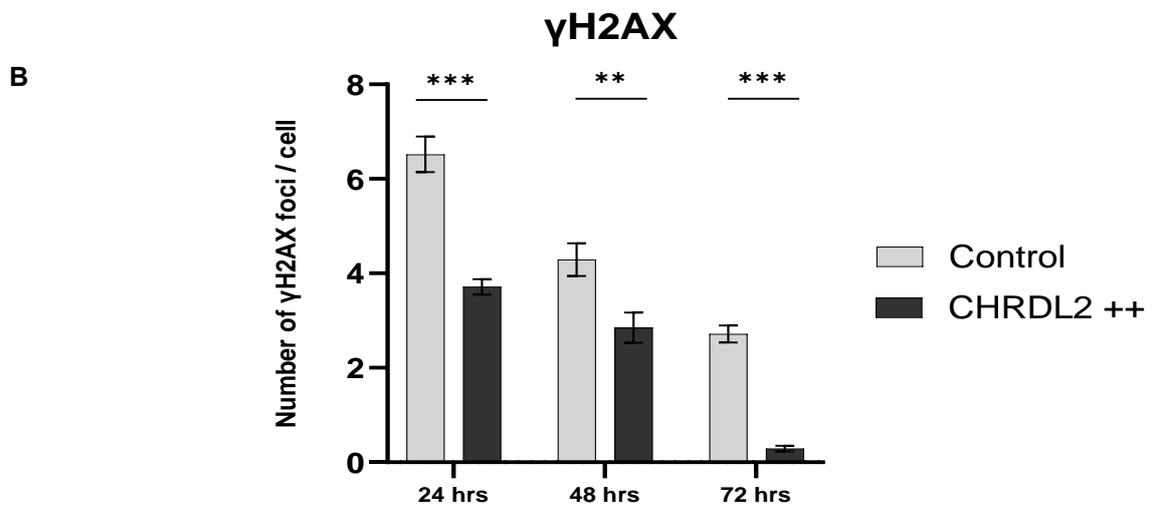
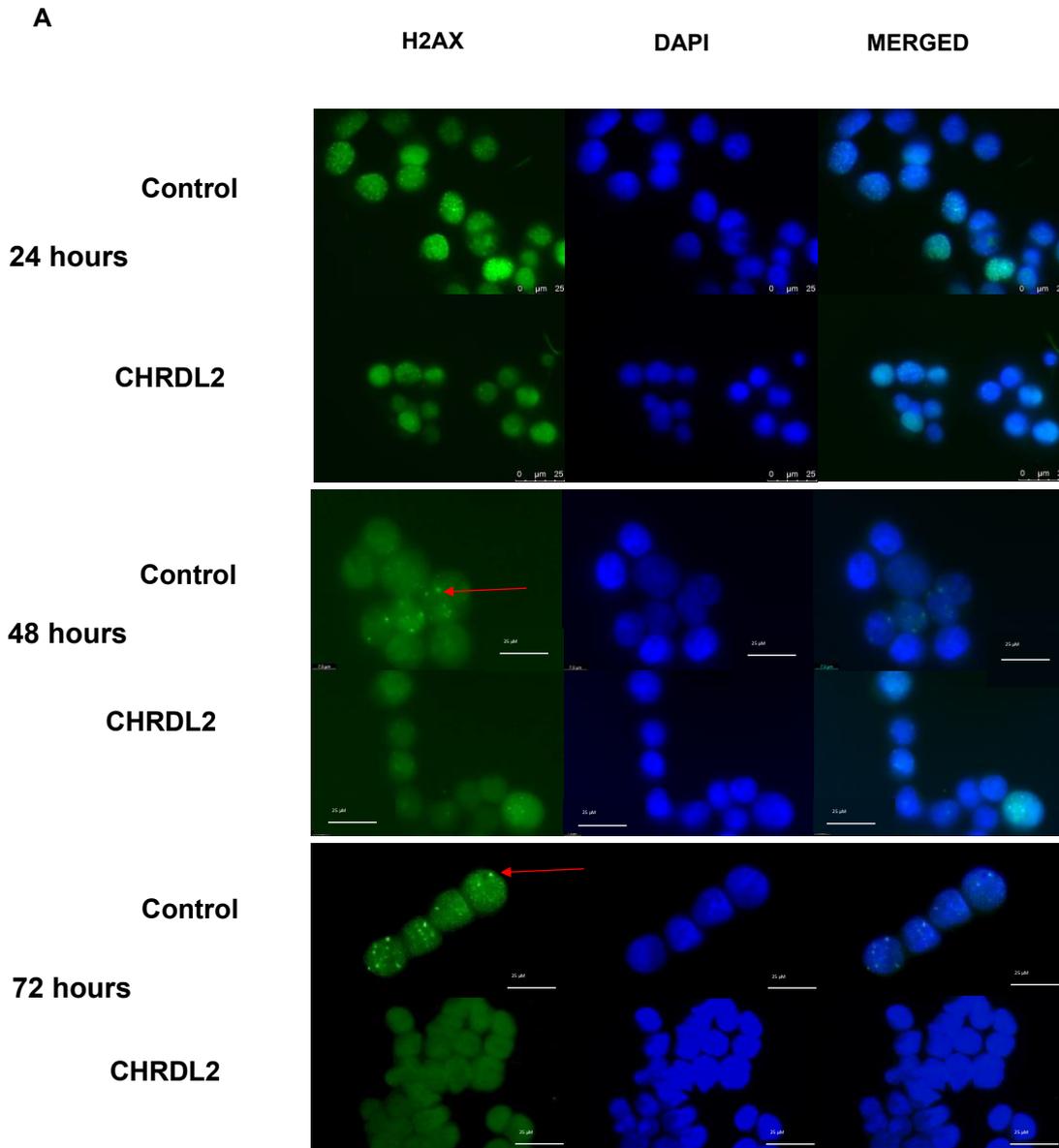
Figure 18: Figure demonstrating the pathways of DNA repair following Chemo and radiotherapy. Bold texts highlights proteins analysed.

Initially, we looked at the presence of DNA damage caused by chemotherapy in our cells lines. To achieve this, quantification of DSBs in cells treated with chemotherapy Oxaliplatin was performed. Cells were treated with 5 μ M Oxaliplatin which is approximately the IC25 value for this cell line and should therefore cause 25% of cells to die. This will allow us to look at damaged DNA without killing 100% of cells. Staining of double stranded breaks was performed using γ H2AX, to measure DSB repair via HR, and Ku70, to analyse DSB repair via NHEJ.

Staining of γ H2AX was performed and foci quantified at 24, 48, and 72 hours post treatment (Image 19, A). As seen in panel A, γ H2AX foci as seen in both CHRDL2 and control treatment, but fewer can be observed in CHRDL2 treated cells at nearly 50% of the control, which is supported by quantification in panel B ($P < 0.0001$). This difference is augmented at 48 and 72 hr post treatment, where CHRDL2 treated cells show a clear reduction of γ H2AX foci at an accelerated level compared to that of the control. At 72hrs the biggest difference can be observed, where CHRDL2 overexpressing cells harbour less than 10% of γ H2AX foci than that on the control. This is shown by quantification in panel B, which shows nearly all foci have been cleared in CHRDL2 cells, whereas many foci are still present in the control. This would suggest that CHRDL2 overexpressing cells are able to repair DNA damage at a faster rate than their non-expressing counterparts. Accelerated DNA repair is a hallmark of CSCs, which further supports our hypothesis that CHRDL2 increases stem-ness in cancer cells [239].

We also quantified the presence of Ku70 expression at 72hrs post treatment. Like H2AX, Ku70 is a marker of DSB with the NHEJ pathway. As seen in figure 19 C, there is a strong reduction in Ku70 expression during CHRDL2 overexpression at less than 50% of control cells, signifying a reduction of DSBs repair by NHEJ. This is quantified in panel D, which shows tis significant reduction ($P < 0.005$).

Next, we looked at the presence of ATM staining (panel E), which binds to γ H2AX, to initiate repair of DSBs. ATM expression was significantly increased in CHRDL2 over-expressing cells, suggesting regulation of DNA repair pathways, which corroborates our findings of reduced DSBs. We also analysed the expression of RAD21, as seen in figure 19 F, which, like ATM, is involved In DSB repair, and was upregulated during CHRDL2 overexpression. Both ATM and RAD21 expression was quantified, and as seen in panel G, had significant overexpression during CHRDL2 treatment and had a fold increase of over 100% ($P < 0.001$, $P < 0.01$, respectively).



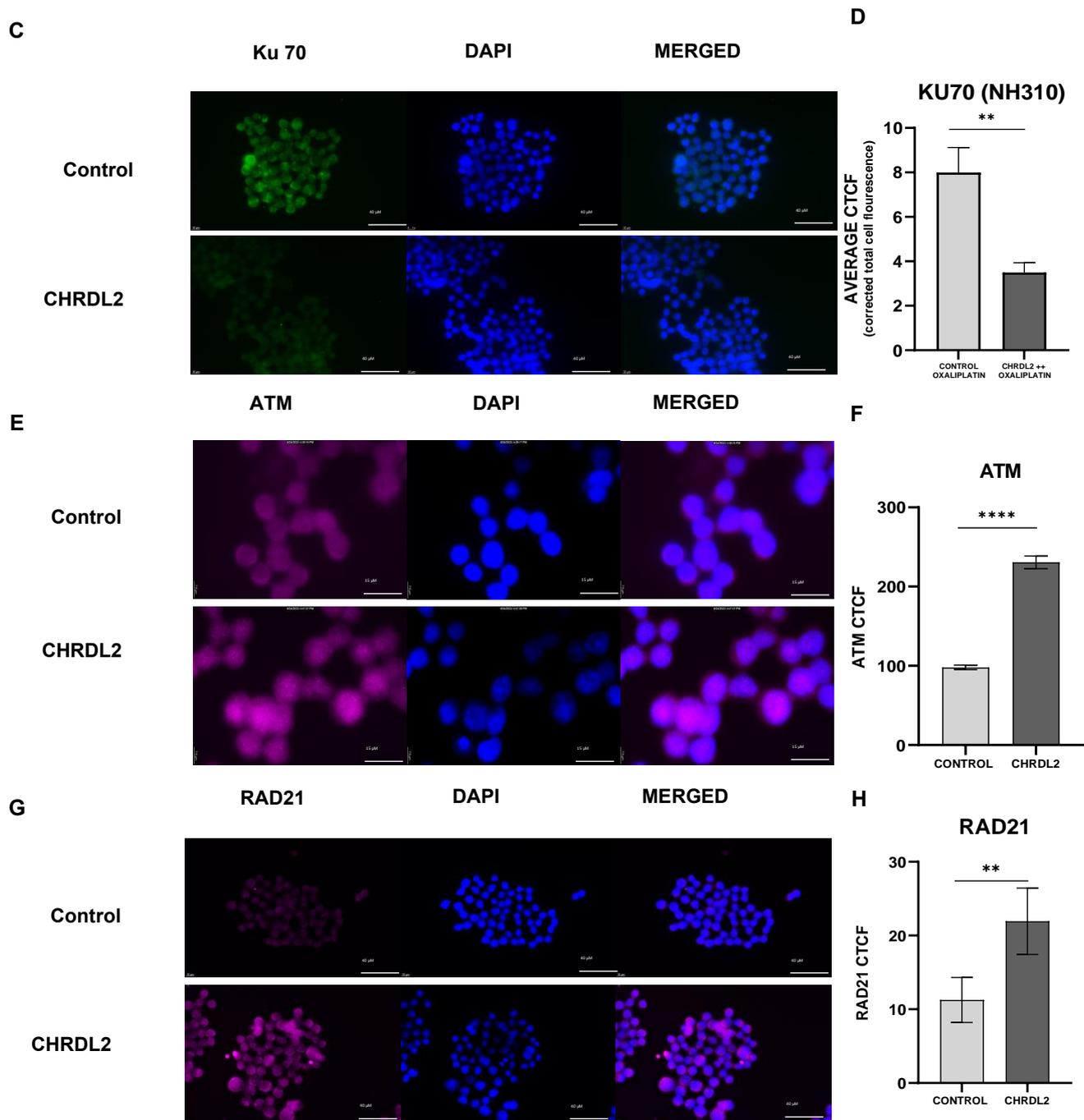


Figure 19: A) Representative immunofluorescence of γ H2AX on COLO320 cells treated with 5 μ M oxaliplatin at 24, 48 and 72 hours. B) Quantification of γ H2AX foci in COLO320 cells overexpressing CHRD2 treated with 5 μ M oxaliplatin at 24, 48, and 72 hours. Cells were treated with DMSO control reagent, or Doxycycline to induce CHRD2++ overexpression. T-test: 24hs $P < 0.0001$, 48hrs $P < 0.01$, 72hrs $P < 0.0001$. $N = 3$. C) Immunofluorescence staining of Ku70 in COLO320 cells treated with CHRD2 overexpression after 72hrs. D) Quantification of Ku70 immunofluorescence in COLO320 cells. Given as CTCF. T-test: $P < 0.01$. $N = 3$. E) Immunofluorescence staining of ATM in COLO320 cells treated with CHRD2 overexpression. F) Quantification of ATM immunofluorescence in COLO320 cells. Given as CTCF. T-test: $P < 0.01$. $N = 3$. G) Immunofluorescence staining of RAD21 in COLO320 cells treated with CHRD2 overexpression. H) Quantification of RAD21 immunofluorescence in COLO320 cells. Given as CTCF. T-test: $P < 0.01$. $N = 3$. Error bars given as \pm SD.

We have further demonstrated the ability of CHRDL2 overexpression to reduce DNA damage during chemotherapy by alkaline comet assay, as observed in figure 20. Comet assays utilize gel electrophoresis of lysed cell DNA, to measure the total DNA damage by strand breaks of a cell, represented as longer comet “tails”. COLO320 cells were treated in the same manner with IC25 Oxaliplatin to that of our γ H2AX and Ku70 assay. We observed cells with CHRDL2 overexpression had shorter “tails” to their comets, showing less fragmented or damaged DNA. Quantification using ImageJ confirmed this, with CHRDL2++ cells having significantly decreased tail lengths at 50 μ M compared to control cells at 75 μ M ($P < 0.0001$).

Furthermore, γ H2AX is also known to accumulate during cellular senescence [240]. So, to test whether our cells are upregulating DNA damage repair pathways or are entering senescence we analysed P53 (a known senescence marker) expression. However, since we found no difference in P53 expression (Figure 20 C, D) in our CHRDL2 overexpressing cells it is more likely that upregulation of DNA damage pathways in CHRDL2++ cells protect against DNA damage by chemotherapy, particularly when compared with our cell cycle data.

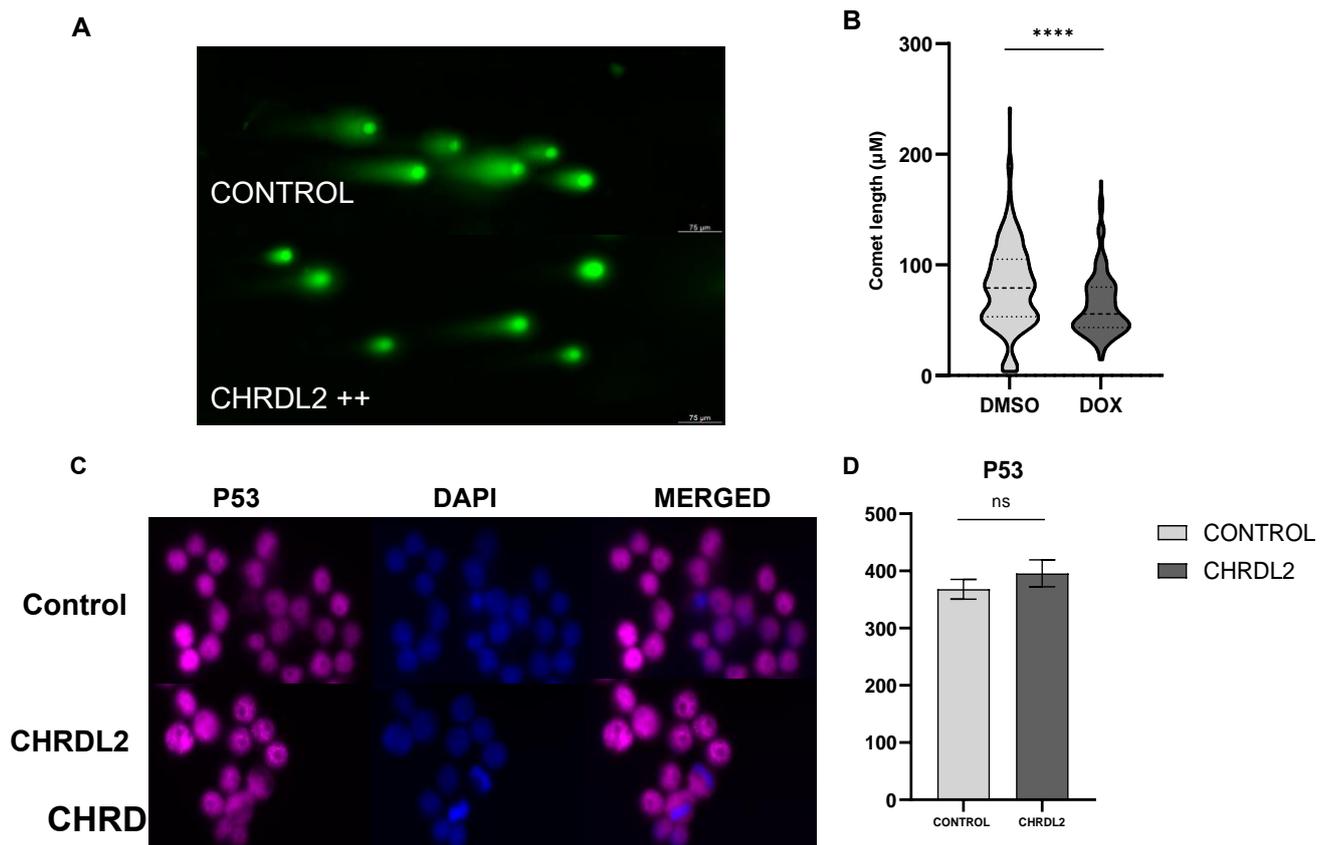
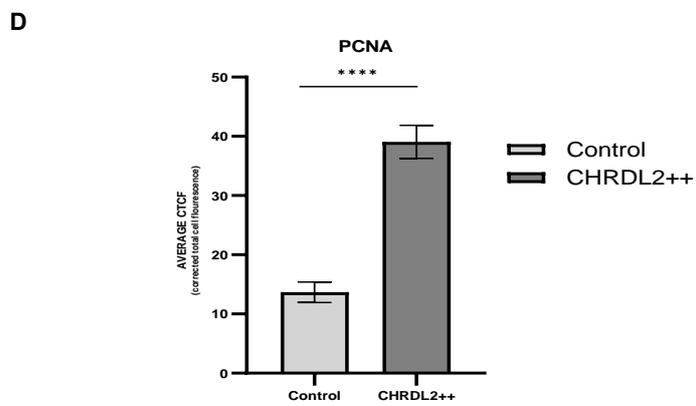
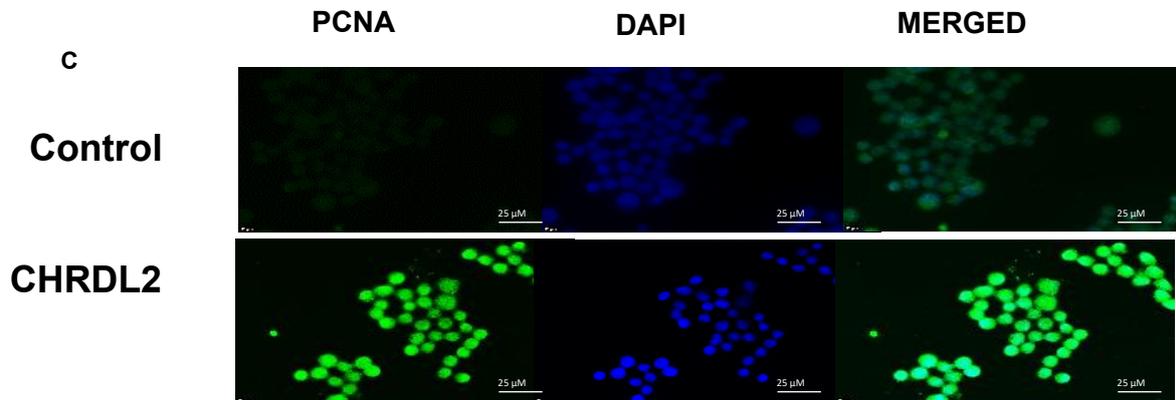
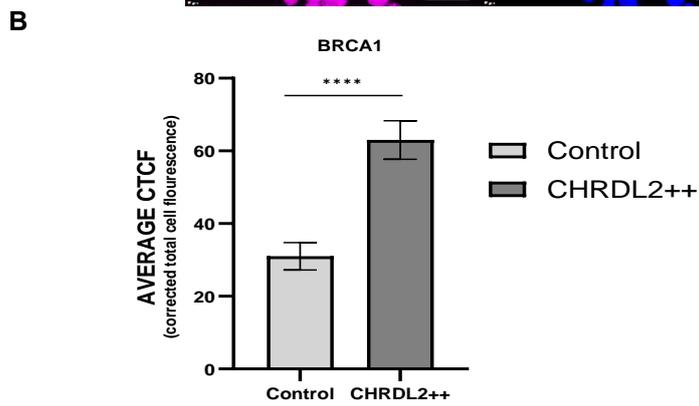
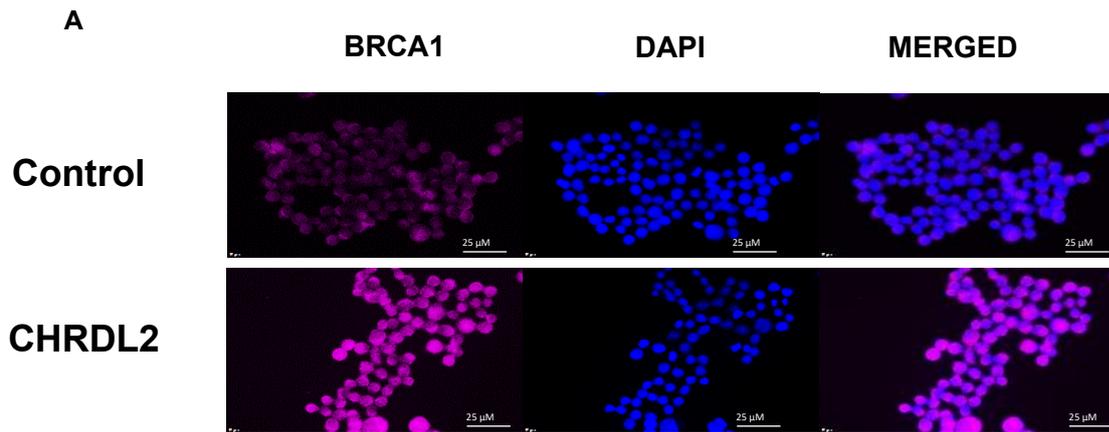


Figure 20: A) Comet assay of RKO cells treated with IC50 Oxaliplatin. Cells were then treated with CHRDL2 ++ overexpression or a control. B) Quantification of Comet assay, T-test $P < 0.0001$. $N = 3$. Quantification carried out using Image J. C) Immunofluorescence staining of P53 on COLO320 cells treated with 5 μ M oxaliplatin. D) Quantification of P53 in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce CHRDL2++ overexpression. NS. $N = 3$. Error bars given as \pm SEM.

Therefore, we investigated expression of repair proteins in CHRDL2 overexpressing cells was analysed. COLO320 cells were given IC25 of Oxaliplatin, and treated with either CHRDL2++ overexpression, or a DMSO control. Immunofluorescence staining was then performed on selected repair proteins.

BRCA1, which recruits RAD51 and is instrumental in the repair of DSBs by HR (Figure 18) displayed a significant increase of 100% fold change in expression during CHRDL2 overexpression (Figure 21 A, B) ($P < 0.001$). The PCNA protein, which is involved in gap filling during single-strand break repair, was also found to be significantly overexpressed with a 400% increase in fold change ($P < 0.0001$) (Figure 21 C, D).

ERCC1, which works as part of the excision complex during repair of DNA adducts, although weakly, was also seen to be overexpressed by CHRDL2 with a 50% increase (Figure 21, E, F) ($P < 0.05$). Furthermore, ARTEMIS, which aides in stabilising DNA stands for exonuclease activity of damage DNA during NHEJ, was also found to be overexpressed by a 200% increase in fold change (Figure 21 G, H) ($P < 0.05$). Together, these expression analyses of these proteins represent an upregulation of all DNA repair pathways that are recruited during repair of DNA after chemotherapy exposure.



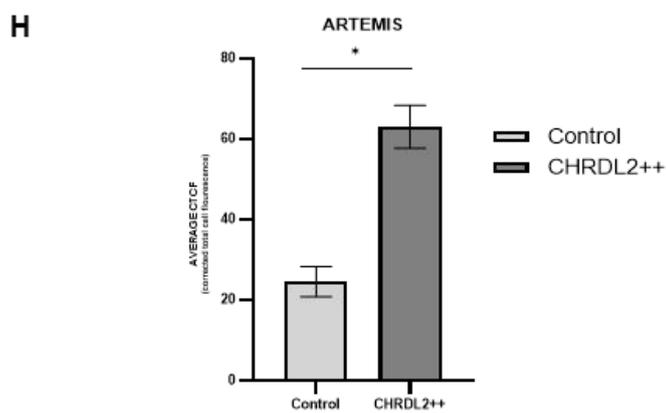
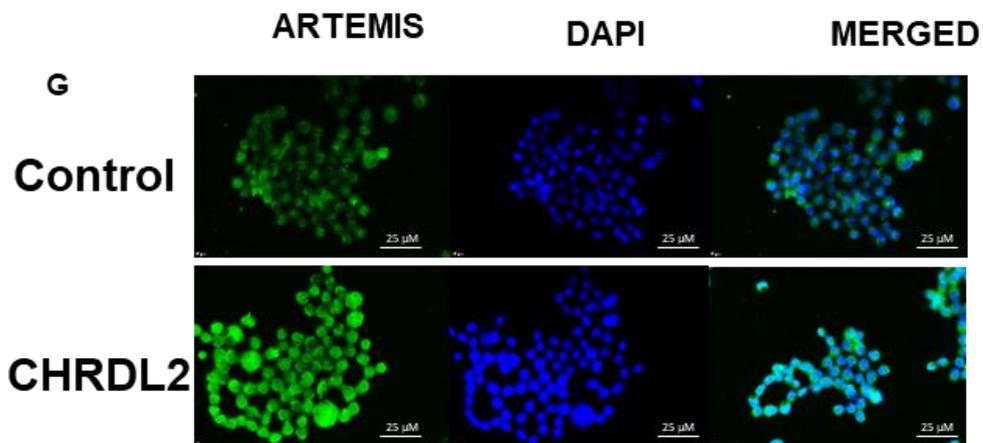
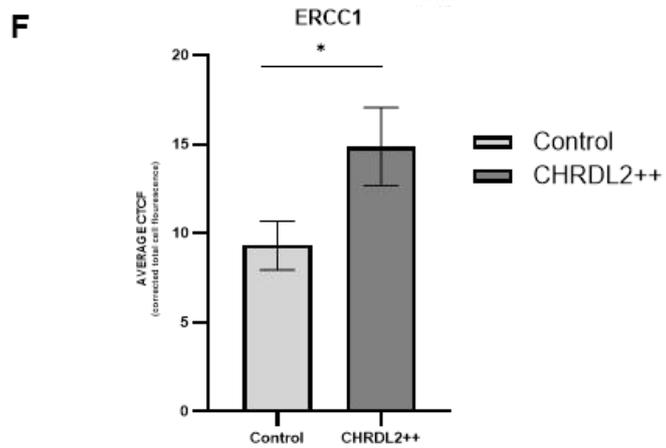
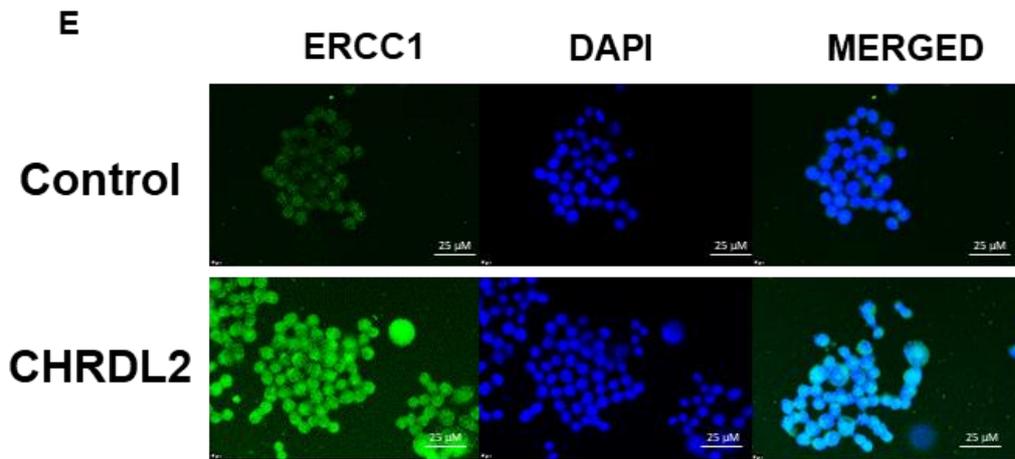


Figure 21: A) Immunofluorescence staining of BRCA1 on COLO320 cells treated with 5 μ M oxaliplatin. B) Quantification of BRCA1 in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce CHRDL2++ overexpression. $P < 0.0001$ N=3 T-test. C) Immunofluorescence staining of PCNA on COLO320 cells treated with 5 μ M oxaliplatin. D) Quantification of PCNA in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce CHRDL2++ overexpression. $P < 0.0001$ N=3 T-test. E) Immunofluorescence staining of ERCC1 on COLO320 cells treated with 5 μ M oxaliplatin. F) Quantification of ERCC1 in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce CHRDL2++ overexpression. $P < 0.05$ N=3 T-test. G) Immunofluorescence staining of ARTEMIS on COLO320 cells treated with 5 μ M oxaliplatin. H) Quantification of ARTEMIS in COLO320 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce CHRDL2++ overexpression. $P < 0.05$ N=3 T-test. Error bars given as \pm SEM.

Discussion

To see how CHRDL2 overexpression may affect chemotherapy resistance, we tested how overexpression of CHRDL2 affected cell survival when subjugated to these 3 chemotherapy agents. As seen in figure 10 A, in each of our CHRDL2 overexpressing cell lines there was an increase in cell survival when treated with doxycycline to induce CHRDL2 expression compared to the DMSO control. This is evidenced by the IC50 values (Figure 10 B) which on average was 2 times higher in CHRDL2 cells than the control. These data would suggest that CHRDL2 overexpression increases cell survival when exposed to chemotherapy agents. This may in part explain the data presented by Sun. et al, which showed that CHRDL2 overexpression in CRC patients resulted in poorer prognosis, compared to non-CHRDL2 overexpressing counterparts [189].

Next, IC50 values were then compiled, and paired t-tests were performed to allow comparisons between IC50 values for CHRDL2 and the control for each cell line (Figure. 10 C). Paired T-tests were performed due to the variability of IC50 values in each experiment repeat. This variability between experiments can be explained by small variations in serial dilutions, different stocks or the age of stocks of chemotherapy drugs, and number of cells plated and even whether the plated cells came from flasks in log phase or nearing confluency. A heat map of relative increase in IC50 valued can be seen in figure 12, D. Cell line COLO320 had the highest increase in survival when CHRDL2 was overexpressed, shown the biggest difference in IC50 values in all chemotherapy drug groups. COLO320 cells do not have mutations in BMP receptors such as RKO and LS180 cells, and unlike CACO2 cells, do not have damaging mutation in Beta-catenin (CTNNB1) which is a WNT pathway transducer. This pattern of mutations may explain why COLO320 showed the highest survival increase when CHRDL2 was overexpressed, as unlike the other cell lines, the downstream pathway from CHRDL2 inhibition of BMP signalling is not affected. Furthermore, we have theorised that WNT signalling is increased by CHRDL2 overexpression, and COLO320 cells no not have damaging mutations in prominent downstream WNT transducers.

We then looked at cell cycle status of our cell lines when subjected to chemotherapy. The chemotherapy agents used in or experiments all affect DNA synthesis, and so an anticipation of stalled S phase is to be expected. Indeed, this was confirmed by our analysis which showed a greater portion of cells in S phase, compared to that of non-chemo treated cells. Interestingly, CHRDL2 overexpressing cells showed less cells stalled in S phase compared to a control, indicating that these cells are able to overcome DNA damage induced cell cycle arrest. Furthermore, there were a greater percentage of proliferating cells (marked by Ki-67+ status) during CHRDL2 overexpression, indicating higher growth during chemotherapy treatment.

Additionally, analysis of apoptotic cells revealed less cells in early apoptosis in CHRDL2 treated cells compared to a control, indicating that these cells have a greater ability to by-pass chemotherapy induced programmed cell death.

These effects of increased survival during chemotherapy treatment were replicated using conditioned media with secreted CHRDL2. This would indicate that CHRDL2 is able to exert chemotherapy-resistance by paracrine signalling, as well as autocrine, which is the proposed *in vivo* mechanism of CHRDL2 signalling.

The chemotherapy drugs used in this assay use different methods to kill cells, with 5FU inhibiting in pyrimidine nucleotide synthesis, Oxaliplatin forming DNA adducts and DSBs, and Irinotecan acting as a topoisomerase inhibitor [228][241]. There was no pattern regarding drug type in increasing survival with CHRDL2 overexpression, however it appeared that Irinotecan, a topoisomerase inhibitor, showed the least difference in cell survival.

Organoids treated with chemotherapy 5FU also showed the same response, with the number of live organoids far higher during CHRDL2 treatment and chemotherapy compared to a control. Therefore, we can see that an increase in chemotherapy resistance occurs in both cancerous 2D cell lines, as well as non-cancerous organoid models.

We then tested whether CHRDL2 overexpression would also increase survival during irradiation treatment, which is often used in conjunction with chemotherapy in patients with rectal cancer. Again, we saw an increase in cell survival after irradiation treatment at 4 and 6 GY radiation. We propose that CHRDL2 therefore not only increases resistance to chemotherapies, but also to radiation, relaying the anti-apoptosis and pro survival effects of CHRDL2 overexpression.

Collectively, this data supports our stem-cell hypothesis, as it has been previously shown that stem-cells resists chemotherapy and irradiation treatment compared to their differentiated tumour cell counterparts [242]. This is in part because of the responses stem cells are able to manifest following chemotherapy treatment, such as induction of the epithelial to mesenchymal transition pathway, increased expression of drug transporters such as solute carriers, upregulation of DNA repair mechanisms, and an increase in hypoxia signalling through upregulation of HIF1- α [243]. There is also growing evidence that cancer-stem cells (CSCs) can resist DNA damage imposed by radiation and chemotherapy agents, making them more resistant to treatment [244]. Resistance to DNA damage by CSCs has been shown to be likely due to increase DNA damage response (DDR) pathway activation, in which DNA lesions are acknowledged and repair by the DNA repair machinery [244][245].

Within the context on Chemotherapy-induced DNA damage, the DDR pathway includes the repair of double stranded breaks (DSBs) caused by Oxaliplatin and Irinotecan, through homologous recombination (HR) or non-HEJ (NHEJ), as well as repair of intra-strand crosslinks caused by Oxaliplatin through nucleotide excision-repair (NER) or interstream crosslink (ICL) repair. Repair of DSBs is initiated by either ATM (HR) or ATR (NHEJ) (Figure 18). Coincidentally, previous studies have shown that POLD3, the gene closely related to CHRDL2 overexpression, is crucial for both repair mechanisms as discussed in chapter 6 [113].

In light of this, we sought to measure whether CHRDL2 overexpression decreased DNA damage, and increased DNA damage repair pathways, when exposed to chemotherapy. When treated with the IC25 value of Oxaliplatin, CHRDL2 overexpression in COLO230 cells reduced the amount of DNA damage visible, shown by a reduction in H2AX foci (Figure 19, A). At 24hrs the number of foci representing DSBs was reduced to 2/3rds compared to control cells, but after 72hrs there was nearly almost no H2AX foci present, indicating a complete clearing of DSBs by CHRDL2 overexpressing cells ($P < 0.0001$, $P < 0.0001$) (Figure 19, B). We also found a decreased expression of Ku70, a marker of DSB by NHEJ (Figure 19, C).

DNA damage repair of DSBs by HR is initiated by ATM, which recruits repair machinery to the site of damage. During CHRDL2 overexpression and treatment with IC50 oxaliplatin, we saw a dramatic increase of ATM, indicating an increase in activation of downstream repair pathways (Figure 19 E). We also saw an increase in RAD21, which is also instrumental in DNA repair initiation (Figure 19 F).

H2AX accumulation is also known to be modulated during cellular senescence, so to confirm that our observed reduction in γ H2AX was due to increased DNA repair and not cellular senescence, we examined the expression of senescence and cell cycle arrest marker, P53, in cells treated with the same chemotherapy (Figure 20 C). However, since we found no difference in P53 expression in our CHRDL2 overexpressing cells it is likely that upregulation of DNA damage pathways in CHRDL2++ cells can protect against DNA damage by chemotherapy.

General DNA damage by both single strand and double strand breaks was also analysed. Comet assays are a commonly used assay, with comet "tail" length used as a measure of fragmented DNA. As seen in figure 20 A, B, CHRDL2 overexpressing cells has comet tails during chemotherapy treatment compared to a control, indicating that CHRDL2 protects against both forms of strand break induced DNA damage. Collectively, this data supports our previous hypothesis that CHRDL2 may increase stem-cell potential of CRC cells, as evidenced by increased DDR activation and reduced DNA damage during conventional therapies.

It should also be considered that these chemotherapy agents target proliferating cells. When taken with our proliferation data, it can be shown that this increase in cell survival is not due to hyper-proliferation of the CHRDL2 cells, and therefore must be through another mechanism. Another process in which CHRDL2 may promote cell survival during chemotherapy treatment, is that CHRDL2 may force the cell into a slower growing more resilient phenotype when under cellular stress. As chemotherapy agents target proliferating cells, if a cell is slow growing it may be more resilient to chemotherapy. This is supported by our low glucose proliferation data, where CHRDL2 overexpression decreased proliferation in low glucose stressed conditions.

To further ensure that survival during CHRDL2 overexpression is due to upregulation of DNA repair pathways, we analysed the expression of known DNA repair pathway proteins. WE found regulation of BRCA1, and ARTEMIS, which are involved in the repair of DSBs by HR and NHEJ respectively. We also found upregulation of ERCC1, which resolves crosslinks and DNA adducts formed by chemotherapy drug oxaliplatin. Finally, we found upregulation of PCNA, which resolved single strand breaks. Together, these represent regulation of the 4 DNA repair pathways used to resolve DNA damage after chemotherapy agents,

Limitations and future work

As with data in chapter 3, a limitation of this study is the use of human CHRDL2 on murine organoids, as previously discussed. Despite high similarity between murine and human CHRDL2, the addition of mouse CHRDL2 on murine organoids would be a more faithful model to uncover the full effects of CHRDL2 on chemotherapy resistance. Western blotting could also be performed to strengthen the findings of changes in protein expression in DNA repair pathways.

Conclusion

Collectively, we have shown that CHRDL2 upregulation increases resistance to common chemotherapies to treat CRC. Furthermore, we have found that CHRDL2 also increases survival of radiation, which is often used in conjunction to chemotherapy to treat CRC. This is evidenced by increased IC50 values in our 4 experimental cell lines, increased survival in organoid models, and reduced DNA damage shown by comet assay. Furthermore, CHRDL2 overexpressing cells display a lower stalling at the S/G2 checkpoint, and higher expression of proliferation markers following chemotherapy treatment.

Additionally, we have found that CHRDL2 increases clearing of DNA damage, shown by reduction in γ H2AX foci over time, and a reduction in Ku70 staining. This suggests that CHRDL2 does not just decrease DNA damage but increases resolution of DNA damage.

To achieve this, we have shown that CHRDL2 overexpression results in overexpression of common DNA repair mechanisms, including HR, NHEJ, BER, and NER. This is shown through upregulation of repair proteins ATM, RAD21, BRCA1, ERCC1, PCNA, and ARTEMIS.

This increased ability of CHRDL2 overexpressing cells to survive chemotherapy, and upregulate DNA repair pathways, is indicative of the cancer-stem cell phenotype, in which cancer stem cells increase in immortality through upregulation of the pathways we have shown. This supports our data shown previously that through BMP inhibition, CHRDL2 increases stem-cell properties through increased WNT signalling.

Chapter 5: Gene expression analysis and mechanisms of CHRDL2

Introduction

In this chapter, we have utilised RNA sequencing methods and online databases to analyse CHRDL2 expression in patient samples, and the effects of CHRDL2 overexpression in 2D cell lines. Previously, we have shown that CHRDL2 enhances the WNT signalling pathway, and decreased BMP signalling. However, other pathways affected are yet to be elucidated.

The WNT signalling pathway includes canonical, and non-canonical signalling methods. The canonical WNT signalling pathway, which we have shown to be enhanced upon CHRDL2 overexpression, signals through sequestering of the β -catenin complex, leading to β -catenin localisation to the nucleus. Here, β -catenin activates target genes via TCF/LEF transcription factors. The canonical WNT pathway has been shown to mainly effect cellular proliferation, whereas the non-canonical WNT signalling pathway, such as the WNT/ Ca^{2+} pathway, tends to affect cell polarity and migration. These two pathways form a network of mutual regulation to affect cell fate [246].

Once the WNT/ β -catenin signalling pathway has been activated, TCF/ β -catenin can act through enhancers that can be hundreds of kilobases away from the proximal promoters of WNT targets. This can be seen in WNT gene regulation where chromatin looping allows enhancer-promoter communication [247]. Some of the most important WNT targets, include cyclin D1 and c-myc, which controls the G1 to S phase transition in the cell cycle, leading to cell division and proliferation [248].

WNT signalling has also been shown to affect pathways involved in migration, invasion and metastasis, as well as increased longevity and enhanced DNA repair pathways [249]. Importantly, WNT signalling has been shown to influence populations of cancer stem-cells, which have an increased capacity of these properties. For instance, one of the hallmarks of cancer stem cells is the ability to maintain telomers through expression of the TERT gene. TERT was found to be directly enhanced by binding of β -catenin to its promoter region, confirming the role of WNT signalling in enhancing cell longevity [250].

WNT activation has also been widely shown to directly effect the expression of LGR5, which is a marker of intestinal stem cells. LGR5+ cellular expansion has been shown to require RAC1, which in turn drives ROS production and thereby activates NF κ B signalling, which enhances WNT signalling [251]. This suggests a co-activation of WNT and NF κ B, leading to enhanced ROS and stem-like qualities in intestinal cells.

SNAI2, another WNT target, is an important transcription factor for EMT initiation. Cytoplasmic SNAI2 concentration is kept in check by GSK3 β phosphorylation and subsequent ubiquitinylation by β -TrCP. WNT signalling stabilizes SNAI2 by inhibiting GSK3 β kinase activity and initiates EMT transcriptional activities, leading to enhanced metastatic capacity [252].

WNT signalling has also been shown to strongly interact with the Hippo pathway, which negatively regulates Yes-associated protein (YAP) and Tafazzin (TAZ). The YAP/TAZ proteins have been shown to be elevated in CRC, and have been identified as oncogenes which are able to reprogram cancer cells into cancer stem-cells [253]. This may be due to YAPs ability to overcome contact inhibition, which allows hyperproliferation of cancer cells. There is evidence that WNT signalling directly stimulates YAP, as YAP is a target for degradation by the β -cat destruction complex [254]. Interestingly, CHRDL2 has been shown to stimulate the YAP/TAZ pathway in gastric cancer, which is a potential method in which CHRDL2 effects gene regulation upon WNT activation[147]

It is clear that overexpression of the WNT signalling pathway can have a multitude of effects on cancer cell fate. Elucidating pathways in which CHRDL2 overexpression may exert its effects through WNT activation and other potential pathways, is key to deciphering the potential role CHRDL2 may play as an oncogene in tumour formation.

CHRDL2 predicts poor survival in late-stage CRC.

Analysis of CHRDL2 expression in patient samples were performed. Data was provided through online publicly available data sets [255], totalling 1167 patient samples. CHRDL2 was found to be overexpressed in tumours from patients with colorectal adenocarcinoma (COAD) compared to normal tissues (Figure 22 A). Overall survival for patients with high CHRDL2 overexpression was not significantly different from that of low CHRDL2 overexpression despite a trend of poorer survival($P<0.078$)($N=1167$)(Figure 22 B). However,, high expression of BMP5, in which CHRDL2 binds to prevent signalling, significantly increases survival of patients ($P<0.0046$) ($N=1167$) [144].

Next, we looked at stage-specific survival of patients with CHRDL2 high expression. In the earliest stages of disease, stage 1 and 2, CHRDL2 high expression did not affect survival($P<0.14$, $P<0.013$) ($N=90$, $N=510$). Early stages of CRC treatment rely on the use on tumour resection, as at this stage polyps have formed that have not yet metastasised. Our findings here may reflect our previous data that CHRDL2 overexpression decreased cell

proliferation, which is an important factor in the earliest stages of CRC progression, and would therefore not contribute to lower survival.

Interestingly, in the later stages of disease, at stage 3, CHRDL2 high expression begins to show a trend of decreased patient survival, although again not significantly ($P < 0.34$) ($N = 408$). At the latest stage of disease, stage 4, where typical treatment relies on the use of chemotherapy and radiotherapy over surgical removal, we found a drastic decrease in patient survival with CHRDL2 high expression (Figure 22 C). CHRDL2 high expression in these patients significantly reduced patient survival ($P < 0.05$) ($N = 71$). These findings may correspond to our previous findings that CHRDL2 overexpression in 2D cell lines resulted in increased survival during chemotherapy and radiotherapy.

Next, we looked at how CHRDL2 expression may affect patient survival in different CMS classified tumours. As seen in figure 22 D, CRC tumours can be classified into 4 CMS groups, CMS1, which characterises immune activated tumours which are hypermutated, MSI+, and CIMP+, CMS2, which is the canonical tumour phenotype with increased WNT activation, CMS3 tumours which are metabolic dysregulated, and CMS4 tumours which arise from mesenchymal cells and display increased stromal infiltration. For this analysis we selected patient samples in stage 3 and 4, and these represented stages in which CHRDL2 high expression resulted in decreased patient survival.

As seen in figure 22, E, patient survival was assessed in these four CMS subtypes with CHRDL2 high or low expression. CMS1, 2 and 4 tumours with high CHRDL2 overexpression did not show decreased patient survival. On the other hand, CMS3 tumours showed that CHRDL2 overexpression resulted in higher patient survival $P < 0.03$. This data appears to show no correlation between tumour subtypes and patient survival with CHRDL2 overexpression.

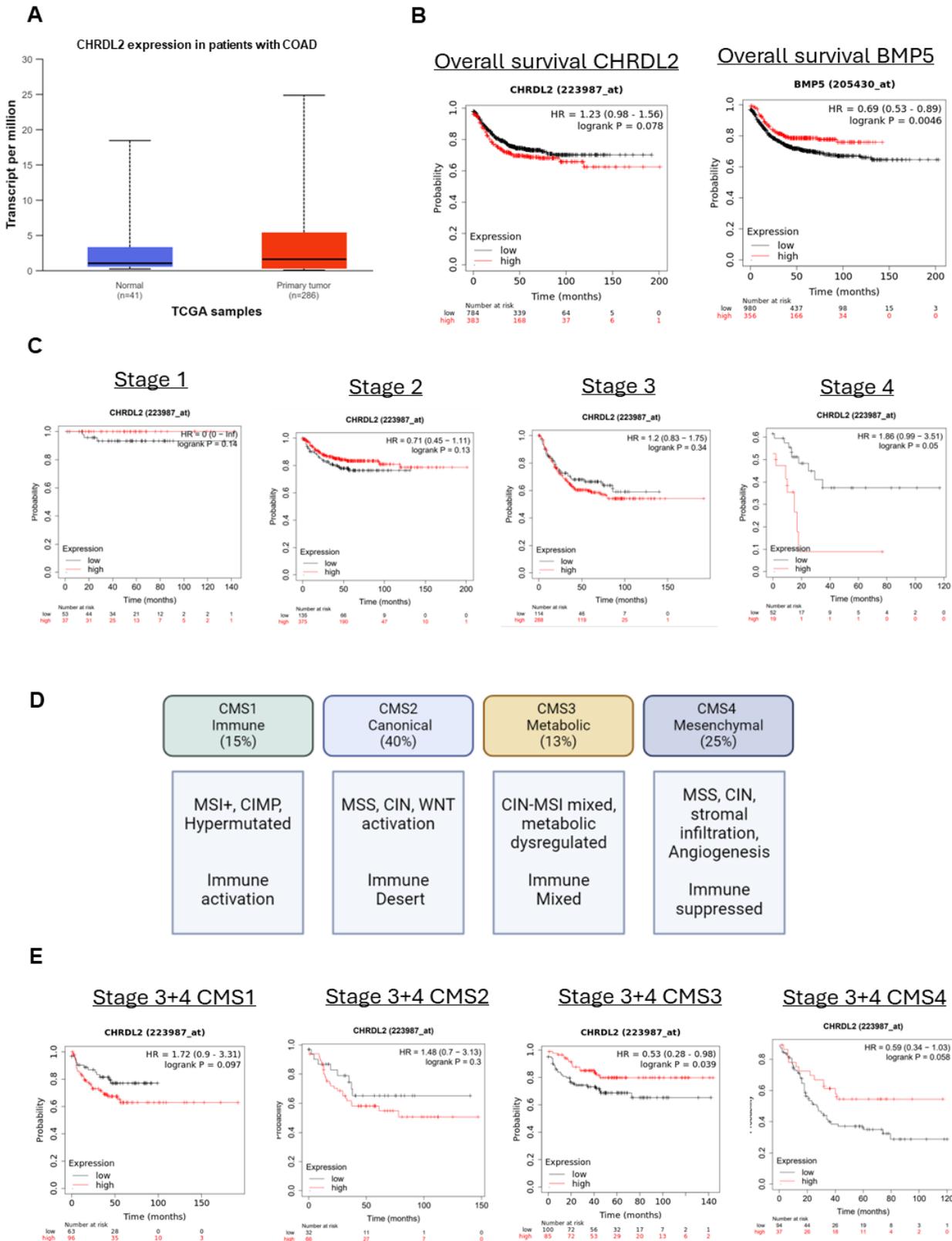
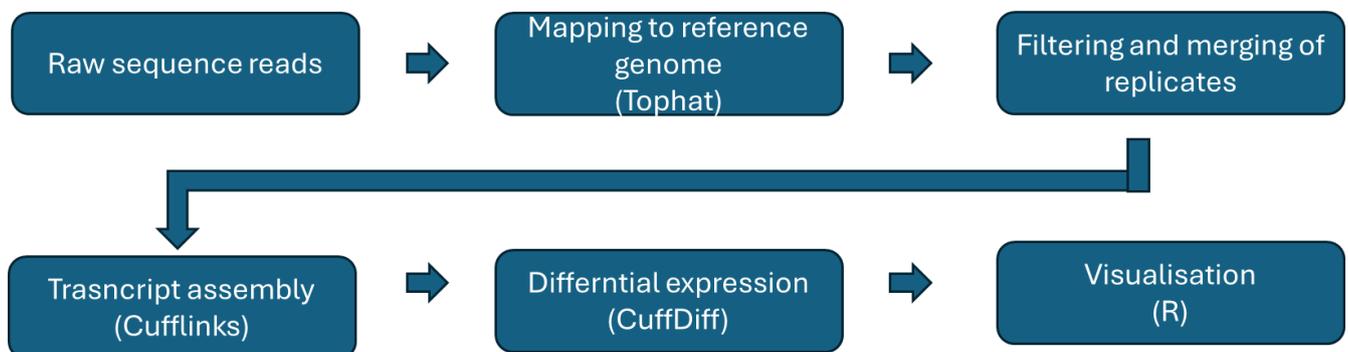


Figure 22: A) CHRD2 gene expression in primary Colorectal adenocarcinoma (COAD) and normal tissue. B) Overall patient survival during CHRD2 and BMP5 overexpression P<0.078, P<0.004. N=1167. C) Stage specific patient survival with CHRD2 high and low expression. Stage 4 P<0.05. Stage 1 N=90, stage 2 N=510, stage 3 N=408, stage 4 N=71. D) CMS classification of CRC subtypes. E) Stage 3+4 patient survival with CHRD2 high and low expression within different CMS subtypes. Stage 3+4 CMS3 P<0.039. N=185.

CHRD_{L2} induces cancer-related genes during overexpression.

Next, to analyse genes and genetic pathways affected by CHRD_{L2} overexpression, we performed RNAseq analysis on our doxycycline inducible CACO2 cell lines. RNAseq analysis allowed us to elucidate changes in gene expression when *CHRD_{L2}* was overexpressed. CACO2 cells underwent doxycycline-inducible expression of *CHRD_{L2}* and were treated with doxycycline at LOW MED and HIGH concentrations, corresponding to our CHRD_{L2}, CHRD_{L2}⁺, and CHRD_{L2}⁺⁺ treatment previously used (Figure 23 A), as well as a DMSO control for 3, 5, and 7 days. Cells were harvested at these time points and the levels of RNA present in the cell were measured through RNA seq analysis. Time points were compiled into 3 replicates to control for temporal variation in gene expression. The workflow used to process raw sequence reads can be seen below:



After obtaining sequence read, differential expression analysis was performed. Differential expression analysis compared differentially expressed genes in our doxycycline-treated groups to the DMSO control (Figure 23, B).

A heatmap was produced to visualise differential expression of significant gene changes by CHRD_{L2}. Figure 19 C shows a heatmap of differentially expressed genes at significance level of $P < 0.05$ and $P < 0.01$ (Figure 23 D). As seen in figure 23 C, the LOW treatment most closely resembled the DMSO control treatments, whereas the MED and HIGH treatments cluster closely together. CHRD_{L2} overexpression can clearly be shown to increase (dark brown colour) with increased doxycycline treatment, which reflects qPCR and western blotting data from our experimental cell lines seen in Figure 4 A. Increasing the significance threshold to $P < 0.01$ in panel D, reduces the number of genes common in all 3 pathways. We also have compiled our 3 DMSO replicates to compare to our LOW, MED, and HIGH treatment groups seen in panel E we can clearly see a group of genes that are significantly differentially expressed by CHRD_{L2}, such as NCALD, TFF1, REG4.

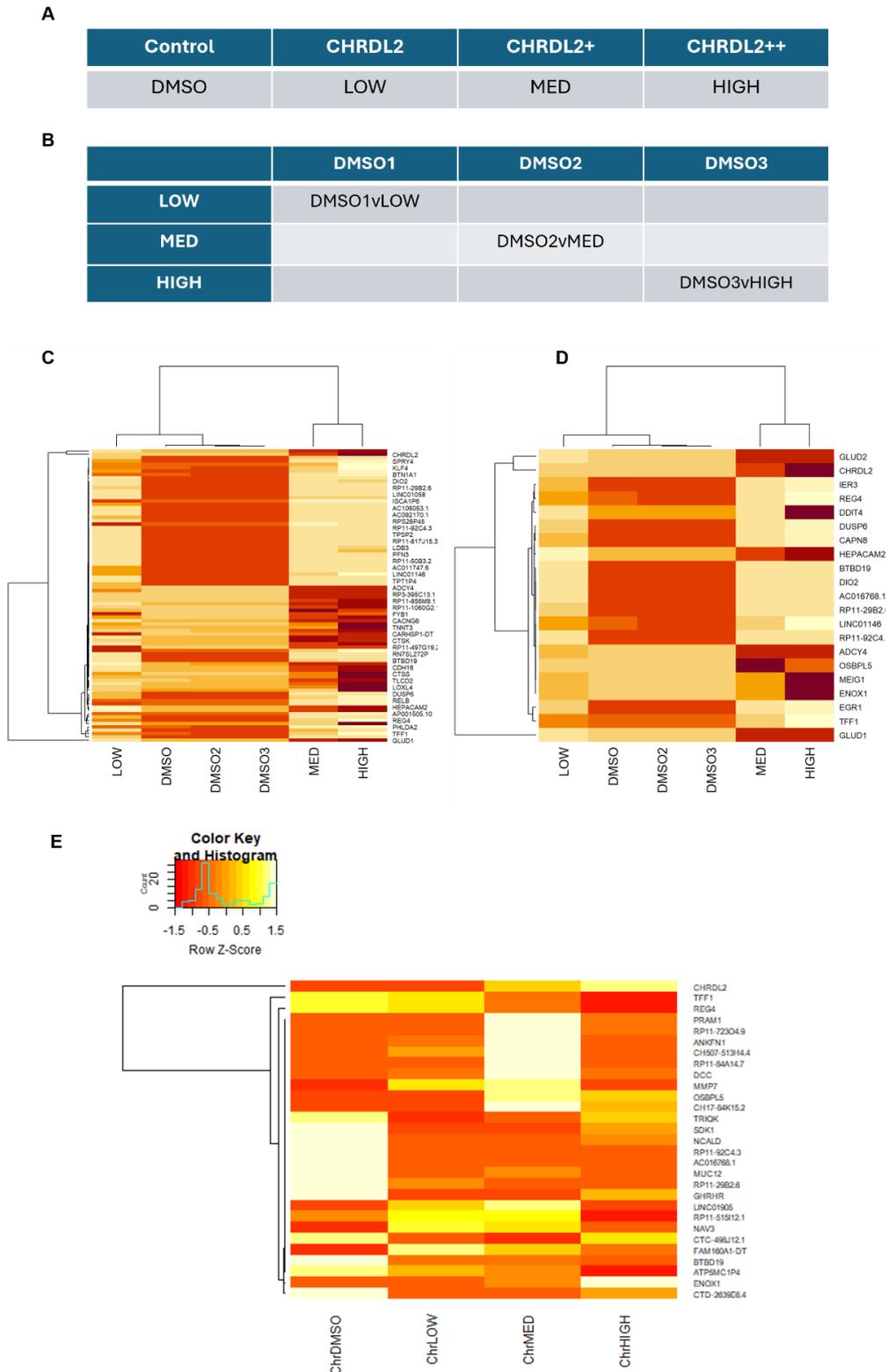


Figure 23: A) Experimental design of treatment through CHRDL2 overexpression. B) Experimental design of differential expression analysis. Differential expression was accomplished through comparing DMSO control V CHRDL2 overexpressing cells. C) Heatmap of differentially expressed genes by CHRDL2 ($P < 0.05$). D) Heatmap of differentially expressed gene by CHRDL2 $P < 0.001$. E) Heatmap of differentially expressed gene by CHRDL2, with DMSO control replicates merged. $P < 0.01$.

Next, volcano plots were generated to display differentially expressed genes by CHRDL2 overexpression. These graphs plotted the fold change of the gene against the pvalue associated (Figure 24). Plots were generated for LOW, MED and HIGH treatments. CHRDL2 can be seen on the right hand side of each graph with a positive fold change. Many genes can be observed to the far right or far left of these plots, as these designate genes that have a fold change of “infinite”. This occurs when a sample in either control and treated groups has a FPKM value of 0, resulting in a infinite fold change. Therefore, the differences of these genes may not be as prominent as those toward the middle of the volcano plots.

In the LOW treatment, upregulation of genes *DDIT4* and *NAV3* can be observed. In the MED treated groups, we can observe upregulation of *GLUD1*, *GLUD2*, *CTSK*, *PNCK*, *ENOX1*, *MMP7* as well as others, and downregulation of *DDIT4*, *NCALD*, *DUSP6*, *REG4*, and *MYC*. In the HIGH treated groups, we can see upregulation of *ENOX1*, *MAPK10*, *CEBPB*, *PCK2*, *DDIT4*, *GLUD1*, and downregulation of *CXCL8*, *MUC12*, *TFF1*, *REG4*, *APOB*, among others.

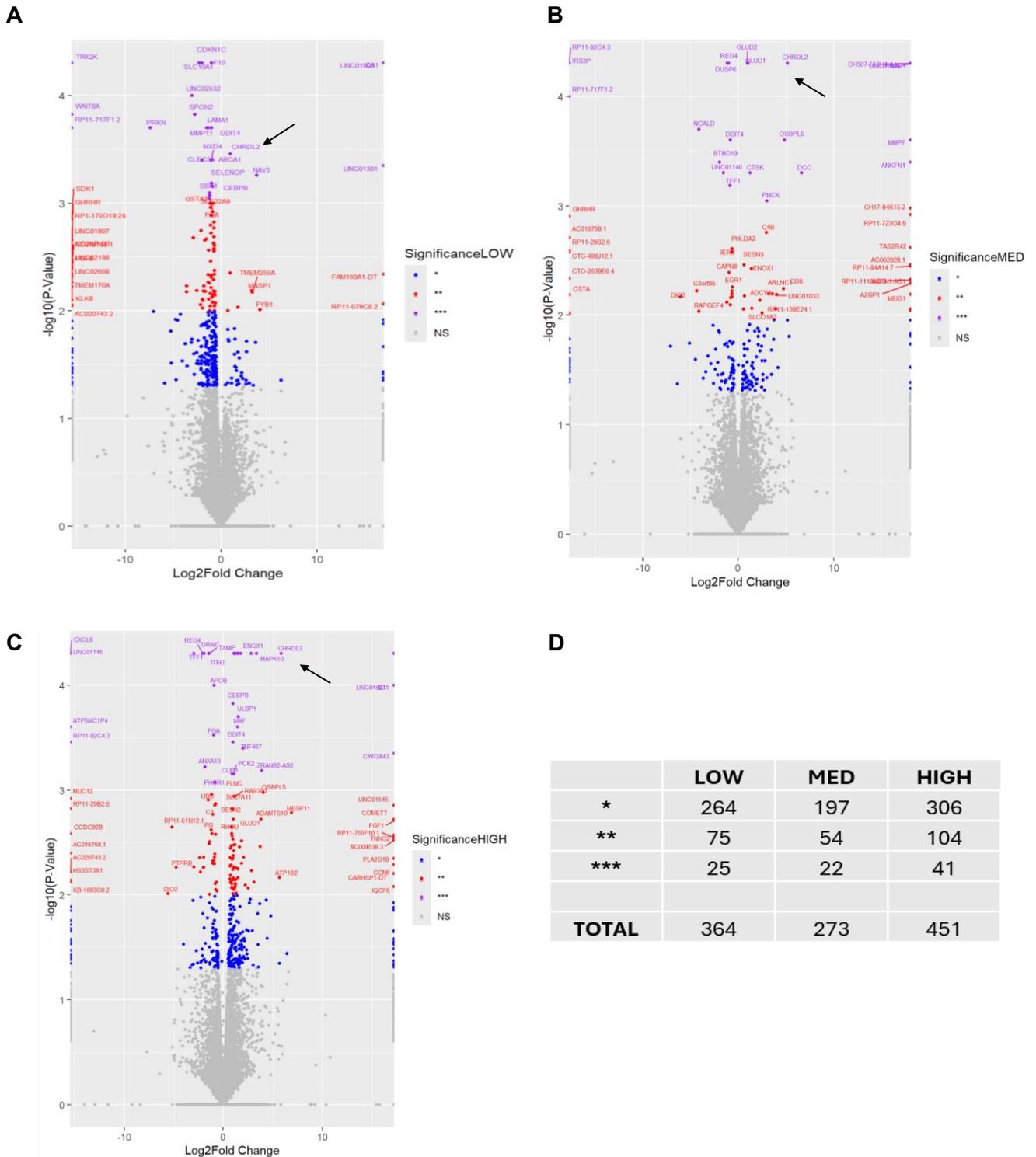


Figure 24: A) Volcano plot of differentially expressed gene by CHRDL2 LOW overexpression. Genes plotted by Log2 fold change in gene expression versus log10 of the P value for the gene. Gene sorted by significance level: * = 0.05, ** = 0.01, *** = 0.001. B) Volcano plot of differentially expressed gene by CHRDL2 MED overexpression. Genes plotted by Log2 fold change in gene expression versus log10 of the P value for the gene. Gene sorted by significance level: * = 0.05, ** = 0.01, *** = 0.001. C) Volcano plot of differentially expressed gene by CHRDL2 HIGH overexpression. Genes plotted by Log2 fold change in gene expression versus log10 of the P value for the gene. Gene sorted by significance level: * = 0.05, ** = 0.01, *** = 0.001. D) Counts of differentially expressed genes at significance levels *, **, and *** in all groups.

Identification and validation of differentially expressed genes by CHRDL2.

We then compiled differentially expressed genes from the MED and HIGH treated groups. 76 and 145 differentially expressed genes were identified in the MED and HIGH groups respectively, ($P < 0.01$) (Figure 25). LOW treated groups were eliminated from this analysis as they showed high similarity to the DMSO controls, and did not cluster next to MED and HIGH groups in the heatmaps. From this we selected the 21 genes that were differentially expressed in both MED and HIGH vs control cells for downstream analysis (Figure 25).

We then sought to validate differential expression by CHRDL2 of these target genes through qPCR analysis. This analysis was performed on COLO320 cells with CHRDL2 overexpression. Cells were harvested 72hrs post treatment and RNA extracted. Gene expression was then quantified, as seen in Figures 26 and 27.

Interestingly, Trefoil factor 1 (*TFF1*) known to inhibit proliferation, migration and invasion was downregulated in the RNAseq data and in qPCR. Glutamate Dehydrogenase 2 (*GLUD2*), a glycolysis related gene, Phosphoenolpyruvate Carboxykinase 2 (*PCK2*), involved in mitochondrial respiration and elevated in tumours, and DNA damage inducible transcript 4 (*DDIT4*), associated with advanced CRC, were all upregulated in both RNAseq and qPCR data. Stem cell markers *LGR5* and *LGR6*, as well as B lymphoma Mo-MLV insertion region 1 homolog (*BMI1*) were also upregulated in our qPCR data, with *LGR6*, a WNT transducer, also upregulated in RNAseq data in the CHRDL2++ treatment (Figure 26).

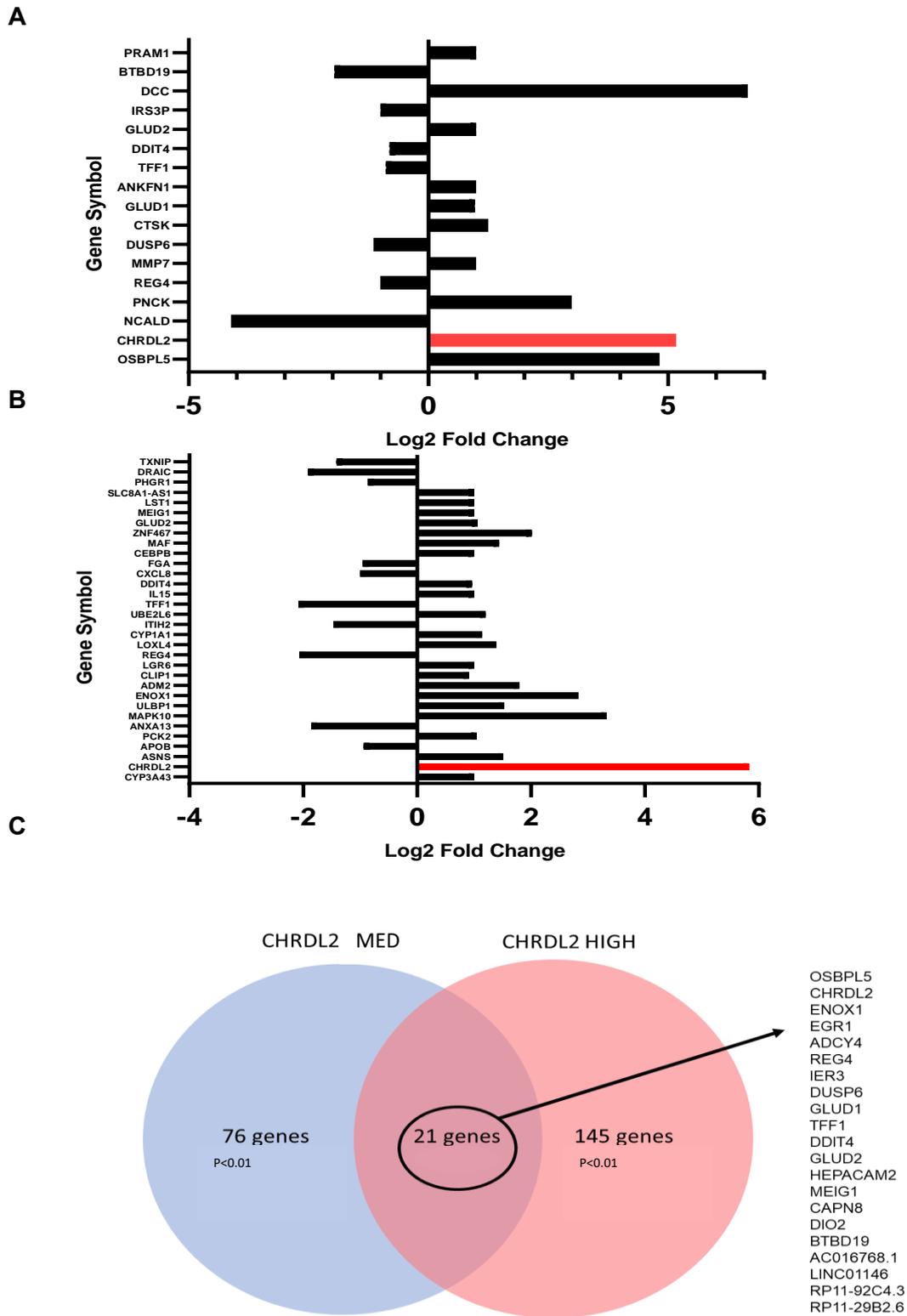


Figure 25: A) Bar-plot of significantly differentially expressed genes by CHRDL2+ and CHRDL2 ++. Genes included pass the threshold of $P < 0.01$ for CHRDL2+ and $P < 0.01$ from CHRDL2 ++. B) Intersect of highly differentially expressed genes in both the CHRDL2 + and CHRDL2++ treated groups $P < 0.01$. 21 genes were differentially expressed in both groups.

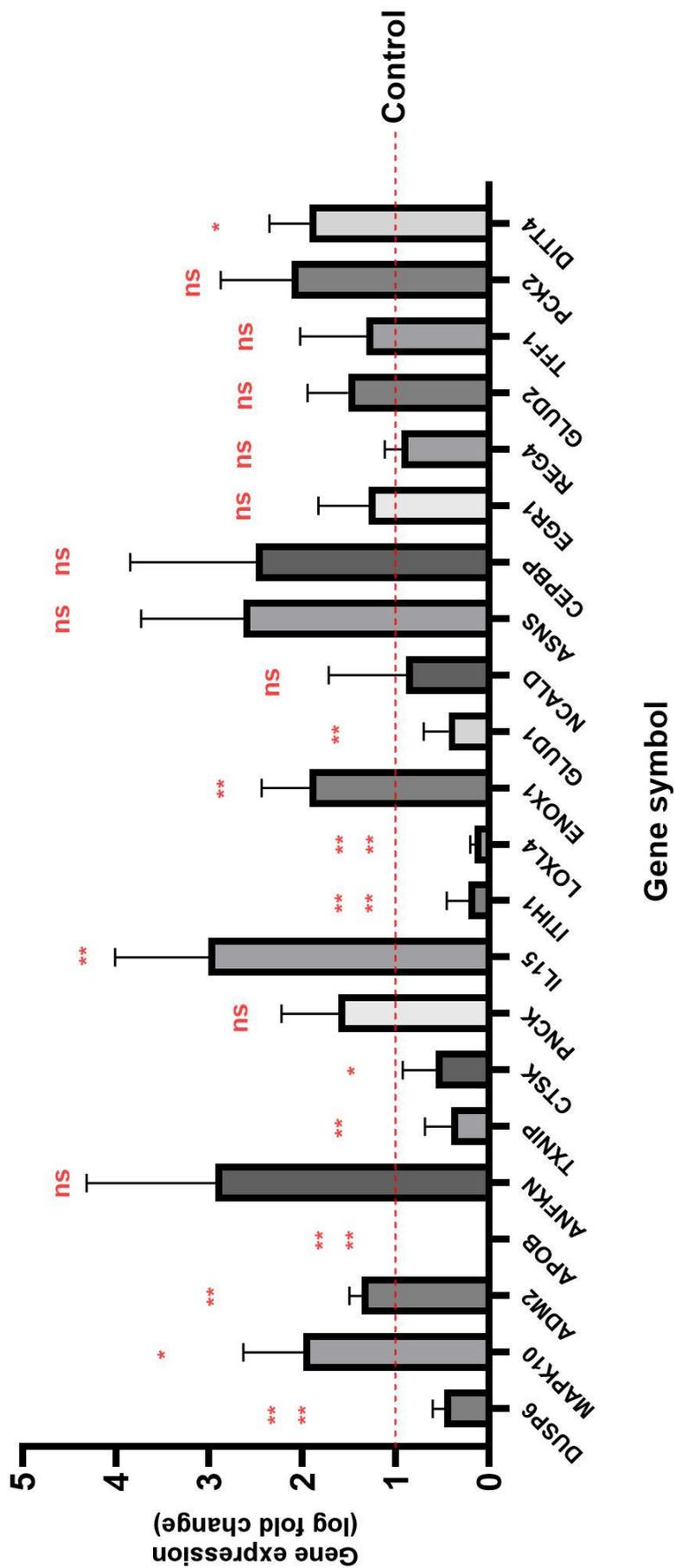


Figure 26: qPCR analysis of differentially expressed gene by CHRDL2 highlighted in RNAseq data. DUSP6 P<0.00014. MAPK10 P<0.019. ADM2 P<0.0083. APOB P<0.000. ANFKN P<0.059. TXNIP P<0.00533. CTSK P<0.021. PNCK P<0.0652. IL15 P<0.00661. ITIH1 P<0.000. LOXL4 P<0.000. ENOX1 P<0.00331. GLUD1 P<0.00142. NCALD P<0.387. ASNS P<0.0512. CEBPB P<0.0949. EGR1 P<0.244. REG4 P<0.326. GLUD2 P<0.106. TFF1 P<0.281. PCK2 P<0.052. DIT4 P<0.018. N>3. Error bars given as \pm SEM.

Gene	Function	RNAseq Change	qPCR Change	Fold Change	Pval
MAPK10	MAP kinase family protein.	UP	UP	0.48	****
DUSP6	Dual phosphatase. Inactivates MAPKs.	DOWN	DOWN	1.56	*
ADM2	Calcitonin family. Secreted by metabolic stress	UP	UP	4.69	-
APOB	Apolipoprotein that organises LDL formation	DOWN	DOWN	0.0002	****
ANKFN1	Predicted to be involved in spindle formation	UP	UP	18.6	-
TXNIP	Regulator of cellular oxidation through Thioredoxin.	DOWN	DOWN	0.55	**
CTSK	Lysosomal cysteine protease, bone resorption.	UP	DOWN	0.57	*
PNCK	Protein serine/threonine kinase	DOWN	DOWN	1.8	-
IL15	Interleukin 15, an inflammatory cytokine.	UP	UP	3.009	**
ITIH1	Glycoprotein of the family of inter-alpha trypsin inhibitor (I α I) serum proteins.	DOWN	DOWN	0.21	****
LOXL4	Lysyl oxidase homolog 4, essential for essential to the biogenesis of connective tissue	UP	DOWN	0.15	****
ENOX1	Ecto-NOX disulfide-thiol exchanger 1. Electron transport chain pathway.	UP	UP	1.9	**
GLUD1	Glutamate dehydrogenase 1,	UP	DOWN	0.78	**

	glutamine metabolism.					
GLUD2	Glutamate dehydrogenase 2.	UP	UP	1.25	*	
NCALD	Neuronal calcium sensor (NCS)	DOWN	DOWN	0.48	**	
ASNS	Glutamine-Dependent Asparagine Synthetase	DOWN	UP	2.63	-	
CEBPB	CCAAT/enhancer-binding protein beta, bZIP transcription factor	UP	UP	2.50	—	
EGR1	Early growth receptor 1 regulates attachment and survival of cells and induces apoptosis in abnormal cells with decreased adhesion	DOWN	UP	1.29	*	
REG4	Regenerating islet-derived 4 (REG4), involved in macrophage polarisation.	DOWN	DOWN	0.93	***	
TFF1	Trefoil factor 1, secretory protein expressed in intestinal mucosa.	DOWN	DOWN	1.31	***	
PCK2	mitochondrial enzyme	UP	UP	2.11	-	
DDIT4	DDIT4 acts as a negative regulator of mTOR	UP	UP	1.92	*	

Figure 27: Genes highlighted by our RNAseq analysis were selected for analysis by qPCR in COLO320 cells with CHRDL2 overexpression. N=3. P values given as students T-test.

We then looked into the background of these differentially expressed genes, and aimed to elucidate any potential role in tumorigenesis that they may play.

Upregulated genes

ADM2 is upregulated in our RNAseq and qPCR data. *ADM2* belongs to the CGRP/calcitonin family, and is released upon metabolic stress, such as ER, mitochondrial, and integrated stress responses. *ADM2* knockdown suppresses proliferation in tumour cells, and upregulation is associated with aggressive cancer pathologies [13]

Ankyrin repeat and fibronectin type III domain containing 1 (*ANKFN1*), is predicted to be involved in mitotic spindle formation, and involved in cell polarity. There is some evidence that it may increase metastatic potential in hepatocellular carcinoma, by stimulating the MEK1/2-ERK1/2 pathway[14]. *ANKFN1* is upregulated in both our RNAseq and qPCR data.

PNCK, or pregnancy up-regulated non-ubiquitous CaM, was upregulated in both our qPCR and RNAseq data. *PNCK* is a member of the calcium/calmodulin-dependent protein kinase family of protein serine/threonine kinases. *PNCK* is expressed during foetal development, and is suggested in vitro that *PNCK* phosphorylates CAMK1A, CREB, ATF1, and Synapsin [15]. *PNCK* is shown to be upregulated in breast (BC), renal cell (RCC), nasopharyngeal, and hepatocellular carcinomas, and is suggested to stimulate proliferation, clonogenicity, and cell cycle progression [16]

Interleukin 15 (IL15) is an inflammatory cytokine, with structural similarity to IL2. IL15 binds and signals through a complex of IL2/IL15 and CD132 or CD122 [17]. IL15 is secreted by mononuclear phagocytes following infection, stimulated the proliferation of Natural killer cells [18]. In cancer, IL15 has been shown to enhance the ant-tumorigenic effect of CD8+ t cells [19]. Interestingly, *IL15* was upregulated in both RNAseq and QPCR data, suggesting increased immune response in CHRDL2 overexpressing cells.

GLUD1 and GLUD2 are Glutamate dehydrogenases, which are enzyme localised in the mitochondria that acts as a hexamer to recycle nitrogen and glutamate for energy homeostasis. There is evidence to suggest Glutamate hydrogenases suppresses cancer formation through inhibiting PI3K/Akt/mTOR pathway activation [20–22]. Both *GLUD 1* and *2* were upregulated in our RNAseq data.

CEBPB is a is a bZIP transcription factor that can bind as a homodimer to DNA regulatory regions that is upregulated in our cell lines. *CEBPB* is important in the regulation of many inflammatory genes, such as IL1, IL6, and binds to several regulatory regions of several acute-phase and cytokine genes. *CEPB* is also known as the master regulator of macrophage

differentiation [23], and therefore has been implicated in the immune response to several cancers [24].

PCK2 catalyses oxaloacetate to phosphoenolpyruvate during gluconeogenesis. PCK2 is associated with several cancers, including lung and colorectal cancer, and promotes tumorigenesis through its gluconeogenic function [25][26] [27]. In low-glucose settings, stress to the endoplasmic reticulum upregulates ATF4, which then upregulates PCK2. As PCK2 allows cells to utilize alternative methods from glycolysis, PCK2 activity can enhance the survival of tumour cells in a reduced glucose environment. *PCK2* was upregulated in our RNAseq and qPCR data, suggesting that *CHRD2* allows cells to alternatively use metabolism methods in low-glucose settings.

ENOX1 is an NADH oxidase that is expressed in endothelial cells. ENOX1 is a reactive-oxygen species-generating enzyme which regulates redox-dependent signalling. ENOX1 is often expressed in tumour vasculature and stroma, and aids in repair of DSBs, increasing cell survival in response to ionizing radiotherapy [28–30]. Interestingly, *ENOX1* is upregulated by both our qPCR and RNAseq data, which may play a role in *CHRD2* overexpressing cells and resistance to chemotherapy.

Downregulated genes

MAPK10, a member of the MAP kinase family, is downregulated in our RNAseq and qPCR data. *MAPK10* has been previously suggested to inhibit tumorigenesis and be involved in apoptosis [31] and is shown to be downregulated in CRC tumour tissue [32].

DUSP6 is downregulated by our RNAseq and qPCR data. However, *DUSP6* has been described as an oncogene, and de-phosphorylates inactivated MAPK family proteins, including ERK1/2, resulting in increased DNA damage response and enhanced chemotherapy resistance [33][34].

Apolipoprotein B (*APOB*) is a marker of oxidation of the low-density lipoprotein (LDL), which transports cholesterol molecules to organs. There are suggestions that ApoB is upregulated in CRC patients, however this is seemingly to do with diet rather than oncogenic signalling [35][36]. *APOB* is downregulated in our RNAseq and qPCR data.

Thioredoxin-interacting protein (*TXNIP*) is downregulated in both our RNAseq and QPCR data. *TXNIP* is a master regulator of cellular oxidation, and binds to and inhibits Thioredoxin, which together with glutathione, are the major thiol antioxidants. Therefore, dysregulation of *TXNIP*-thioredoxin axis can result in metabolic diseases through increase in reactive oxygen-

species[37][38]. Additional *TXNIP* has been identified as a tumour suppressor gene, as its expression is reduced in a variety of cancers[38].

ITIH1 is a glycoprotein from the of inter-alpha trypsin inhibitor (IαI) serum protein family. ITIH1 covalently binds to hyaluronic acid molecules. ITIH1 presence in serum is proposed to play a roll in ECM stabilisation at the onset of inflammation [39][40]. ITIH1 expression has been found to be downregulated in many cancers, including CRC [41]. Interestingly, *ITIH1* is also downregulated in both our RNAseq and qPCR data.

NCALD (neurocalcin delta) is involved in calcium and G-couples receptor signalling and is a member of the neuronal calcium sensor family. NCALD is abundant in the small intestine, as well as the brain and testis. NCALD has been proposition as a biomarker in myeloid leukaemia and lung and ovarian cancer, [42–44], however the role it may have in tumour progression is conflicted. In ovarian cancer, low expression of NCALD is proposed to increase chemotherapy resistance, and repression of NCALD promotes proliferation in lung cancer [42,44], however it has been suggested high NCALD expression predicts poor prognosis in myeloid leukaemia [43]. *NCALD* reduced expression in *CHRD2* overexpressing cells in our RNAseq and qPCR data.

REG4, or Regenerating islet-derived 4 (REG4), belongs to the calcium-dependent lectin (C-type lectin) gene superfamily, and is abnormally expressed in many cancers, including colorectal [45]. Interestingly, REG4 is activated by GATA6, which also activates LGR5 [46]. REG4 activation enhances the growth of colon cancer cells under adherent conditions and is required for their tumorigenicity. REG4-positive cancer cells show more frequent resistance to chemoradiotherapy, especially 5-FU-based chemotherapy [45]. However, *REG4* was downregulated in our data sets.

TFF1 is a secretory proteins expressed in gastrointestinal mucosa. It has been proposed that TFF1 protect the mucosa from insults and stabilizes the mucus layer. TFF1 expression is frequently lost in gastric carcinoma, probably through mechanism of DNA methylation, and it is therefore considered as a tumour suppressor [47]. Indeed, *TIFF1* is downregulated by *CHRD2* in our RNAseq data.

Conflicting genes

Cathepsin K, or CTSK, is a cysteine cathepsin, or a lysosomal cysteine protease, which is involved in bone remodelling and resorption. CTSK has been shown to be expressed in colorectal, breast cancer and glioblastoma, and may contribute to invasiveness and the formation of bone-residing tumour cells [48][49][50]. *CTSK* is upregulated in our RNAseq data, however, was seemingly downregulated in COLO320 cells when measured by qPCR.

LOXL4, or Lysyl oxidase homolog 4, is a member of the Lox family of proteins, which aid in collagen and elastin crosslinking, as well may play a role in tumour development and metastasis [51][52]. Data suggests that LOXL4 may promote metastasises in several cancer types through activation of the FAK/src pathway to enhance cellular adhesion [52][53]. *LOXL4* was upregulated in our RNAseq data, but downregulated in COLO320 cells, as shown by qPCR.

ASNS is a glutamine-Dependent Asparagine Synthetase which catalyses the synthesis of the non-essential amino acid asparagine (Asn) from aspartate (Asp) and glutamine (Gln). Lack of ASNS expression is a hallmark of Acute Lymphoblastic Leukaemia (ALL) blasts, which, therefore, are auxotrophic for Asn. However, *ASSN* was upregulated in our cell lines, and in some solid tumours it has been shown to overexpress ASNS, which may contribute to proliferation, chemoresistance, and a metastatic behaviour through its metabolic activity [44,54].

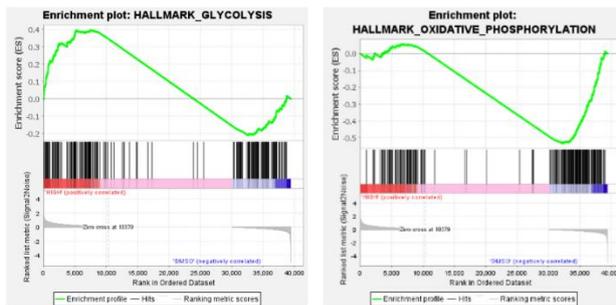
EGR1, or early growth response 1, is a transcription factor that regulated cell survival and attachment and has been implicated in a variety of biological processes, such as cell proliferation, differentiation, invasion, and apoptosis. EGR1 transcription depends on the RAS–RAF–MEK1/2–ERK1/2 signal transduction pathway, which is upregulated by our RNAseq data [55]. *EGR1*, surprisingly, was downregulated in our RNAseq data, however upregulated in COLO320 cells shown by qPCR.

DDIT4, or DNA-inducible transcript 4, acts as a negative regulator of MTOR, which is upregulated in many cancer types. Although *DDIT4* was downregulated in our RNAseq data, it was upregulated in our qPCR data.

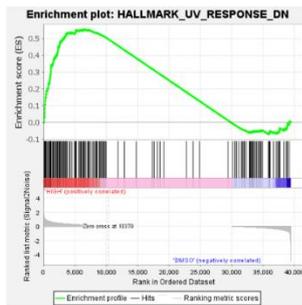
CHRDL2 regulates cancer pathways during overexpression.

Gene-set-enrichment analysis (GSEA) on the entire RNAseq dataset was then performed to analyse genetic pathways disrupted by CHRDL2 overexpression (Figure 28). GSEA analysis revealed upregulation of the hallmark WNT signalling pathway ($P < 0.001$) and BMP regulation ($P < 0.05$), suggesting an increase in WNT signalling and decrease in BMP signalling, which verifies CHRDL2's role as a BMP antagonist in colon cancer cells. The MYC signalling pathway and LEF1 signalling, which are downstream transducers of WNT signalling, were also upregulated. GSEA revealed upregulation of the cancer hallmark pathways, epithelial to mesenchymal transition (EMT) ($P < 0.001$), and angiogenesis, which are frequently upregulated in metastatic colorectal cancers DNA repair pathways were also significantly upregulated including key DSB repair genes BRCA1, RAD51 and RAD52, supporting our findings with respect to chemotherapy resistance ($P < 0.05$). There was also significant upregulation of RAF and MTOR signalling, which are often modulated during cancer progression. Furthermore, cell cycle-related genes upregulated by Rb knockout were also upregulated by CHRDL2, suggesting an increase in cell-cycle protein signalling. We noted that, BMI1 pathways were also highlighted by GSEA a further stem-cell defining pathway and correlating with our Q-PCR data.

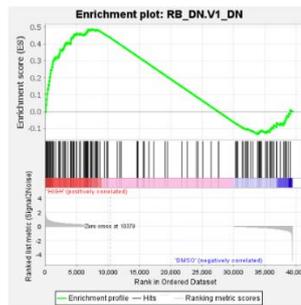
Metabolism



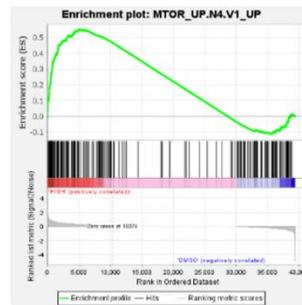
DNA damage repair



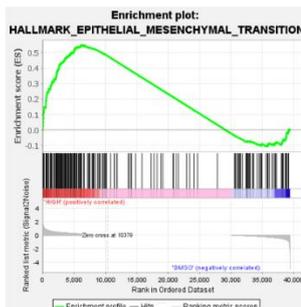
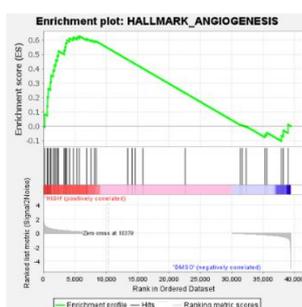
Cell cycle



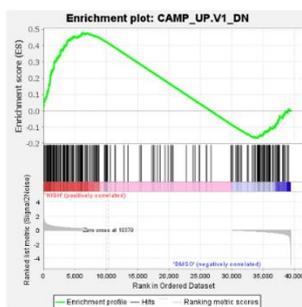
MTOR signalling



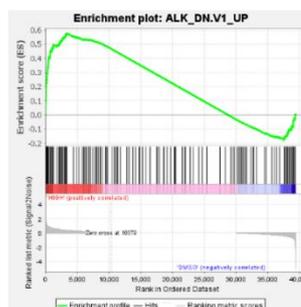
Hallmark cancer pathways



cAMP signalling



ALK signalling



IL2/STAT5 signalling

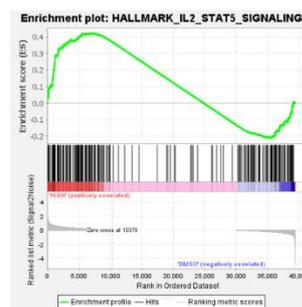


Figure 28: GSEA plots of RNAseq analysis from CHRDL2++ cells (HIGH) compared to DMSO control. P values given as FWER. WNT_SIGNALLING NES= 1.15 P<0.0. MYC_UP.V1_DN NES=1.25 P<0.01. LEF1_UP.V1_UP NES=1.26 P<0.01. CORDENOSI_YAP_CONSERVED_SIGNATURE NES=1.36 P<0.01. YAP1_UP NES=1.32 P<0.01. BMI1_DN.V1_UP NES=1.33 P<0.01. BMI1_DN_MEL18_DN.V1_UP NES=1.38 P<0.01. MEL18_DN.V1_UP NES=1.41 P<0.0. RAF_UP.V1_UP NES=1.2 P<0.01. EGFR_UP.V1_UP NES= 1.25 P<0.0. TBK1.DF_UP NES=1.2 P<0.01. HALLMARK_GLYCOLYSIS NES=1,13 P<0.01. HALLMARK_OXIDATIVE_PHOSPHORYLATION NES=-1.1 P<0.01. HALLMARK_UV_RESPONSE_DN NES=1.199 P<0.01. RB_DN.V1_DN NES=1.2 P<0.01. MTOR_UP.N4.V1_UP NES=1.31 P<0.01. HALLMARK_ANGIOGENESIS NES=1.21 P<0.01. HALLMARK_MESENCHYMAL_TRANSITION NES=1.29 P<0.01. CAMP_UP.V1_DN NES=1.18 P<0.01. ALK_DN.V1_UP NES=1.36 P<0.01. HALLMARK_IL2_STAT5_SIGNALLING NES=1.15 P<0.01. N=3.

We then ran our highly differentially expressed genes through the GO ontology PANTHER analysis, to further elucidate any upregulation of biological pathways (Figure 29). We found a positive upregulation of adhesion pathways and cell spreading, indicating a more invasive cellular phenotype, supporting our findings of increased EMT signaling. Interestingly, we also found cell enrichment for regulation of apoptotic processes, and programmed cell death. Taken with our previous findings of enhanced survivability under chemo and radiotherapy, this would indicate a downregulation in apoptotic pathways, indicative of pro-survival cancer-stem cell phenotype. We also observed an effect on cellular differentiation pathway, which again may support our hypothesis of a de-differentiated phenotype.

GO biological process complete	Homo sapiens - REFLIST [20592]	DIFFERENTIALLY EXPRESSED BY	over/under	fold Enrichment	raw P-value	FDR
		CHRD12 [130]				
triglyceride mobilization (GO:0006642)	3	2	+	> 100	1.18E-04	4.29E-02
neutrophil homeostasis (GO:0001780)	17	3	+	27.95	1.57E-04	4.87E-02
positive regulation of substrate adhesion-dependent cell spreading (GO:1900026)	43	4	+	14.73	1.55E-04	5.02E-02
proteino-genic amino acid biosynthetic process (GO:0170038)	54	5	+	14.67	2.29E-05	2.05E-02
L-amino acid biosynthetic process (GO:0170034)	54	5	+	14.67	2.29E-05	1.94E-02
cellular response to glucocorticoid stimulus (GO:0071385)	55	5	+	14.4	2.51E-05	2.01E-02
cellular response to corticosteroid stimulus (GO:0071384)	65	5	+	12.18	5.66E-05	3.32E-02
alpha-amino acid biosynthetic process (GO:1901607)	65	5	+	12.18	5.66E-05	3.20E-02
amino acid biosynthetic process (GO:0008652)	67	5	+	11.82	6.55E-05	3.12E-02
regulation of lipid biosynthetic process (GO:0046890)	175	7	+	6.34	1.25E-04	4.32E-02
regulation of small molecule metabolic process (GO:0062012)	324	11	+	5.38	6.80E-06	1.48E-02
carboxylic acid biosynthetic process (GO:0046394)	278	9	+	5.13	6.94E-05	3.21E-02
organic acid biosynthetic process (GO:0016053)	281	9	+	5.07	7.53E-05	3.19E-02
regulation of hormone levels (GO:0010817)	550	14	+	4.03	1.02E-05	1.29E-02
positive regulation of transport (GO:0051050)	855	18	+	3.33	7.24E-06	1.38E-02
carboxylic acid metabolic process (GO:0019752)	801	16	+	3.16	4.58E-05	3.04E-02
oxoacid metabolic process (GO:0043436)	823	16	+	3.08	6.31E-05	3.10E-02
organic acid metabolic process (GO:0006082)	830	16	+	3.05	6.98E-05	3.13E-02
response to hormone (GO:0009725)	785	15	+	3.03	1.31E-04	4.45E-02
positive regulation of developmental process (GO:0051094)	1337	25	+	2.96	8.50E-07	4.32E-03
positive regulation of cell differentiation (GO:0045597)	871	16	+	2.91	1.22E-04	4.33E-02
regulation of apoptotic process (GO:0042981)	1476	24	+	2.58	1.58E-05	1.72E-02
lipid metabolic process (GO:0006629)	1238	20	+	2.56	9.71E-05	3.79E-02
regulation of programmed cell death (GO:0043067)	1521	24	+	2.5	2.58E-05	1.97E-02
response to endogenous stimulus (GO:0009719)	1411	22	+	2.47	7.04E-05	3.07E-02
response to biotic stimulus (GO:0009607)	1429	22	+	2.44	8.48E-05	3.40E-02
response to other organism (GO:0051707)	1382	21	+	2.41	1.51E-04	4.99E-02
response to external biotic stimulus (GO:0043207)	1385	21	+	2.4	1.55E-04	4.93E-02
regulation of developmental process (GO:0050793)	2448	35	+	2.26	3.44E-06	1.05E-02
regulation of localization (GO:0032879)	2028	28	+	2.19	7.58E-05	3.12E-02
anatomical structure morphogenesis (GO:0009653)	2239	30	+	2.12	5.43E-05	3.45E-02
regulation of multicellular organismal process (GO:0051239)	2966	39	+	2.08	4.15E-06	1.05E-02
response to external stimulus (GO:0009605)	2290	30	+	2.08	1.06E-04	3.95E-02
regulation of biological quality (GO:0069008)	2849	37	+	2.06	1.62E-05	1.65E-02
response to organic substance (GO:0010033)	2467	32	+	2.05	5.91E-05	3.10E-02
cell differentiation (GO:0030154)	3654	44	+	1.91	8.65E-06	1.32E-02
cellular developmental process (GO:0048869)	3657	44	+	1.91	8.78E-06	1.22E-02
system development (GO:0048731)	3546	41	+	1.83	5.84E-05	3.18E-02
multicellular organism development (GO:0007275)	3964	45	+	1.8	4.41E-05	3.05E-02
anatomical structure development (GO:0048856)	5231	59	+	1.79	7.85E-07	5.98E-03
developmental process (GO:0032502)	5739	63	+	1.74	6.11E-07	9.32E-03
cell communication (GO:0007154)	5262	55	+	1.66	2.92E-05	2.12E-02
signal transduction (GO:0007165)	4789	50	+	1.65	9.84E-05	3.75E-02
signaling (GO:0023052)	5109	53	+	1.64	5.97E-05	3.03E-02
positive regulation of cellular process (GO:0048522)	5707	59	+	1.64	1.84E-05	1.75E-02
multicellular organismal process (GO:0032501)	6745	69	+	1.62	1.96E-06	7.46E-03
positive regulation of biological process (GO:0048518)	6210	63	+	1.61	1.26E-05	1.47E-02
cellular response to stimulus (GO:0051716)	6448	63	+	1.55	5.57E-05	3.40E-02
response to stimulus (GO:0050856)	8182	77	+	1.49	8.29E-06	1.40E-02

Figure 29: PANTHER Overrepresentation Test on CHRD12++ differentially expressed genes. Provided by GO ontology database.

Discussion

Here, through online database sources and a comprehensive RNAseq analysis, we have aimed to elucidate potential pathways modulated by CHRDL2 overexpression.

We have shown that CHRDL2 is upregulated in CRC compared with normal tissues, and high expression is associated with poor patient survival outcomes in stage 4 of the disease. We have also shown that high expression of BMP5, which is inhibited by CHRDL2, enhances patient survival, perhaps through pro-differentiation pathways. When looking at stage specific survival, we see that CHRDL2 high expression only predicts poor patient outcome at later disease stages. This could correlate with our data found previously that CHRDL2 enhances chemotherapy survival, which is used in later disease stages in CRC treatment. Furthermore, CHRDL2 does not affect survival at the earliest disease stages where hyper-proliferation plays an important role, which is supported by our findings of a slower growth after CHRDL2 overexpression. It could be seen that at initiation of tumour progression, CHRDL2 reduces proliferation, resulting in smaller polyps that are removed easily, but at later stages of the disease where treatment relies of chemo and radiotherapy, CRHDL2 overexpression enhances the cancer stem phenotype and enhances chemotherapy resistance, increasing the likelihood for incomplete tumour abolition. There did not appear to be any significant differences in CMS classification in later stages CRC with CHRDL2 overexpression, however it appeared that CHRDL2 decreased survival most strongly in CMS2 tumours, which are WNT dysregulated, which may support our hypothesis of WNT activation by CHRDL2 overexpression.

Next, we utilised RNAseq analysis to uncover potential pathways dysregulated by CHRDL2 overexpression. CHRDL2 overexpression was given as CHRDL2, CHRDL+ and CHRDL2++, also termed Low Medium and High. Heat maps showed differentially expressed genes, and the clustering of Medium and High treatments. The Low treatment appeared genetically very similar to the DMSO control, so our analysis focused on the Medium and High treated groups. Volcano plots were generated to find differentially expressed genes, and genes that were significantly differentially expressed in both groups were taken for further analysis and validation by qPCR. Differentially expressed genes were also put through GSEA analysis, to elucidate differential expression of common cancer-associated molecular pathways, and PANTHER analysis for common biological pathways.

Through GSEA analysis we also saw an increase in WNT signalling pathways, such as general WNT signalling, MYC, and LIF1, as well as BMI1 signalling, which are known to mark slow-cycling stem cells. This supports our hypothesis that CHRDL2 increases a slow-cycling stem

cell phenotype through WNT activation. Furthermore, we saw an increase in DNA repair signalling, which corroborates our previous finds of increased DNA damage repair. This was also seen by PANTHER analysis, which showed regulation of the apoptotic process and cell death, supporting our anti-apoptotic pro survival phenotype.

Interestingly, several genes that are involved in metabolism and oxidative stress were dysregulated by CHRDL2. For example, TXNIP was downregulated in our data, which causes dysregulation of oxygen metabolism and production of reactive oxygen species. ENOX1, which generates reactive oxygen species, was also found to be upregulated by CHRDL2. Reactive oxygen species have been found to aid in repair of double stranded breaks [256–258], which we have previously shown to be a feature of CHRDL2 overexpression. ADM2 and PCK2, which are also involved in metabolic stress, were also upregulated by our CHRDL2 overexpressing cells, and have been shown to enhance proliferation. PCK2 in particular, is utilised by cancer cells for glucose-independent metabolism, which we have shown to be important in our CHRDL2 proliferation data. GSEA analysis also showed a dysregulation in glycolysis signalling, and a reduction in oxidative phosphorylation.

PNCK and ASNS have also been implicated as enhancing cancer cell proliferation, which were both upregulated by our CHRDL2 overexpressing cell lines. However, previously we have shown that CHRDL overexpression reduces cellular proliferation.

CTSK and LOXL4 were upregulated and have both been shown to enhance an invasiveness cancer cell phenotype. This fits with our findings on increased migratory and EMT markers found in CHRDL2 overexpressing cells.

Interestingly, we saw stimulation of the RAS/RAF/MEK/ERK pathway through upregulation of both ANKFN1 and EGR1. The RAS/RAF/MEK/ERK pathway has been well documented in cancer, with overexpression resulting in proliferation, migration, and metastasis [259]. This is corroborated by GSEA analysis, which shows an enrichment for the RAF signalling pathway.

There was also evidence of dysregulation of several immune and inflammatory pathways by CHRDL2 overexpression. IL15 was up in our data set, which is involved in immune cell infiltration to tumour sites. CEBP was also upregulated, which stimulates IL5 and IL6, who show structural and functional similarities to IL15. However, previous finds have shown that immune infiltration to tumour sites is beneficial for patient survival, and IL15 has been proposed as enhancing cancer-cell immunity [260]. IL2 signalling was also upregulated by GSEA, suggesting an increase in immune infiltrator. ITH1, which has been shown to stabilise ECM during inflammation, was downregulated in our data, and has also been shown to be lost in many cancer types. This would suggest an increase in immune and inflammation pathways by CHRDL2 overexpression.

We also found upregulated of MAPK10, which has been shown to inhibit apoptosis, which links to our data on increase survivability. However, we also saw a downregulation of DUSP6, which activates de-phosphorylated MAPK10, perhaps acting as a negative feedback loop.

There was an observed downregulation of genes that are highly expressed in the healthy intestinal lining and mucosa, such as NCALD and TIFF1. NCALD, a calcium sensor, is found abundantly expressed in the intestinal tract, as well as TIFF1, which aides in generation of the intestinal mucosa. Both of these genes were down regulated by CHRDL2 overexpression. This could signify a de-differentiation of our cancer epithelial cells, due to evidence of loss of normal intestinal markers.

Metastatic pathways were also upregulated, with an increase in EMT signalling, and angiogenesis by GSEA, and cell spreading by PANTHER analysis. This is supported by positive increase in genes that effect the RAS/RAF/MEK/ERK pathways, and CTSK and LOXL4 which promote invasiveness. Together, this data would indicate a more migratory, invasive phenotype, which we observed in our 2D cell line experiments.

Limitations and future work

To strengthen the findings from our RNAseq data and qPCR validation, western blotting could be performed to confirm protein overexpression of the genes deregulated by CHRDL2. This would aid in elucidating the pathways in which CHRDL2 acts, to further uncover the mechanisms in which CHRDL2 exerts in functional affects observed in previous chapters.

Furthermore, we have only analysed publicly available data from TCGA that showed expression of CHRL2 in patient tumour samples. To confirm these findings, further work could be done for staining CHRDL2 protein expression in patient samples. We have shown that CHRDL2 negatively affects patient survival at later stages of the disease, however this was done using only 19 patients with CHRDL2 high expression at stage 4, so expanding upon this data set would strengthen the conclusions found in this study.

Conclusion

CHRDL2 is overexpressed in cancer tissues, and high CHRDL2 expression predicts poor prognosis in the later stages of disease. CHRDL2 overexpression leads to an increase in stem-cell WNT signalling, and leads to the upregulation of DNA repair, migratory, and invasiveness pathways.

Chapter 6: The role of POLD3 knockdown in CRC

Introduction

DNA replication, and its ability to do so without error, is one of the most important factors governing cell division. For a cell to replicate its DNA to pass onto daughter cells, polymerases (Pol) are required whose enzymatic activity synthesises nucleic acid chains. Polymerases are also vitally important in ensuring correct replication is maintained, by repairing damaged DNA through DNA damage repair pathways [261]. The polymerase Delta complex (Pol δ) is one such polymerase whose role is to both synthesise DNA and repair it, making it a clear functional target for maintaining DNA integrity [262][177][263].

Pol δ is a heterodimer comprising subunits POLD1, POLD2, POLD3, and POLD4. During DNA replication, Pol δ is responsible for lagging strand DNA synthesis, as well as repair of DNA double-strand breaks (DSBs) via homologous recombination (HR). Pol δ works as a gap-filling polymerase, which is often a site for DNA fragility during replication [262].

Cancer cells intrinsically rely on this ability to replicate and repair DNA at rapid rates, while maintaining enough DNA integrity to prevent cellular collapse. Colorectal cancer (CRC) is one of the most prevalent and deadly cancers globally, and accounts for 10% of all new cancer cases each year [264].

CRC is a heterogeneous disease and can arise from numerous mutations and varying pathways of genetic instability. Genetic instability is a hallmark of CRC, and CRC can be classified by different forms of genome or epigenetic instability, including chromosomal instability (CIN), microsatellite instability (MSI) (caused by faults in the mismatch-repair (MMR) system, and high single nucleotide mutation rates (hypermutation-ultra mutation)[265]. However, patients who do not carry germline faulty MMR mutations have mutations in the DNA replication and proof-reading mechanisms, such as with Pol δ , which increase genetic instability and decrease DNA proof-reading mechanisms, leading to activation of oncogenic pathways.

Recent attention has been brought to the loci surrounding the POLD3 gene this subunit of Pol δ , as increasing the risk for the development of CRC [43] [266][267]. Furthermore, the Pol δ has been shown to be overexpressed in CRC, as well as the POLD3 gene frequently found to be amplified in many CRC tumours [268]. In light of this, POLD3 makes an interesting target for studying the effects of Pol δ mutations in the context of CRC development.

POLD3 was first identified using proliferating cell nuclear antigen affinity chromatography (PCNA column) and glycerol gradient centrifugation from mouse and calf thymus, which revealed a distinct subunit that reacted strongly with both Pol δ complex and the PCNA binding domain [174] [175]. POLD3 has dual roles in Pol δ complex, first as stabilizer the POLD1-POLD2 interaction, and second as a facilitator to the binding of the POLD complex to PCNA through a C-terminal PIP box. POLD3 also exhibits 3' to 5' exonuclease activity, which increases the processivity of DNA synthesis during replication as seen in figure 2. POLD3 mediates the binding of PCNA to Pol δ through a canonical PCNA-binding sequence located in its C terminus (83). Removal or malfunction of the POLD3 subunit prevents PCNA binding to Pol δ , and therefore ineffective DNA replication.

In addition, Pol δ has roles in DNA double strand break repair via homologous recombination, an important aspect of DNA repair. The Pol δ complex also participates in DNA mismatch and base-excision repair, which is a key process shown to be defective in many hereditary CRC susceptibility disorders, such as lynch syndrome [176]. Phosphorylation of POLD3 by cyclin-dependent kinases shows that POLD3 activity may be dependent on cell cycle regulation, an important factor when considering cancer cell cycle progression [177].

POLD3 has also been implicated in the telomere maintenance pathway. Telomeric maintenance is required by cancer cells in order to increase cell longevity after successive cell divisions. Therefore, cancer cells require methods of increasing/sustaining telomere lengths. The majority of cancer cells maintain telomeres through telomerase reactivation, where activation of *TERT*, the active catalyst in telomerase, mediates telomere elongation. A small subset (10~15%) of cancer cells do not use *TERT* reactivation to maintain telomeres, but rely on the alternative lengthening of telomeres (ALT) mechanism [269][270]. However, recent publications have discovered these two methods of telomere maintenance may also occur concurrently [271][272].

The ALT mechanism of telomere maintenance is poorly understood. It is known that promyelocytic leukaemia protein (PML) bodies, also known as PML bodies, cluster in ALT+ cells, termed ALT-associated PML bodies (APBs) [273]. These APBs are associated with heterogeneous telomere length, abundant extrachromosomal telomere repeat (ECTR), C-circles at telomere ends, and high levels of telomere sister chromatid exchange (T-SCE), which together present hallmarks for ALT+ cells [274]. At sites of telomere ends, R-loops, G-quadruplexes, and DNA single-strand breaks cause the collapse of replication forks which induce SUMOylation of telomere proteins, which recruit PML and trigger APB formation through SUMO/SIM-mediated liquid-liquid phase separation LLPS. APBs then function to cluster telomeres together, and recruit repair proteins, which initiate DNA repair by break-

induced repair (BIR) mechanisms. BIR then takes place through RAD52 dependent and independent pathways, promoted by BLM helicase, and mediated by POLD3/POLD4 which form C-circles to resolve telomere ends [274] [274].

Previously, studies have shown that POLD3 loss in mouse embryonic stem cells results in rapid telomere shortening, chromosomal abnormality and aneuploidy, and increased number of chromosome or chromatid breaks [275]. Furthermore, POLD3 has recently been shown to be required for break-induced repair and the conservative replication of telomeres in human alternative lengthening of telomeres (ALT) cell lines [276][277], suggesting POLD3 is required for ALT telomeric maintenance using homology-directed DNA repair to maintain telomeres.

Due to the recent attention surrounding the POLD3 loci as increasing CRC risk, and the define roles POLD3 plays in DNA synthesis, repair, and telomere maintenance, we sought to discover the functional effects of POLD3 in 2D cell culture systems. Despite the genetic studies on POLD3, and the understanding of its functional affect it is still unknown in POLD3 is oncogenic. POLD3 is upregulated in many tumour tissues [267][268], however it is unknown if this is due to oncogenic potential of due to pressure as a protection mechanism triggered to improve tolerance to replication stress [278]

Establishing *POLD3* knockdown cell lines

To analyse the effects of *POLD3* expression, we used a dox-inducible shRNA delivered by lentiviral vector to knockdown *POLD3*. Knockdown of *POLD3* was chosen over *POLD3* overexpression, due to the already high presences of *POLD3* in CRC cell lines [268]. Two shRNA oligos were designed with complementary target sequences to control for off-target effects, and resuspended and annealed through thermocycling (Figure 30 A).

ShRNA oligos were then treated with T4 poly-nucleotide kinase to phosphorylate annealed oligo ends. Vector EZ-Tet-pLKO-Puro (Addgene, #85966) was used to allow doxycycline inducible expression of the shRNA.

To validate the incorporation of the target sequence within the vector, colony PCR was performed. Sequencing primers were developed at the 5' site 398, and the 3' site at 861, which encompasses the target sequence within the vector. If the vector successfully ligated with the target sequence, this would produce a fragment of 423 bp, whereas a re-ligated vector without the sequence would be 363, and the fragment produced from the uncut original vector would be 584 bp.

As seen in Figure 30 B, colony PCR showed that samples in lanes 1,2,4,6,7,8,9, and 10 showed successfully ligated vector containing the target sequence. The sample in lane 3 showed a fragment of size ~500 bp that was equivalent to the uncut vector in lane 10, indicating that this colony did not contain the target sequence, but the uncut vector. The sample in lane 5 showed bands in both 400 bp and 500 bp, indicating that this sample contained both successfully ligated vector with the target sequence and the uncut vector. Samples in lane 9 and 10 also showed a faint band at 500 bp, however this may be caused by spill over from lane 11, which contained the uncut vector, as seen in lane 12 which was used as a water control. Clones 2 and 6 from shRNA 1 and 2 were validated by sanger sequencing and shown to contain the correct target sequence of the ShRNA, confirming incorporation of the annealed vector containing our shRNA sequence.

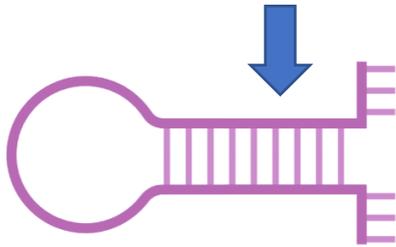
To assess the effects of *POLD3* depletion in cancer cells, we transduced 2 well characterised CRC cells lines CACO2 and RKO, with our virally packaged doxycycline-inducible shRNAs targeting the *POLD3* gene. 2 shRNAs targeting different loci of the *POLD3* gene were generated which were transduced into separate cell lines, termed *POLD3*-1 for shRNA 1 and *POLD3*-2 for shRNA 2 (Figure 30 A). Our vector was a doxycycline inducible vector, which allowed measuring of *POLD3* knockdown in an inducible and level dependent manner (Figure 30, C). Doxycycline was administered to cells lines at 0.1 µg/ml (*POLD3*), 1 µg/ml (*POLD3*-) or 10 µg/ml (*POLD3*--) to analyse the effects of *POLD3* depletion at different levels. qPCR and western blotting confirmed mRNA and protein reduction as seen in Figure 30, D-E. *POLD3* knockdown was most efficient using shRNA-1, with 10 µg/ml doxycycline treatment of CACO2 cell lines showing reduction of *POLD3* mRNA significantly ($P < 0.01$) and RKO ($P < 0.01$). shRNA2 also showed significant reduction of *POLD3* expression in CACO2 cells at 10 µg/ml ($P < 0.005$), however this reduction was less than that of shRNA 1. In RKO cells, shRNA 2 showed a greater reduction of *POLD3* expression than shRNA 1, with 0.1 µg/ml ($P < 0.01$) and 10 µg/ml ($P < 0.01$), so was chosen in preference to shRNA 1.

A

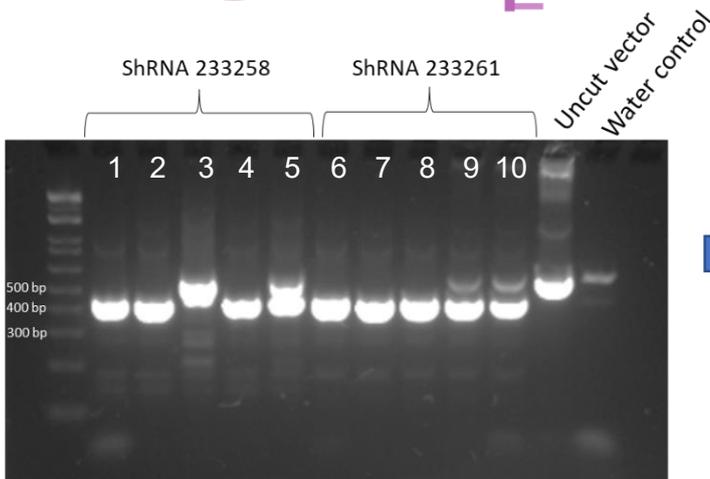
shRNA 1 (233258) 5'-CTAGCACTGACTATGACATCCTTAAATACTAGTTTTAAGGATGTCATAGTCAGTTTTTTG-3'

shRNA 2 (233261) 5'-CTAGCCGAGTCAGCATTGACGATATTACTAGTAATATCGTCAATGCTGACTCGTTTTTTG-3'

Nhe1 Overhang (yellow box)
 Spe1 Loop (cyan box)
 Termination (red box)
 EcoR1 overhang (green box)

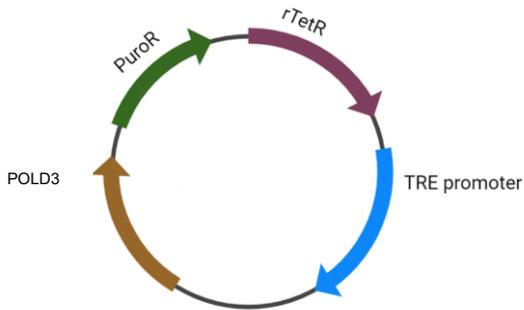


B

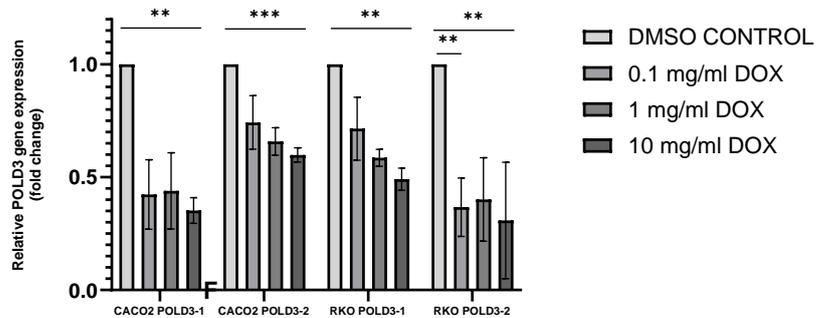


Clones 2 and 6 were validated through sanger sequencing and virally packed before transfection

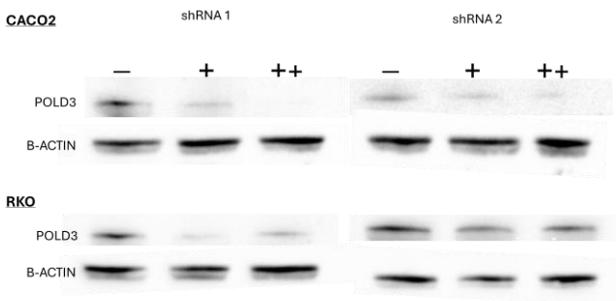
C



D



E



F

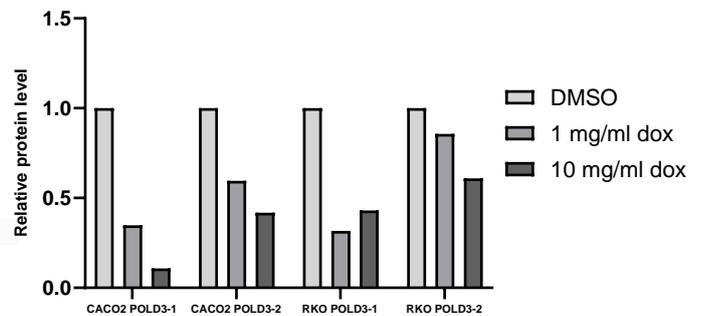


Figure 30: A) shRNA sequences used of POLD3 knockdown. B) Agarose gel electrophoresis of colony PCR from ligated clones contain the shRNA POLD3 sequences. Predicted fragment size for successfully ligated clones is 423 bp, clones harbouring re-ligated vector without the sequence is 363 bp, and clones with uncut vector is 584 bp, as seen in lane 11. C) qPCR of mRNA levels of POLD3 expressed as fold change in 4 experimental cell lines. Cell lines were grown with doxycycline at: 0.1 µg/ml, 1 µg/ml or 10 µg/ml to induce expression of shRNA 1. CACO2 POLD3-1-10µg/ml $p < 0.01$, RKO POLD3-1-10µg/ml $p < 0.01$, N=3. D) qPCR of mRNA levels of POLD3 expressed as fold change in 4 experimental cell lines. Cell lines were grown with doxycycline at: 0.1 µg/ml, 1 µg/ml or 10 µg/ml to induce expression of shRNA 2. CACO2 POLD3-2-10µg/ml $p < 0.001$, RKO POLD3-2-10µg/ml $p < 0.01$, RKO POLD3-2-1µg/ml $p < 0.01$, N=3. E) Western blotting of corresponding protein levels of POLD3 in cell lines with lentiviral overexpression. F) Quantification of POLD3 protein levels as measure by western blot using Image J software.

POLD3 knockdown reduces proliferation and clonogenic ability.

Proliferation of cancer cells inherently relies on the use of DNA polymerases to synthesise DNA. Therefore- we measured the effects of POLD3 knockdown on proliferation in our cancer cells lines. Both CACO2 and RKO cell lines showed a reduction in cellular proliferation after 72hrs treatment with 10 µg/ml doxycycline to induce shRNA expression ($P < 0.001$). Growth curves were generated, which showed a reduction in proliferation over time, which was significant in all tested cell lines ($P < 0.0001$) (Figure 31 A).

Next, colony formation ability (clonogenicity) was assessed. Cells treated with 10 µg/ml doxycycline to induce POLD3 knockdown showed smaller colonies, however the number of total colonies was not reduced (Figure 31 B, C). This signifies that POLD3 depletion does not affect the ability of cells to form colonies, but limits growth. Colonies over a size threshold were counted, and POLD3 depletion in both cells during POLD3-1 induction showed a reduction in larger colonies in both CACO2 and RKO cell lines at 1 µg/ml ($P < 0.001$ and $P < 0.01$) and 10 µg/ml ($P < 0.001$ and $P < 0.05$) respectively.

Next, we analysed the effects of POLD3 on cell cycle progression. As seen in Figure 31 D, both shRNAs tested significantly increase the proportion of cells stalled in s phase ($P < 0.005$). This would signify a slow replication ability during POLD3 knockdown, perhaps due to the lack of binding of the polymerase delta complex to the PCNA clamp. Insufficient DNA replication machinery binding will result in slower replication, and accumulation of errors which may cause cells to stall in S phase. Furthermore, we saw a reduction of cells in G2 phase, indicating that POLD3- cells were unable to pass the S phase checkpoint as readily as control cells. This is further seen in Figure 31 E, where an increase in S phase cells can be observed.

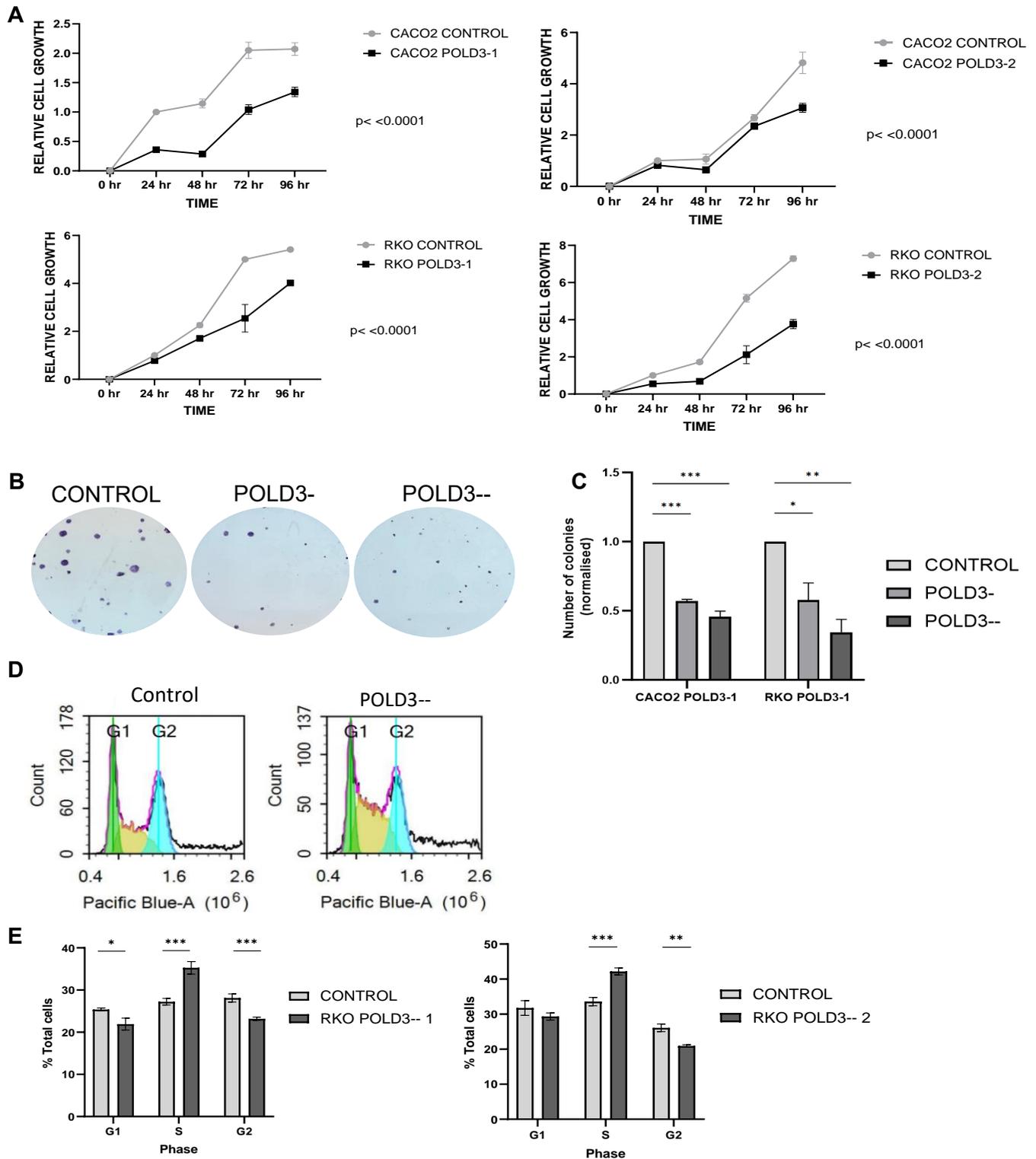


Figure 31: A) Time course of MTT assay of cellular proliferation on CACO2 POLD3-1, CACO2 POLD3-2, RKO POLD3-1 and RKO POLD3-2 cell lines. Two-way RM ANOVA $P < 0.0001$, $P < 0.0001$, $P < 0.0001$, $P < 0.0001$. B) Crystal violet staining of colonies of RKO cells treated with 1 $\mu\text{g/ml}$ and 10 $\mu\text{g/ml}$ doxycycline to induce POLD3-1 knockdown. C) Quantification of colonies established in our 2 experimental cell lines with POLD3-1 knockdown. CACO2 and RKO cell lines both showed reduced colony formation in the 1 $\mu\text{g/ml}$ and 10 $\mu\text{g/ml}$ treated groups, $p < 0.001$, $p < 0.001$, $p < 0.01$, $P < 0.05$. $N = 3$. D) Flow cytometry analysis of RKO cells given POLD3 knockdown by shRNA 1. Student T-test. shRNA 1: G1 $P < 0.05$, S $P < 0.005$, G2 $P < 0.005$. Flow cytometry analysis of RKO cells given POLD3 knockdown by shRNA 2. Student T-test. shRNA 2: S $P < 0.005$, G2 $P < 0.001$. E) Histograms of cell cycle status in RKO POLD3-1 cells. Error bars given as \pm SEM

POLD3 knockdown induces senescence and apoptosis

Next, we sought to uncover the reason for lower cell confluency in POLD3 knockdown cells. To do this, we looked at senescence and apoptosis markers in our cell lines.

The ki67 status of our POLD3 knockdown cells was analysed to assess the proportion of cells actively proliferating cells. As seen by representative images in figure 27 A, Immunofluorescence staining of Ki67 was present in a greater proportion of control cells compared to that of in our POLD3 knockdown cell lines. Quantification of RKO cells with shRNA 1 and 2 showed that Ki67 + cells dropped from 80% to 40% and 30% respectively ($P < 0.05$, $P < 0.0001$).

As seen in figure 32 C and D, flow cytometry staining by Annexin V revealed that POLD3 knockdown dramatically increased the proportion of cells that had begun to enter apoptosis. This was found in both RKO shRNA 1 and 2 cell lines, $P < 0.01$, $P < 0.01$. This data collectively suggests that POLD3 knockdown causes reduction in actively proliferating cells, as well as causing cells to enter apoptosis.

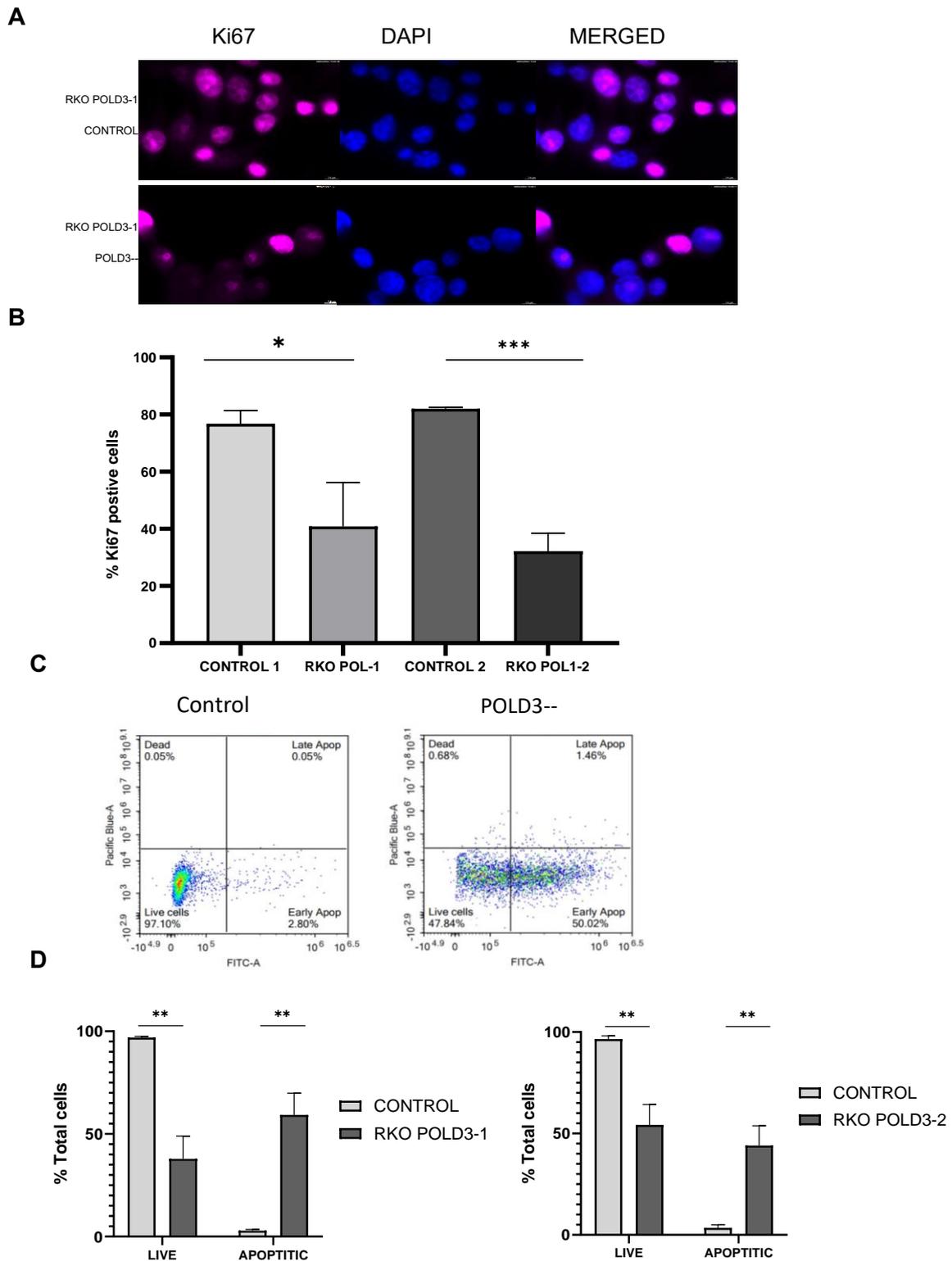


Figure 32: A) Ki-67 immunofluorescence staining of RKO cells given 10 $\mu\text{g/ml}$ shRNA1 knockdown of POLD3. B) Quantification of the % of Ki-67+ cells by immunofluorescence in RKO cells given 10 $\mu\text{g/ml}$ shRNA 1 and 2 POLD3 knockdown. Students T-test: POLD3—1 $P < 0.01$, POLD3—2 $P < 0.01$. C) Scatter plots of apoptotic cells upon POLD3 knockdown by shRNA1 in RKO cells, stained with apoptotic marker Annexin V. D) Quantification of apoptotic cells marked by Annexin V staining via flow cytometry. RKO cells treated with 10 $\mu\text{g/ml}$ shRNA 1 and 2. Students T-test: shRNA1 live cells $P < 0.01$ apoptotic cells $P < 0.01$, shRNA2 live cells $P < 0.01$ apoptotic cells $P < 0.01$. Error bars given as \pm SEM

POLD3 knockdown sensitises cells to chemotherapy.

Next, sensitization to chemotherapy during POLD3 depletion was assessed. Chemotherapy remains the cornerstone of CRC treatment within the UK, with the chemotherapy drug 5FU used in combination with Irinotecan or Oxaliplatin, the most common regimen in the UK. These chemotherapy agents target DNA synthesis, leading to DNA damage and cell death, whereas DNA polymerases can help cells survive DNA damage through their involvement in DNA repair mechanisms. [17]. Therefore, we analysed how POLD3 knockdown would affect cells during chemotherapy.

POLD3 has previously been shown to act to repair DSBs [18]. Therefore, for our analysis we chose Irinotecan to test chemo-sensitisation, as Irinotecan acts as a topoisomerase 1 inhibitor, resulting in double-stranded breaks (DSBs). POLD3 depletion by shRNA 1 and 2 significantly reduced the half maximal inhibitory concentration (IC50) of Irinotecan in both CACO2 and RKO cell lines ($P < 0.018$, $P < 0.0001$, $P < 0.014$, $P < 0.0001$) (Figure 33 A-D), indicating sensitisation to chemotherapy.

Next, we looked at radio-sensitization. Radiotherapy by irradiation is commonly used in conjunction with chemotherapy, particularly for later stage rectal cancers, Targeted ionising radiation causes DNA damage in tumour cells, also in the form of DSBs. When we subjected out POLD3- cell lines to 5 varying levels of irradiation; 2, 4 6 and 8 GY, we observed an increase in cell death compared with controls (Figure 33, E, F). As seen in Figure 2 E, RKO cells treated with shRNA 1 showed a significant reduction in viable cells numbers compared to a control during irradiation, $P < 0.005$, which was replicated by the second shRNA ($P < 0.001$) (Figure 33 F).

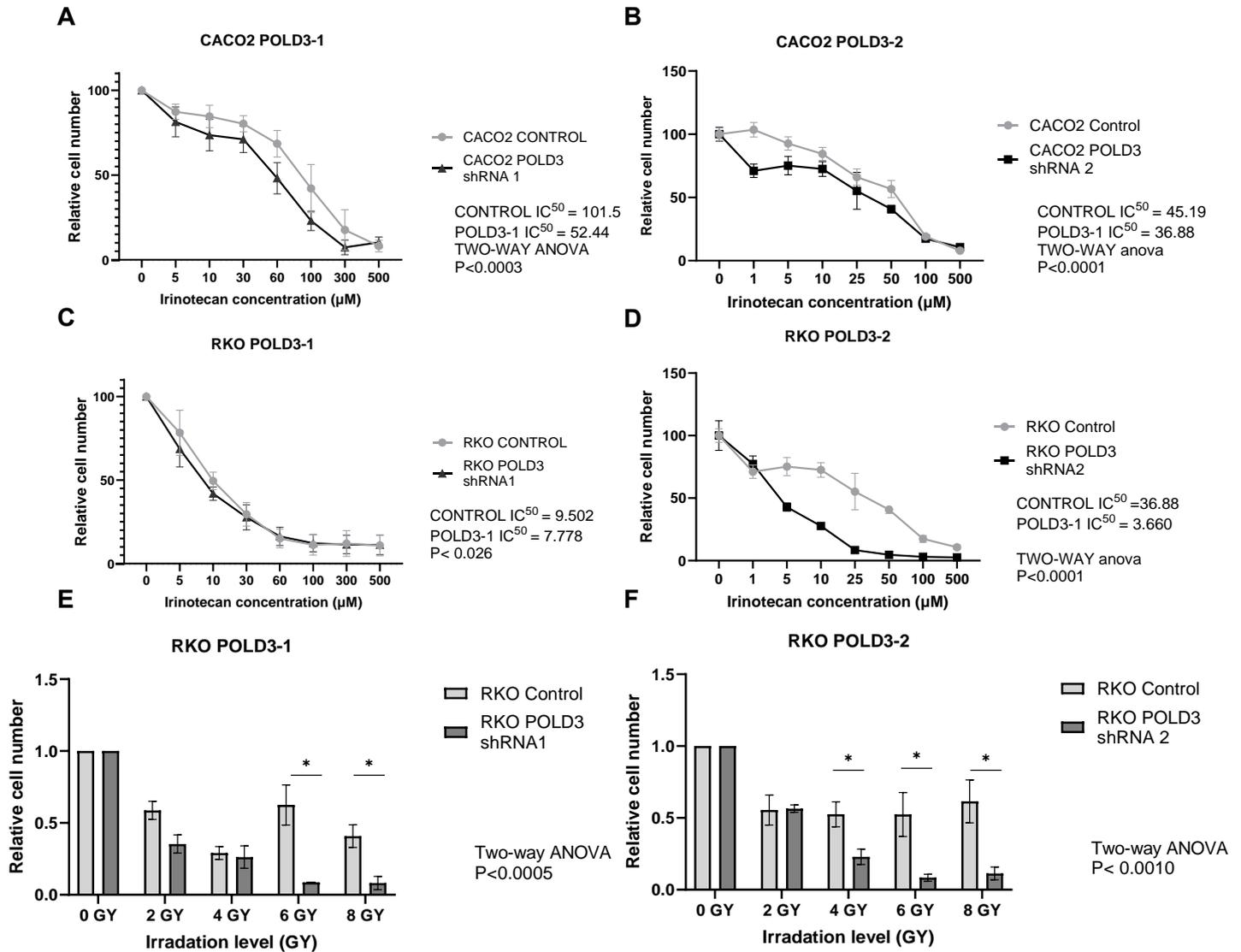


Figure 33: A) Drug dose response curves of CACO2 cells with POLD3 knockdown by shRNA 1 and Irinotecan. N=3. Two-way ANOVA was used to find differences between curves, $P < 0.01$. Linear regression analysis was used to find the IC_{50} of these cell lines with baseline = 0. B) Drug dose response curves of CACO2 cells with POLD3 knockdown by shRNA 2 and Irinotecan. N=3. Two-way ANOVA, $P < 0.0001$. Linear regression analysis was used to find the IC_{50} of these cell lines with baseline = 0. C) Drug dose response curves of RKO cells with POLD3 knockdown by shRNA 1 and Irinotecan. N=3. Two-way ANOVA, $P < 0.05$. Linear regression analysis was used to find the IC_{50} of these cell lines with baseline = 0. D) Drug dose response curves of RKO cells with POLD3 knockdown by shRNA 2 and Irinotecan. N=3. Two-way ANOVA, $P < 0.0001$. Linear regression analysis was used to find the IC_{50} of these cell lines with baseline = 0. E) Cell count after irradiation of RKO cells with POLD3 knockdown by shRNA 1. N=3. T-test 6GY: $P < 0.05$, 8GY: $P < 0.05$. F) Cell count after irradiation of RKO cells with POLD3 knockdown by shRNA 2. N=3. T-test 4GY: $P < 0.05$, 6GY: $P < 0.05$, 8GY: $P < 0.05$. Error bars given as \pm SEM.

POLD3 knockdown causes DNA damage

Next, we evaluated the effect of POLD3 depletion of DSBs. POLD3 has been extensively shown to be required for repair of double stranded breaks and is instrumental in repair of single stranded breaks through break-induced replication (BIR) [275] [263] [276]. Therefore, we analysed the effect of POLD3 depletion on the initiation of DSBs. As seen in figure 3 A and B, in both CACO2 and RKO cell lines POLD3—knockdown resulted in accumulation of DSBs shown by accumulation of γ H2AX foci. As seen in Figure 34 C and D, in CACO2 cells, DSBs were linearly increased with increased levels of POLD3 knockdown shown by doxycycline induced shRNA 2 at 1 μ g/ml and 10 μ g/ml ($P < 0.05$ and $P < 0.0001$). In RKO cells, the highest level of POLD3 knockdown also significantly increased the presence of γ H2AX foci ($P < 0.001$).

Next, strand breaks (SBs) were analysed through alkaline-comet assay. CACO2 and RKO cells with POLD3—knockdown were treated for 2 weeks to induce SBs through POLD3 depletion. As seen in Figure 34, E CACO2 POLD3-- cells showed increased comet “tail” length after gel electrophoresis, indicating increased SBs by POLD3 depletion. Tail length was quantified, and Figure 34, F clearly shows increased tail length significantly in both CACO2 and RKO cell lines ($P < 0.0001$ and $P < 0.01$ respectively)

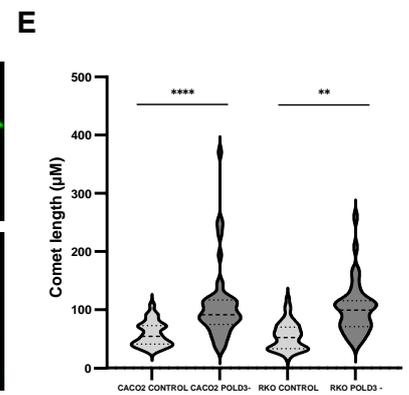
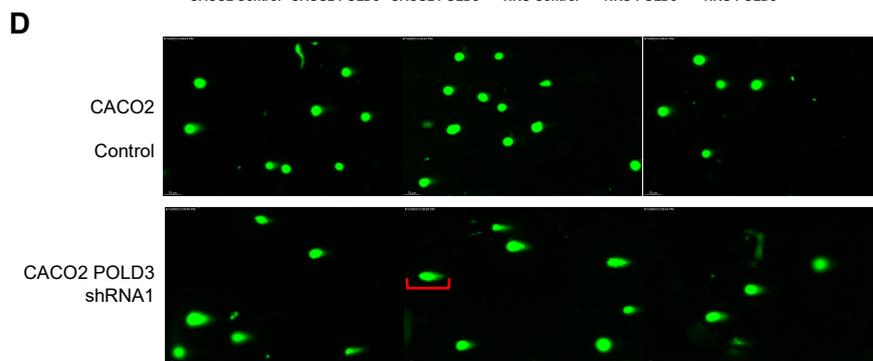
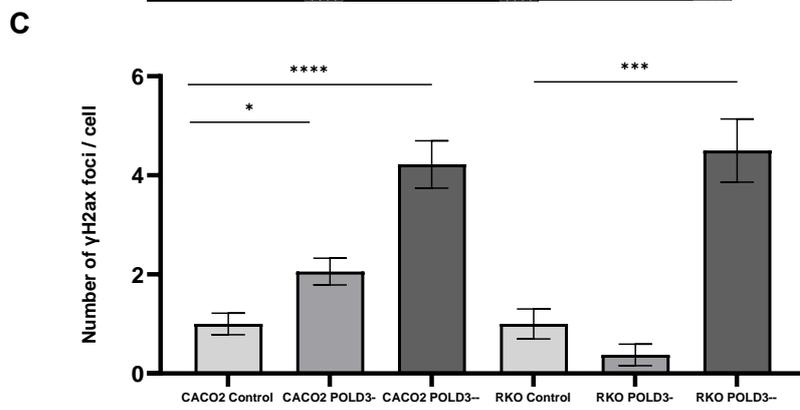
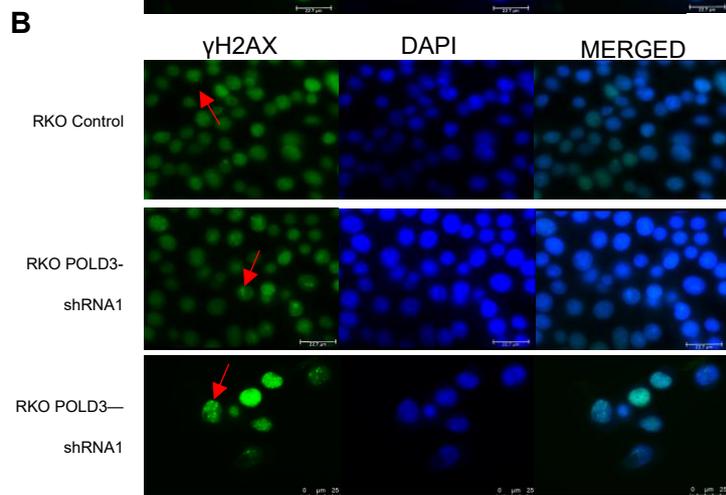
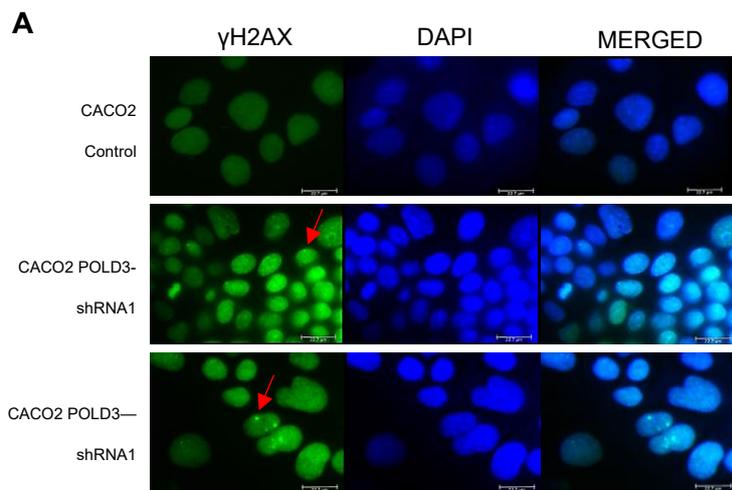


Figure 34: A) Immunofluorescence staining of γ H2AX on CACO2 POLD3-1 cells treated with doxycycline to include expression or a control. Cells were incubated for 72hrs before staining. B) Immunofluorescence staining of γ H2AX on RKO POLD3-1 cells treated with doxycycline to include expression or a control. Cells were incubated for 72hrs before staining. C) Quantification of γ H2AX foci in CACO2 POLD3-1 cells treated for 72 hours. Cells were treated with DMSO control reagent, or Doxycycline to induce POLD3 shRNA knockdown. POLD3- P< 0.05, POLD3-- P<0.0001. N=3. D) Quantification of γ H2AX foci in CACO2 POLD3-1 cells treated for 72 hours. Cells were treated with DMSO control reagent, or Doxycycline to induce POLD3 shRNA knockdown. POLD3-- P<0.001. N=3. E) Comet assay of CACO2 and RKO POLD3- cells induced for POLD3—knockdown and a DMSO control reagent. F) Quantification of Comet assay, CACO2 POLD3-- T-test P<0.0001. RKO POLD3-- T-test P<0.01. N=3. Error bars given as \pm SEM. Quantification carried out on Image J.

POLD3 knockdown induced DNA damage results in upregulation of DNA repair.

We then assessed whether POLD3 knockdown would lead to an increase in DNA repair pathway signalling due to the formation of DSBs. RAD21 is central mediator of DNA repair via HR. Therefore, we analysed whether RAD51 expression increase during POLD3 knockdown. As seen in figure 26 A, RAD21 signalling increased during POLD3 knockdown, P<0.001 (Figure 35, B). This was replicated in the second shRNA, as seen in figure 4 C and D, P<0.0.01.

We also analysed the presence of ATM protein, which serves as a central transducer during DSB repair. Again, ATM signalling was increased during POLD3 knockdown in both shRNAs, P<0.0001, P<0.005 (Figure 35 E-H). Together these data indicates that POLD3 is essential for successful DNA replication, and knockdown of the PLD3 gene results in the formation of DSBs, and therefore initiation of the DNA repair pathway.

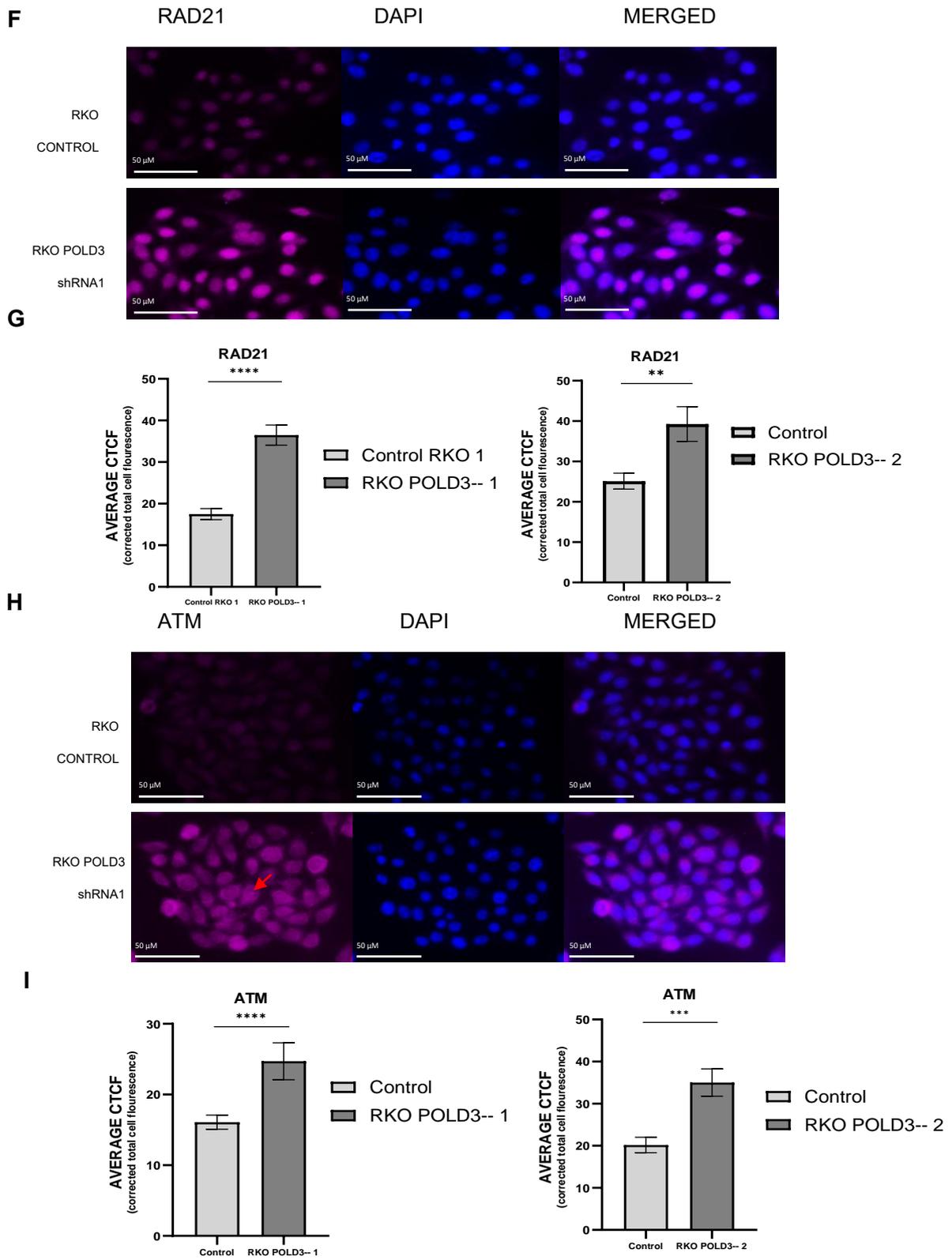


Figure 35: A) Immunofluorescence staining of RAD21 on RKO cells treated with POLD3-1 or a control. B) Quantification of RAD21 staining on RKO POLD3-1 and POLD3-2 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce POLD3 knockdown. T-test: $P < 0.0001$ and $P < 0.01$ respectively. $N=3$ C) Immunofluorescence staining of ATM on RKO cells treated with POLD3-1 H) Quantification of ATM staining on RKO POLD3-1 and POLD3-2 cells. Immunofluorescence given as Corrected Total Cell Fluorescence (CTCF). Cells were treated with DMSO control reagent, or Doxycycline to induce POLD3 knockdown. $P < 0.0001$ and $P < 0.005$ respectively. $N=3$ T-test. Error bars given as \pm SEM.

POLD3 knockdown reduces telomere length.

Telomere length is regulated by *TERT* reactivation and the ALT telomere pathway. There is conflicting evidence as to whether these pathways occur strictly separately, or if they can both be active within cell lines. We wished to investigate if POLD3 loss modifies telomeres, and if so by which mechanism. Therefore, we analysed how POLD3 depletion may affect telomere length in cancer cell lines. As seen in figure 36 A and B, both CACO2 and RKO cells showed a dramatic decrease in telomere length after 15 days treatment with POLD3 knockdown. Telomere length was decreased in line with increased POLD3 depletion, suggesting a direct link between POLD3 levels and telomere length. This was significant in CACO2 cells at 1 µg/ml (POLD3-) and 10 µg/ml (POLD3--) doxycycline induction of the POLD3 shRNA ($P < 0.05$, and $P < 0.01$). Similarly, in RKO cells the same pattern was observed ($P < 0.05$, $P < 0.01$).

Next, *TERT* levels in our POLD3 knockdown cell lines were tested. Surprisingly, in all 4 of our tested cell lines, we found a significant increase in *TERT* expression during POLD3 knockdown compared to control cells (Figure 36 C) ($P < 0.05$). This would conflict with the shorter telomere ends observed, however, previous research has shown that when telomeres become short in normal human cells, *TERT* is reactivated to resolve shortening telomeres [279]. Therefore, POLD3 knockdown may be causing telomeres shortening through other mechanisms, which in turn activates *TERT* expression.

Therefore, we sought to analyse the presence of ALT telomere maintenance in our POLD3 knockdown cell lines. To achieved this, we measured the presence of C-circles at telomere ends, which are required by POLD3 mediated ALT telomere maintenance. As seen in figure 36 D, there was an observable decrease in the presence of c-circles in POLD3 knockdown cell lines. C-circles were presented as a % of c-circle's present in cell line U2OS, which is known to harbour high ALT activity [280]. This was significant in CACO2 cells with shRNA 1 ($P < 0.05$). One way ANOVA across all cell lines revealed a significant change in the presence of C-circles ($P < 0.05$). Lack of significance in individual cell lines may be attributed to the high variability of this method of C-circle analysis. Future work could be done using blotting techniques to further analyse c-circle presence, as previously described [192]. This data would

suggest that POLD3 knockdown negatively effects the ALT mechanism, which may decrease telomere length, and therefore increase activation of *TERT* in compensation.

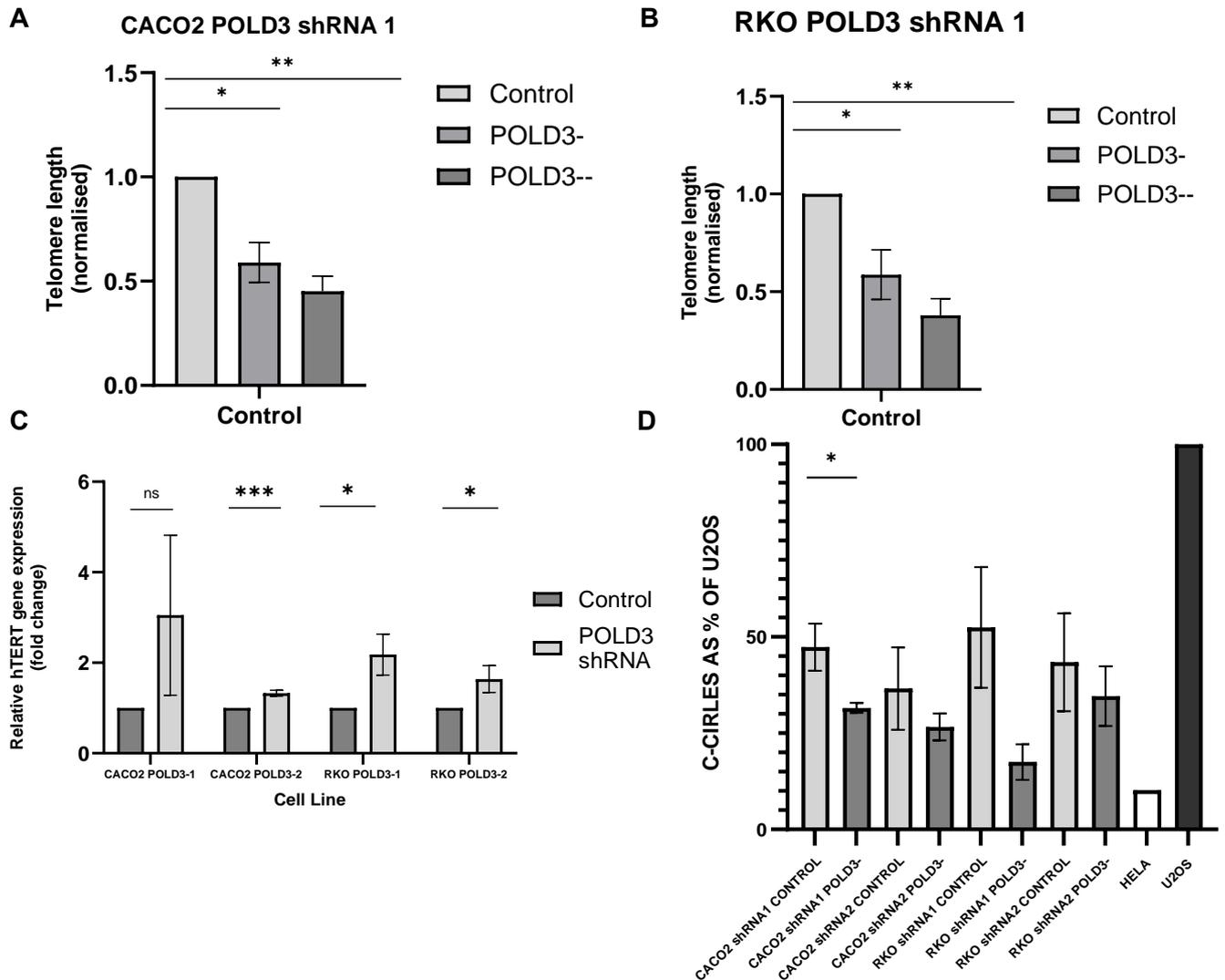


Figure 36: A) Telomere length analysis of CACO2 cells with POLD3 knockdown by shRNA 1. Doxycycline was administered at concentrations of 1 $\mu\text{g/ml}$ or 10 $\mu\text{g/ml}$ and refreshed every 72hrs for 15 or 30 days. Control samples were given DMSO and treated in the same manner. T-test; 15 days control to POLD3- $P < 0.05$, POLD3— $P < 0.01$, 30 days control to POLD3— $P < 0.01$. $N = 3$. B) Telomere length analysis of RKO cells with POLD3 knockdown by shRNA 1. Doxycycline was administered at concentrations of 1 $\mu\text{g/ml}$ or 10 $\mu\text{g/ml}$ and refreshed every 72hrs for 15 or 30 days. Control samples were given DMSO and treated in the same manner. T-test; 15 days control to POLD3- $P < 0.05$, POLD3— $P < 0.01$, 30 days control to POLD3— $P < 0.0001$. $N = 3$. C) qPCR analysis of hTERT mRNA expression in CACO2 and RKO cells given POLD3 knockdown by shRNA 1 or 1 compared to a control. CACO2 POLD3-2 $P < 0.003$, RKO POLD3-1 $P < 0.029$, RKO POLD3-2 $P < 0.04$. D) C-circle analysis on CACO2 and RKO cells given 10 $\mu\text{g/ml}$ induction of shRNA 1 and 2 to knockdown POLD3 expression after 7 days incubation. CACO2 shRNA 1: students T-test $P < 0.0477$. One way ANOVA across all cell lines: $P < 0.035$. $N = 4$. Error bars given as \pm SEM.

POLD3 expression in patient tumours and survivability

Finally, we sought to look at POLD3 expression and survival rates of CRC patients (Figure 37). For this analysis, data was provided through online publicly available data sets [255], totalling 1336 patient samples. *POLD3* high expression in a cohort of Colorectal adenocarcinoma (COAD) patients totalling 419, comprising all stages of cancer development, resulted in a higher survival than *POLD3* low expressing patients in a cohort of 917 patients. This was significant to an FDR value on 0.01. This would suggest that patients have a better outcome probability with high *POLD3* expression. We then looked at survival outcomes in a stage specific manner. In stage 1, there appeared to be a reduction in survival with high *POLD3* high expression, but this was not significant at FDR: 1. In stages 2,3, and 4, this was reversed, and *POLD3* high expression resulted in better survival in patients, which was significant at stage 2 with FDR: 0.05.

We then looked at whether POLD3 protein and gene expression was increased or decreased in patient samples. Data was acquired by publicly available datasets with genetics data obtained from TCGA and proteomics data acquired from CPTAC. *POLD3* gene expression was significantly increased in tumour samples compared with normal samples in patients with CRC ($P < 0.025$). Furthermore, POLD3 protein expression was significant increase in tumour samples ($P < 0.0005$).

Finally, we looked at promoter methylation, and we found that *POLD3* promoters were demethylated, or Hypo-methylated in tumour tissue compared to surrounded normal tissues, indicating an increase in expression levels in tumour samples.

Collectively, this data suggests that both POLD3 protein and gene expression is increased in CRC tumours, however higher levels of POLD3 are beneficial to patient survival outcomes. Although high POLD3 levels may predict tumour formation, low levels in POLD3 in tumour tissue is a negative indicator for survival, suggesting an interesting role for POLD3 in tumour initiation and progression.

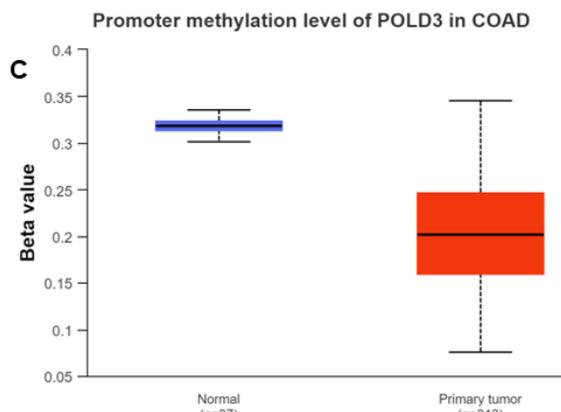
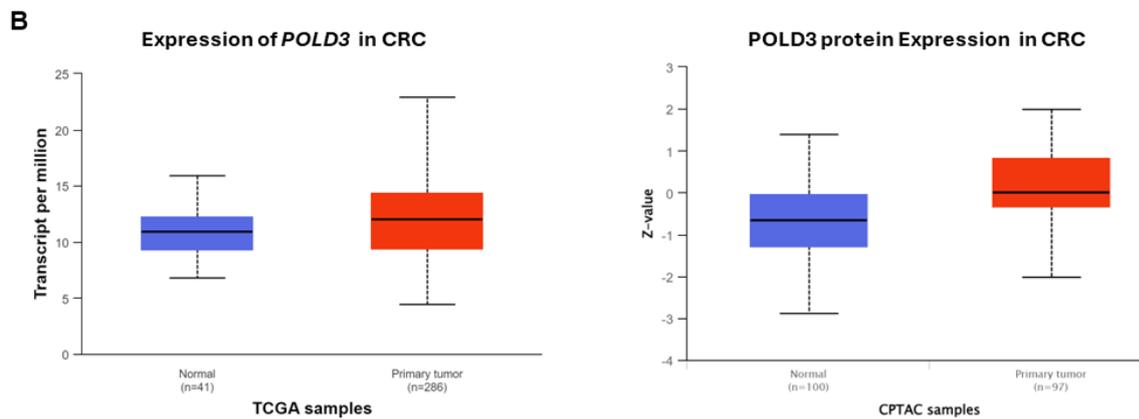
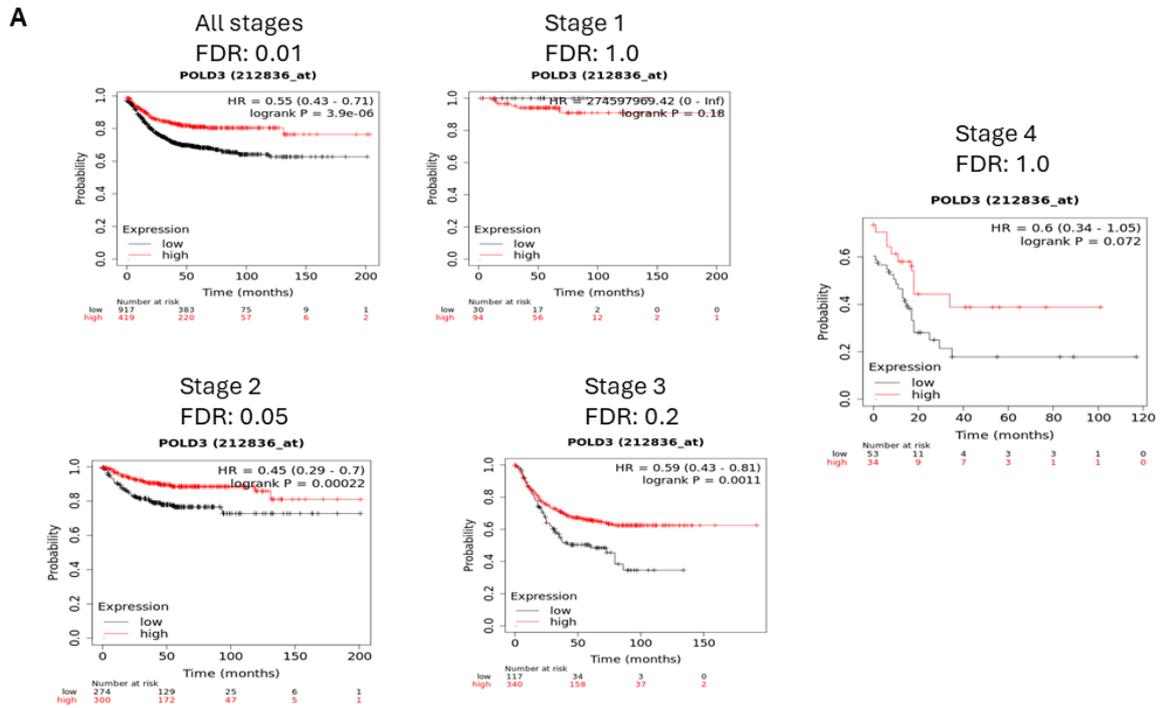


Figure 37: A) Kaplan-Meier survival curves for POLD3 High and low expression in patients with Colorectal adenocarcinoma. Data presented as stage specific, or all stages combined. In all stages combined POLD3 high expression resulted in increases survival, FDR: 0.01 N= 1336. Stage 1 no significant difference, FDR: 1 N=124. Stage 2 POLD3 high expression increased survival, FDR: 0.05 N=574. Stage 3 no significant difference, FDR: 0.2 N=457. Stage 4 no significant difference, FDR: 1 N=87. B) Expression of *POLD3* and POLD3 protein in normal tissue compared with primary tumour ($P < 0.025$ and $P < 0.0005$ respectively). Promoter methylation of *POLD3* in normal and primary tumour ($P < 1.624e-12$). Data obtained from TCGA.

Discussion

DNA polymerases are crucial in DNA synthesis and are required for cellular proliferation and DNA damage repair. Furthermore, DNA polymerases may be influential on telomere maintenance and longevity, a hallmark of cancer cells. Here, we have shown that POLD3 knockdown not only reduces proliferation but sensitises cells to chemotherapy through loss of DNA repair mechanisms. However, we have also shown the POLD3 knockdown results in DNA damage, which is associated with genome instability, a hallmark associated with tumorigenesis. We have further shown the POLD3 knockdown reduces telomere length, another important marker of genome instability, suggesting an important role for POLD3 in growth and maintenance of cancer cell lines.

Here we have clearly shown that POLD3 loss results in impaired proliferation of CRC cells, evidenced by slower growth over time, reduced colony forming capacity, and stalling of cell cycle progression. This stalling in S phase progression is perhaps due to the formation of DSBs observed. The intrinsic need for cancer cells to proliferate to form tumours relies on DNA polymerases, including POLD3, suggesting that loss of POLD3 may result in slower growing tumour phenotypes. Indeed, complete knockout of POLD3 in mice has been shown to cause embryonic lethality, suggesting POLD3 is essential for maximum DNA replication ability [166].

POLD3 knockout was achieved to approximately $\frac{1}{2}$ of the control mRNA expression before cell death occurred (Figure 30). This would suggest complete removal of POLD3 is impossible, as POLD3 is essential for replication and complete loss would result in cell death. POLD3-depleted cells were then treated for 96hrs, and cell numbers were assessed every 24 hours to produce growth curves under POLD3 depletion. As seen in figure 30, upon POLD3 depletion, cell numbers were reduced to half that of a control in CACO2 POLD3-2 and RKO POLD3-1 and 2, whereas in CACO2 POLD3-1 cells were reduced by $\frac{2}{3}$ rds ($P < 0.0001$). This reflects mRNA levels of POLD3 depletion, which could not be sustained beyond approximately $\frac{1}{2}$ – $\frac{1}{3}$ knockout. Colony forming potential was also observed, and upon POLD3 knockdown colonies were far smaller in size and number, suggesting inability of cells to proliferate.

Upon release from G1 phase synchronisation by serum addition, POLD3 cells were delayed in S phase progression, according to the essential function of Pol δ [177]. POLD3 depleted cells had a higher percentage stalled in S phase and consequently a lower percentage of cells had transition to G2 phase. This is likely due to DNA damage inflicted by POLD3 depletion, which activated S/G2 checkpoints to stall cell cycle progression.

This is supported by Figure 34, which shows an accumulation of γ H2AX foci, indicating increased DNA damage manifested by DSBs. In CACO2 cells, γ H2AX accumulation correlated with the level of POLD3 depletion, suggesting that lower POLD3 expression resulted in increased DNA damage. In RKO cells this difference was not as apparent, and γ H2AX foci only accumulated during the highest treatment of POLD3 knockdown, suggesting that perhaps POLD3 is less instrumental in DNA repair compared to other DNA repair pathway proteins. DNA damage was also apparent by increased comet tail length, as seen in figure 34 E.

This increase in DNA damage resulting from POLD3 loss exacerbated that of DNA damage caused by chemotherapy agents. As seen in figure 33, treatment with chemotherapy drug Oxaliplatin, which causes the formation of DSBs, resulted in cell death of CRC cell lines with increasing intensity. The addition of POLD3 knockdown enhanced this effect, perhaps through the inability for cancer cells to repair damaged DNA, resulting in a decrease in the IC50 Concentrations of this drug. This was most evident in RKO POLD-2 cell lines, where the number of live cells at the IC50 level of oxaliplatin was 1/10 in POLD3 knockdown cells to that of a control. This suggests that not only does POLD3 cause DNA damage but enhances DNA damage caused by chemotherapy agents. Notably, this was replicated by irradiation treatment, which also saw a decrease in cell survival after radiation treatment in POLD3 depleted cells, suggesting a synergy between DNA de-stabilising agents such as chemotherapy and radiation, and POLD3 knockdown.

Depletion of POLD3 was also shown to increase common DNA repair pathway proteins, ATM and RAD21, which are essential in repair of DSBs by HR. This increase in ATM and RAD21 expression could be due to increased DNA damage by POLD3 depletion stimulating initiation of DNA repair pathways. ATM binds to and aids in the phosphorylation of γ H2AX, thus serving as a central transducer to recruit further repair complexes, such as MRN [236] Damaged DNA is removed by an exonuclease, and repair is initiated by RAD21, before new DNA synthesis is mediated by POLD3 [281]. As seen in figure 35 A-E there was a stark increase in both ATM and RAD21 expression, suggesting activation of DNA repair pathway by HR. However, since we observed the presence of γ H2AX foci after 72hrs in POLD3 depleted cells (Figure 34), we can theorise that POLD3 depletion prevents DNA damage repair from succeeding, despite upregulation of repair proteins earlier in the sequence.

Next, we analysed the effects of POLD3 depletion on telomere length. Telomeres are majorly maintained by the Telomerase pathway, mediated by the catalytic TERT protein, which is activated in 90% of tumours. Alternatively, telomeres can be maintained by the ALT pathway,

which utilises BIR, in which POLD3 is essential. Our cell lines have previously been shown to be TERT positive, with high levels of TERT found in CACO2 and RKO cells [282][283].

Interestingly, we have shown that POLD3 depletion resulted in telomere length shortening, despite high TERT expression. Telomere length was approximately 1/2 of that of the control upon POLD3 depletion after 15 days consecutive treatment. Surprisingly, after 30 days, although telomeres are still shortened in POLD3 depleted cells, this reduction is less than that of 15 days, suggesting two potential causes. 1) Our population of cells are heterozygous, and many of those with POLD3 knockdown had died, increasing average telomere length. 2) POLD3 cells adapted to shortened telomeres and increased telomerase activation to rebuild telomere ends.

Surprisingly, we have shown that POLD3 depleted cells had increased *TERT* expression. This is perhaps caused by activation of the *TERT* gene in response to short telomere recognition [279]. We can therefore speculate that telomere shortening due to POLD3 depletion does not rely on Telomerase mediated telomere maintenance, as *TERT* was found to increase upon POLD3 knockdown.

We then looked at activation of the ALT telomere pathway. This was performed via analysis of the presence of c-circles, which are an ALT pathway hallmark. Interestingly, we observed a decrease in c-circles upon POLD3 knockdown, suggesting that POLD3 loss results in a lack of ALT activity. This reduction in ALT provides an explanation for the apparent reduction in telomere length observed.

Therefore, we can speculate that in our cell lines ALT occurs consecutively with telomerase-mediated maintenance, due to the presence of C-circles and *TERT* expression. This would suggest that when telomeres are shortened due to ALT failure or DNA damage, TERT is activated in response, however it is apparent this is not sufficient to sustain telomere lengths.

We have shown that POLD3 protein and gene expression are over-expressed in tumour tissues compared to adjacent normal tissues. It is unclear whether this is due to oncogenic effect of POLD3, or the selective pressure for high POLD3 expression for cancer cells to deal with the replication load that results from hyper-proliferation. Interestingly, we have also shown that high POLD3 expression in tumour tissue confers with increased survival in CRC patients. This may be explained by our findings on POLD3 depletion causes genome instability manifested by increased DNA damage and shortened telomeres, which can contribute to the genetic instability of cancer cells often associated with tumour progression [284].

CIN+ patients have been shown to have worse outcomes compared to non-CIN and MSI+ patients, and high genetic instability resulting in more-aggressive tumours [7][285]. Therefore, patients with low POLD3 levels may have worse outcome due to this increase in genetic instability, presenting POLD3 loss an indication of worse patient outcomes, despite its over-expression in CRC.

Limitations and future work

In order to confirm our findings of slower replication theorised to be from stalling in S phase of the cell cycle, DNA combing of fork velocity could be performed. Further to this, analysis of markers of S-phase progression impairment, such as cyclin D and cyclin E, while imbalances in the cyclin A/CDK2 and cyclin B/CDK1, may reveal stalling of the cell cycle in S phase as theorised in this study.

We have also shown that POLD3 knockdown reduced telomere length, potentially through a reduction in the LAT telomere pathway in which POLD3 is essential. To strengthen our findings of shortened telomeres in POLD3 knockdown cells, fluorescence in situ hybridization (FISH) could be used to visually assess telomere lengths in POLD3 knockdown cells.

Conclusion

Together, this data presents an interesting role of POLD3 in the context of CRC. POLD3 depletion caused decreased proliferation, clonogenicity, and stalling of cell cycle progression, which is often a positive target for the treatment of CRC. However, low POLD3 expression results in poor patient survival, and POLD3 knockdown results in a manifestation of DNA instability, shown by increased presence of DSBs and activation of DNA repair pathways. DNA instability is a hallmark of CRC and is a known increase for tumorigenicity. Furthermore, we have shown that POLD3 depletion results in shortened telomeres, which may increase genome instability in our cell lines.

However, it should be noted that targeting of telomere maintenance is a therapy currently utilised [286], in which cancer cells are abolished through removal of telomere lengthening pathways. Telomere maintenance abilities/activation is an almost universal principle of cancer cells, which allows them to gain immortality status. Therefore, POLD3 depletion mediating telomere shortening may pose a beneficial effect in the treatment of CRC.

In conclusion, POLD3 presents an interesting role in CRC developments, by both increases and decreases tumorigenic characteristics. The decrease in proliferation observed by POLD3 knockdown cells would suggest an anti-tumorigenic affect in earlier stages of disease, however the decrease in genome stability may potentially lead to hyper-mutated cancer cells which could develop more-pro-tumorigenic affects in later stages of the disease.

Chapter 7: The effects of *POLD3* and *CHRD2* in Colorectal cancer

POLD3 and *CHRD2* lie adjacent to one another on chromosome 11, and have both been highlighted, along with 77 other loci, as harbouring areas containing risk SNPs that have been predicted to significantly increase the risk of CRC (Figure 38) [287] [288,289]. Given that only 79 such loci have been identified, and *POLD3* and *CHRD2* lie together, we decided to look at whether these genes are differentially expressed in a coordinated fashion, or any potential interaction these genes may have at the transcriptional or functional level.

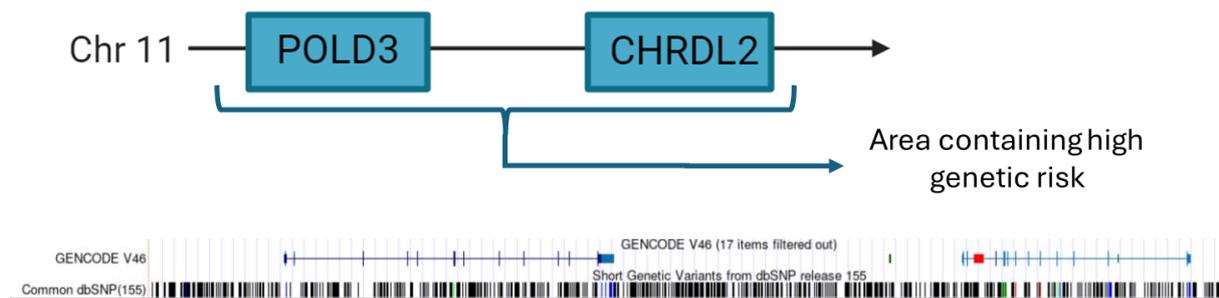


Figure 38: Diagram depicting *POLD3* and *CHRD2* region, and SNPs in the *POLD3*/ *CHRD2* region. Blue indicates non-coding transcript or Untranslated Region (UTR) variants, green indicates synonymous variants, and red indicates Protein-altering variants and splice site variants.

POLD3 and CHRDL2 do not correlate in expression

First, we determined whether *POLD3* and *CHRDL2* gene expression would correlate in patients with CRC. TCGA and GTEX data were utilised, and protein and transcript expressions were measured. As seen in Figure 33 A and B, both *CHRDL2* and *POLD3* are overexpressed in tumour tissues of patients with COAD. This supports previous findings that have shown *CHRDL2* is overexpressed in CRC [144], and also presents evidence of previously un-known *POLD3* overexpression in CRC. Furthermore, *POLD3* promoter regions are hypo-methylated in tumour tissues compared to normal tissues, suggesting increased transcription of the *POLD3* gene during cancer.

We then looked at whether these genes would correlate in expression due to their close proximity. However, we found no correlation between *POLD3* and *CHRDL2* in tumour tissues or normal tissues (TCGA) and only a slight positive correlation in normal tissues (Gtex) (Figure 39 C,D, E). Furthermore, *POLD3* promoters were hypomethylated in CRC, whereas *CHRDL2* was hyper-methylated (Figure 39 F, G). This would suggest there is no correlative effect between these two genes in CRC at the expression level.

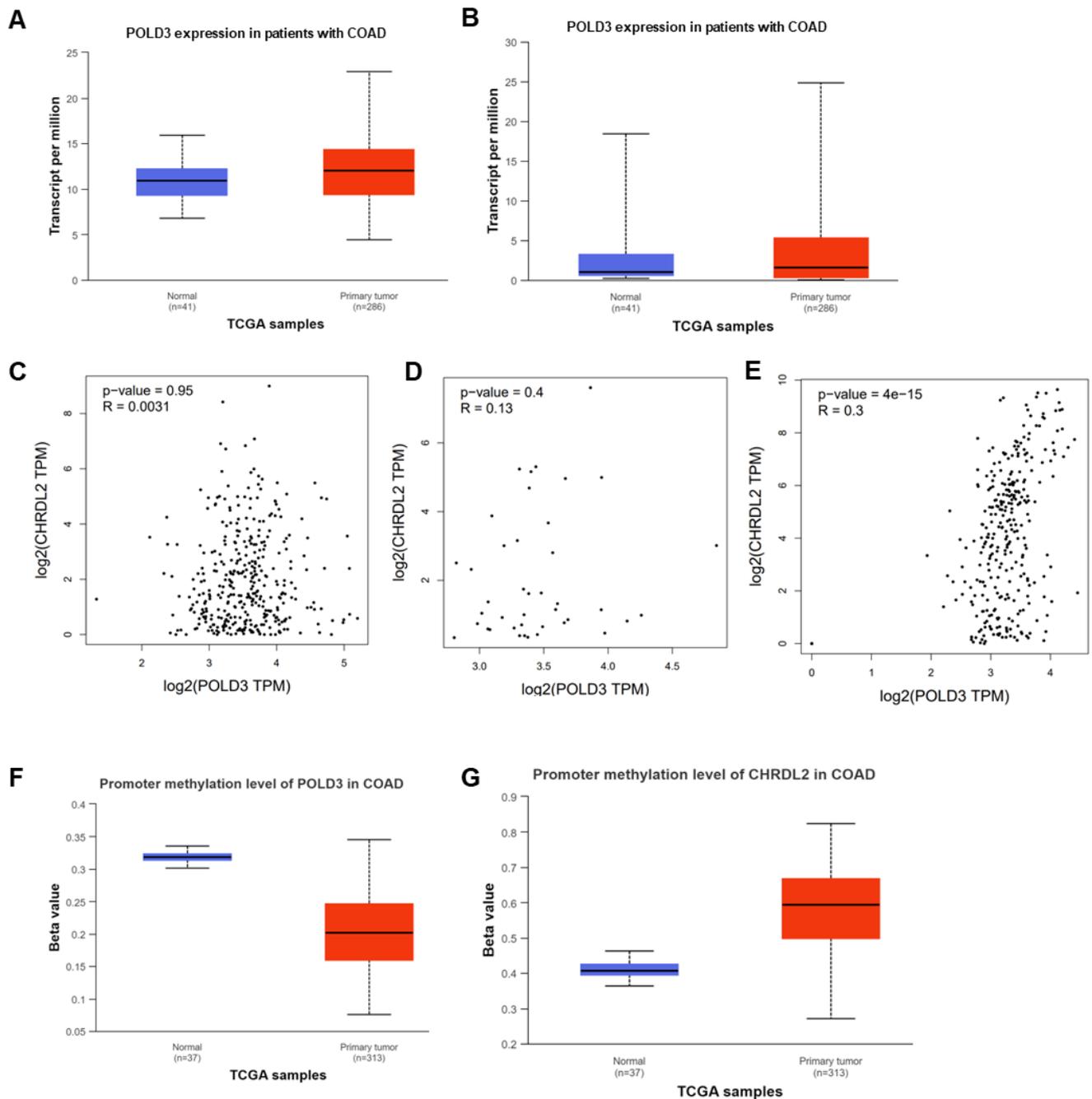


Figure 39: A) POLD3 gene expression in primary tumour and normal tissue $P < 0.025$. B) CHRDL2 gene expression in primary tumour and normal tissue $P < 0.854$. C) Correlation of POLD3 and CHRDL2 gene expression in tumour tissue in patients with COAD (TCGA). $P < 0.95$, $R = 0.0031$. D) Correlation of POLD3 and CHRDL2 gene expression in normal tissue (TCGA). $P < 0.4$, $R = 0.13$. E) Correlation of POLD3 and CHRDL2 gene expression in normal tissue (Gtex). $P < 4e-10$, $R = 0.3$. F) Promoter methylation levels of POLD3 In COAD and normal tissues. $P < 1.624e-12$. G) Promoter methylation levels of CHRDL2 in COAD and normal tissues. $P < 1 E-12$.

POLD3 knockdown and CHRDL2 overexpression results in a reduction of cell proliferation.

Next, we combined our *CHRDL2* overexpression vector with our *POLD3* knockdown cell lines to study any potential affects these genes may play together. As seen in figure 40 A, *CHRDL2* overexpression was successful in all cell lines. However, only in CACO2 POLD3-1 and RKO POLD3-2 we saw a reduction in *POLD3* expression. This is perhaps due to loss of the shRNA expression vector in these cell lines after multiple passages and puromycin selection. Therefore, we used CACO2 POLD3-1 and RKO POLD3-2 in future experiments, termed CHRDL2+/POLD3-.

As seen in panel B, CHRDL2+/POLD3- cells showed a dramatic decrease in cellular proliferation, to less than half of the control cell lines confluency at 10 µg/ml doxycycline. Furthermore, clonogenic assays showed near complete abolition of colonies at 10 µg/ml doxycycline induction of POLD3 knockdown and CHRDL2 overexpression. This would suggest an additive effect of POLD3 knockdown and CHRDL2 overexpression, as this reflects a greater decrease in proliferative ability compared to CHRDL2+ or POLD3- alone (Figures 5 A, 31 A).

Next, we looked at chemotherapy resistance in these CHRDL2+/POLD3- cell lines. Previously, we have seen that CHRDL2 overexpression increased chemotherapy resistance due to activation of repair pathways, and an increase in the cancer-stem-cell phenotype. We have also seen that POLD3 knockdown sensitizes cells to chemotherapy, perhaps due to the inability of cells to repair chemotherapy-induced DNA damage, due to lack of POLD3 which is a known key player in HR repair. Interestingly, when together, CHRDL2 overexpression and POLD3 knockdown cancelled out their respective influences on chemotherapy response, and no observable difference was seen between control and CHRDL2+/POLD3- cell lines (Figure 40 E, F).

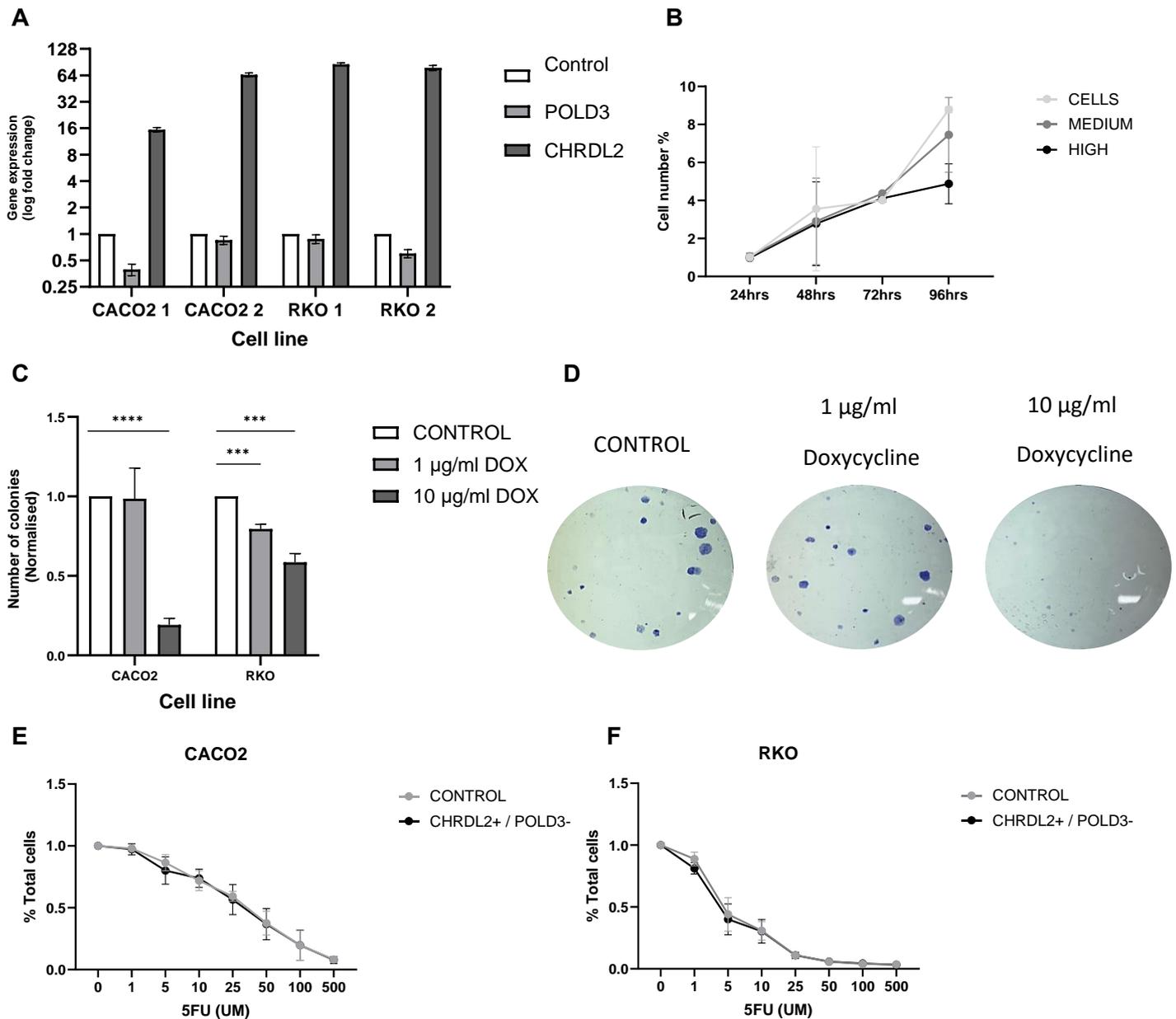


Figure 40: A) qPCR of POLD3 and CHRDL2 mRNA expression given as fold change in CACO2 and RKO cells lines with POLD3 knockdown by shRNA 1 or 2 and CHRDL2 overexpression. Cell lines were grown with a DMSO control of doxycycline at 10 µg/ml to induce expression. T-test between control and doxycycline: CACO2 1 POLD3 $P < 0.005$, CACO2 1 CHRDL2 $P < 0.00001$, CACO2 2 POLD3 $P < 0.1802$, CACO2 2 CHRDL2 $P < 0.0001$, RKO 1 POLD3 $P < 0.286$, RKO 1 CHRDL2 $P < 0.0001$, RKO 2 POLD3 $P < 0.0033$, RKO 2 CHRDL2 $P < 0.0001$. N=3. B) Time course of MTT assay of cellular proliferation on CACO2 POLD3-1/CHRDL2+ and RKO POLD3-2/CHRDL2+ cell lines. Two-way RM ANOVA $P < 0.0001$. C) Quantification of number of colonies in CACO2 POLD3-1/CHRDL2+ and RKO POLD3-2/CHRDL2+ cell lines after 1 µg/ml and 10 µg/ml doxycycline induction. Students T-test: CACO2 10 µg/ml $P < 3.85573E-05$, RKO 1 µg/ml $P < 0.002261$, 10 µg/ml $P < 0.00156$. E) Drug dose response curves of CACO2 cells with POLD3 1 knockdown and CHRDL2 overexpression and 5FU. N=3. F) Drug dose response curves of RKO cells with POLD3 2 knockdown and CHRDL2 overexpression and 5FU. N=3. Error bars given as \pm SEM.

Discussion

Here, we have attempted to elucidate any correlation between the overexpression of *CHRD2* and *POLD3* in the context of Colorectal cancer. Previously, genome-wide meta-analysis highlighted several loci containing risk SNPs significantly increasing the risk of CRC. Several of these SNPs lie in the region of *POLD3/CHRD2*, at 11q13.4. SNP rs3824999, which lies in the intronic region of *POLD3*, was ($3.65E-10$) predictive of increased CRC risk [287]. Imputation indicated a further, stronger association at SNP rs72977282, which is upstream of both *POLD3* and *CHRD2*, mapping 5' to *POLD3*, which was also significantly associated with CRC risk. Subsequently a further independent risk SNP, rs4944940 was identified within 11q13.4, within the intronic region of *CHRD2*. Meta analysis further clusters these SNPs together, with high linkage disequilibrium, making it difficult to pinpoint a causal SNP.

Previously we have shown how overexpression of *CHRD2* may increase the risk of CRC through enhancing the cancer stem cell-phenotype, increasing chemotherapy resistance, and upregulation of the EMT pathway. We have also shown that *POLD3* loss may contribute to genome instability, a hallmark of many CRC cases, however it also contributes to slower replication, sensitization the chemotherapy, and shortened telomeres. Finally, we sought to combine our experimental models for *CHRD2* overexpression and *POLD3* knockdown to elucidate any potential cumulative effects these genes may have in the development of CRC.

We saw no correlation between *POLD3* and *CHRD2* overexpression in tumour tissue, despite both genes showing overexpression in CRC. We did see a slightly positive correlation between these two genes in normal tissue (GTEx) however this was very small ($R=0.3$). Therefore, we can elucidate that these genes are not regulated together. Furthermore, there was no correlation in promoter methylation of these genes, with *POLD3* hypo methylated, and *CHRD2* hyper methylated.

We then looked at the functional role of these genes in 2D CRC cell lines. *POLD3* knockdown and *CHRD2* overexpression resulted in less than half cell confluency compared with normal control cell lines, and nearly abolished colonies formed. However, *POLD3* knockdown and *CHRD2* overexpression appeared to cancel out the effects upon chemotherapy response, indicating that *CHRD2*'s pro-survival mechanism is counteracted by the inability to activate repair mechanism mediated by *POLD3*. It could be seen that overexpression of *POLD3* may aide in the survivability of *CHRD2* overexpressing CRC cells, through intrinsic repair mechanisms.

Limitations and future work

This preliminary data does not suggest any correlative relationship between POLD3 and CHRDL2 in CRC. However, it should be noted that the risk SNPs for the genes are rare, particularly those in the CHRDL2 locus, and by taking whole patient data we may not see a correlative relationship. Further analysis would be required on patients with the identified risk SNPs, to see if this generated a correlation between POLD3 and CHRDL2. Functional analysis on risk SNPs could also be implemented to identify variation in RNA expression or splicing.

Conclusion

This data does not suggest any correlative relationship between the expression patterns on POLD3 and CHRDL2. CRC cells depleted of POLD3 and with CHRDL2 overexpression showed a reduction in cellular proliferation and clonogenicity, but no effects on chemotherapy resistance were observed.

Chapter 8: Thesis Discussion

Colorectal cancer is one of the leading causes of cancer related death worldwide, with 1.5 million reported cases annually [1]. CRC develops from aberrant proliferation of epithelial cells, and often arises from pathways of genetic and chromosomal instability. WNT signalling is one of the most commonly affected pathways in CRC and is present in >90% of all cases. CRC can be caused by inherited mutations that account for 10-16% of instances, however the majority of cases are formed through sporadic mutation in epithelial cells [4]. Recent genetic screening by GWAS of CRC patients has attempted to elucidate causal SNPs which may increase the risk for CRC development. 2 loci containing significantly associated SNPs were identified as potentially causing increased CRC risk, *POLD3* and *CHRD2* [287][184]. These genes offer potential targets for CRC treatment, due to their functional role as a Polymerase Delta subunit and a BMP inhibitor respectively. Therefore, within this thesis we have attempted to elucidate a functional role of these two genes in the context of CRC.

We have shown that *CHRD2*, a BMP antagonist, enhances the WNT signalling pathway through BMP inhibition. *CHRD2* has previously been identified as a BMP antagonist, which binds to the active domain of BMPs to prevent signalling through SMAD1/5 phosphorylation [144][146]. Through protein analysis we have shown that *CHRD2* overexpression indeed reduces phosphorylation of SMAD1/5, therefore decreasing BMP signalling. We have also shown through RNAseq analysis that there is a global modulation of BMP signalling, further confirming our findings.

Canonically, BMP and WNT signalling work in a counter gradient in the intestinal epithelium in order to maintain homeostasis. BMP, a pro differentiation pathway, has been shown to be highly expressed at the tips of villi, aiding in epithelial differentiation, whereas WNT signalling is located at the crypt base, in order to maintain stem cell populations [95]. We therefore theorised that by inhibiting BMP signalling through *CHRD2* overexpression, we would see an increase in WNT signalling. We confirmed this through staining of Beta-catenin in human CRC cell lines and mouse organoid models, in which *CHRD2* overexpression increased nuclear localisation of Beta-catenin, which is a hallmark of WNT signalling. GSEA analysis also supported these findings, in which WNT signalling was shown to be enriched in our 2D CRC cell lines.

WNT driven cells in the crypt are marked by the presence of LGR5, as well as other stem cells markers (SOX9, BMI1, MSI1) [82]. Analysis of stem cell markers supported our findings of increased WNT signalling, with an increase in LGR5, BMI1, OLFM4, MSI1 and SOX9

upregulated in both 2D cell lines and organoid models. RNAseq analysis has also shown an upregulation of BMI1 signalling, further supporting our findings.

Through increased WNT signalling, we have proposed that overexpression of CHRDL2 may alter cancer epithelial cells into a more robust cancer-stem cell phenotype. Conventional theory dictates that colon cancer originates from stem and progenitor cells that lie within the crypt the colonic epithelium [290][291]. The oncogenic transformation of the intestinal stem cell allows hyper-proliferation, self-renewal and ability to differentiate, which is crucial in the initiation and progression of the adenoma-carcinoma sequence for tumour formation. This is evidenced by knockout mutations of the APC gene in mice, which results in activation of WNT signalling and the formation of microadenomas within 3-5 weeks of activation. However, this response is only seen in LGR5+ cells in the intestinal crypt, and not seen in the committed lineages higher up in the villi structure or in the transit-amplifying zone, where microadenomas rapidly stall [291]. This is known as the bottom-up model, where activating APC mutations in intestinal epithelial stem-cells spread to the top of the crypt to form microadenomas [292].

However, another model of cancer initiation has been proposed, known as the top-down model. This model proposes that insult to differentiated epithelial cells in the inter-cryptal zone results in a stem-like transformation, possibly due to activation of inflammatory pathways such as NF- κ B, leading to the lateral spread of cancer cells which fills the intestinal crypt [293][294]. In cases of chronic inflammation, perhaps caused by western-diet habits, NF- κ B modulates WNT signalling inducing dedifferentiation of non-stem cells that acquire tumour-initiating capacity. This is not seen in cryptal stem cells, where inflammatory NF- κ B activation of WNT signalling restricts stem cell expansion, thus proposing a top-down cascade which does not originate from the intestinal stem-cell [294]. This is supported by further work which showed through genetic analysis that dysplastic cells at the tops of the crypts exhibited genetic alterations of APC and neoplasia-associated patterns of gene expression. Whereas the cryptal stem cells did not show these alterations and were not clonally related to the transformed cells of the villi [293].

Most recently, studies have shown that differentiated Paneth cells, upon inflammatory associated APC loss, initiate the formation of tumours, in both patients with inflammatory bowel disease and sporadic colon cancer [295]. Even without inflammatory stimuli, Paneth cells with targeted mutation in APC and KRAS led to a highly aggressive tumour phenotype with invasive morphology. These Paneth cells were able to de-differentiate, and under enhanced WNT signalling formed revival stem cells which initiated tumour formation. This data collectively suggests an additional origin for colon tumour formation, not from the

transformation of cryptal stem-cells as previously described, but from the de-differentiation of colonic epithelial cells which gives rise to genetically distinct tumours.

This top-down model would support our hypothesis that activation of the WNT signalling pathway through BMP inhibition, may cause a phenotypic switch of epithelial cells into a de-differentiated cell-type, which may cause a more aggressive, stem-like cell with enhanced survival capabilities. This is supported by our observed increase in stem-cell markers, and stem-cell marker pathways.

We then sought to investigate the response this potential phenotypic switch to a stem-cell phenotype may produce. We observed an increase in migratory potential in CHRDL2 overexpressing cells, as well as an increase in the production of migratory markers, such as IQGAP1 and EPCAM. RNAseq also confirmed an upregulation of the EMT pathway, which is associated with increased cell motility. Importantly, cell motility is a hallmark of cancer stem cells, which lends to their ability to invade and metastasise distant tissues, supporting our theory of a cancer-stem phenotype [219][20]. We also observed an increase in stem-cell qualities in intestinal organoid models which showed a more spherical, less differentiated phenotype.

Interestingly, we also observed a reduction in proliferation and ki67 expression in our CHRDL2 overexpressing cell lines and reduced size in our organoid models. Traditionally, when considering cancer stem cells we think of a hyper-proliferative state, which is especially true in the context of WNT driven-stem cells [84]. However, recent studies have given light to a subset of cancer stem-cells which are slow -cycling, which may explain the observed reduction in proliferation [226][220]. These slow-cycling stem cells are marked by BMI1, which we found to be upregulated in both our qPCR and RNAseq analysis. Slow-cycling stem cells have enhanced survivability and are proposed to be the route of cancer recurrence through their ability to evade traditional therapies.

In our CHRDL2 overexpressing cell lines and organoid models we observed an increase in survival when subjected to chemo and radio therapy. There are multiple mechanisms in which cancer cells may evade chemo and radio therapy. One mechanism is through the overexpression of ABC transporters which effluxes chemotherapy compounds out of the cell. An example of this is doxorubicin, which effluxes from breast cancer cells by ABCB1 and ABCG2 [296][227]. However, we saw no upregulation of transporter genes by CHRDL2 in our RNAseq data, suggesting that this is not the case. Another mechanism of chemotherapy resistance is ALDH activity. ALDH (Aldehyde dehydrogenase) is a cytosolic enzyme that oxidizes aldehydes and converts them into carboxylic acids [297], however there is only evidence that ALDH converts resistance to cyclophosphamides, which were not used in this

study. Cancer cells are also able to evade apoptosis induced by chemo and radio therapy, by over-expressing B-cell lymphoma-2 (BCL-2), which binds to apoptosis proteins, BCL2-associated-X-protein (BAX) and BCL-2 homologous antagonist killer (BAK) and impairing their ability to release apoptogenic proteins such as cytochrome c from the mitochondria [298]. BCL-2 has been shown to be important in colon-cancer resistance to 5FU, which is perhaps activated by the Akt signalling pathway [227]. We did not see an increase in expression of BCL-2 by CHRDL2 overexpression, but we did see an upregulation of the Akt signalling pathway, which may therefore lend to the increase in chemotherapy resistance.

Importantly, previous studies have shown that upregulation of the WNT signalling pathway may contribute to chemotherapy resistance. This fits with our presented data, which shows an increase in WNT signalling in our RNAseq data, and also through increased nuclear localisation of Beta-catenin. Studies have shown that activation of the WNT signalling pathway in cancer cell leads to resistance to cisplatin (orthologue to oxaliplatin) and 5FU [299][300]. However, it is unclear exactly how WNT signalling induces chemotherapy resistance. One method may be the activation of DNA damage response pathways, the final method of chemotherapy resistance found in cancer stem cells. We have extensively shown an upregulation of DNA damage response pathways in our cell lines upon CHRDL2 overexpression, such as the increase in ATM, RAD21, PCNA, ARTEMIS, and ERRC1. Together, these genes represent the 4 mechanisms of DNA repair following chemotherapy, indicating a global upregulation of DNA repair pathways by increased WNT signalling.

Chemotherapy resistance may also be due to this slow cycling phenotype. The action of chemotherapy relies on the synthesis of new DNA. For example, 5FU is a pyrimidine synthesis inhibitor, inhibiting the production of pyrimidines, required to synthesise new DNA [241]. Irinotecan is a topoisomerase inhibitor, which prevents the resolution of supercoiling during DNA synthesis. Oxaliplatin causes crosslinking through displacement of a chloride ion by a guanine, which interferes with DNA replication [301]. We have shown the CHRDL2 has a lower proliferation rate, shown by growth curves and a reduction in ki67 staining. This may suggest a method of resistance to chemotherapy. Indeed, other studies have shown that slow-cycling stem cells are chemo-resistant due to this very reason, making slow-cycling stem cells a problematic source of cancer reoccurrence [226][220]. However, we do see an upregulation of DNA repair pathways, so their may be a combinational cause of increased chemotherapy resistance by CHRDL2 overexpression.

Interestingly, POLD3 knockdown and CHRDL2 overexpression together had no effect on chemotherapy resistance. This supports the evidence that CHRDL2 overexpression causes CRC cells to have enhanced survival ability through upregulation of DNA damage repair

pathways, and not just due to slower growth. If CHRDL2 overexpression caused chemotherapy resistance purely based on a slow-cycling phenotype, then slower growth of POLD3 knockdown would enhance this effect. However, we see no difference in survivability when these models are together, which could be due to lack of repair due to POLD3 knockdown. This could also suggest that POLD3 is killing cells through apoptosis, however we see no difference in drug dose response curves, meaning this does not seem likely. Therefore, it appears CHRDL2 increases chemotherapy resistance mainly through DNA repair upregulation.

Previous research has shown the potential impact of BMP dysregulation and chemotherapy resistance. Knockdown of BMP2 has been shown to increase chemoresistance of the MCF-7 breast cancer cell line [302]. Similarly, it has been suggested that hypermethylation of BMP6 contributed to acquisition of drug resistance in breast cancer cells [303]. Furthermore, it has been reported that BMP2-based treatment increased the response to temozolomide in hypoxic drug-resistant glioblastoma derived cells [304]. It was also found that a BMP7 variant may reduce tumour growth and stem cell marker expression in subcutaneous and orthotopic glioblastoma stem-like xenografts, which is similar to what we have found in our CHRDL2 overexpressing cell lines [305]. Additionally, knockdown of BMP6 in breast cancer cells showed an increase in chemoresistance to doxorubicin through upregulating of multiple drug resistance (MDR)-1/ P-glycoproteins and activation of the ERK signalling pathway [306]. It is therefore clear that disruption of BMP signalling plays an important role in the acquisition of stem-like phenotypes of cancer cells, and the increase in chemotherapy resistance. This would indicate a potential therapeutic target, through BMP treatment or knockdown of BMP antagonism, to lower chemotherapy resistance acquisition in cancer cells.

Indeed, the use of antibodies to target BMP antagonists has also been investigated. Recent attention to the use of an anti-GREM1 antibody in various diseases has been discovered as a potential therapeutic target. Initially, an anti-GREM1 antibody that prevents GREM1-BMP interaction was developed by Novartis for the treatment of pulmonary artery hypertension [307]. Another study also suggested that anti-GREM1 could reduce GREM1-mediated A549 lung cancer, through reduction in cellular proliferation, invasion, and migration [308]. In CRC, a GREM1 human IgG4p monoclonal antibody, Ginisortamab, has recently been developed by UCB, and has been shown to neutralise GREM1 antagonism of BMP2. In a pre-clinical murine trials, Ginisortamab has been shown to have anti-tumour effects [309]. Transcenta have also developed a monoclonal GREM1 antibody (TST003) that could inhibit GREM1-mediated non canonical activation of FGFR1 phosphorylation [310]. It is clear that targeting GREM1 has produced exciting results, and the potential for

targeting GREM1 in cancer, especially in patients with HMPS, is undoubtable. Therefore, there is potential for the development of further CRC specific targeting of BMP antagonist, such as CHRDL2, for the treatment of CRC.

Together, this data suggests that BMP antagonism through CHRDL2 overexpression increases WNT signalling, and therefore increases a stem-like phenotype in 2D cells and organoid models. This manifests in increased resistance to chemo and radiotherapy, migration, and enhancement of EMT markers. However, we also observe a slower growth phenotype, suggesting a complex role for CHRDL2 in CRC initiation and progression.. Despite no significant changes in patient survival with CHRDL2 overexpression at the early stages of disease, as CRC develops, survivability drops with CHRDL2 overexpression. This is likely due to increased migration and invasive properties, and enhanced resistance to chemo and radiotherapy. This would suggest that CHRDL2 is an important biomarker for how patients will respond to CRC treatment.

We then investigated the role POLD3 may play in the context of CRC. POLD3 is a subunit of the polymerase Delta complex, which is responsible for lagging-strand DNA synthesis, as well as homologous repair of DNA breaks. POLD3 facilitates the binding of the active polymerase enzyme to the PCNA clamp, which increases DNA processivity, and may also help to stabilise the polymerase delta complex [177][166]. Interestingly, we found that POLD3 loss resulted in reduced cellular proliferation, and stalling of the cell cycle progression, which has also been reported previously [177]. This would signify a beneficial role in POLD3 knockdown, as loss of POLD3 could aid in the reduction in cancer-cell proliferation. Furthermore, POLD3 knockdown sensitized cells to chemo and radiotherapy, again, a potential positive benefit to tackling cancer cell treatment. However, we also saw an emergence of DNA damage upon POLD3 knockdown, which may lead to DNA instability, enhanced mutational load, and therefore an increase in oncogenic potential. It is well known that CRC is a hetero-genetic hypermutated cancer, with a number of pathways of genetic and chromosomal instability often associated with CRC. Therefore, POLD3 loss may result in an increase in DNA instability through DNA damage and worsen the possibility of cancer cell formation. We have also shown that POLD3 knockdown reduces telomere length through the ALT mechanism. Despite cancer cells relying on the ability to lengthen telomeres to sustain longevity, the shortening of telomeres may also present detrimental effects, with the possibility of coding DNA being lost at telomere ends, further contributing to DNA and chromosomal instability. It is therefore unclear if POLD3 expression is positive or detrimental to cancer cell formation.

While we have shown that *POLD3* is highly expressed in CRC, this may be due to selection bias for cells with high levels of DNA replication machinery in order to compete with replication stress posed by a highly proliferative state. Interestingly, *POLD3* high expression was associated with better patient outcomes, which may be explained by *POLD3* loss resulting in high DNA damage. It has recently been proposed that *POLD3* depletion results in a high frequency of genome duplications, which increases chromosomal instability [276]. Furthermore, BIR, in which *POLD3* is essential, is a physiological process which accounts for chromosomal translocation, extensive loss of heterozygosity or telomere elongation which are common features of cancer cells [311,312]. Additionally, *POLD3* has been suggested as driving DNA repair synthesis following replication stress during mitosis [313]. It is well known that perturbation of DNA replication is often linked with genome instability, which is associated with tumorigenesis.

Further work would be needed on the analysis of *POLD3* risk SNPs, to understand if these genetic abnormalities lead to the upregulation or dysfunction of *POLD3*, which may give more understanding in its role as a risk gene. However, it appears germline mutations or risk SNPs in the *POLD3* gene may cause *POLD3* misfunction, resulting in the genetic instability we have shown here [184][150][314]. The most probable outcome for *POLD3* risk SNPs in CRC is therefore dysfunction or under-expression of the *POLD3* protein, resulting in high genetic instability, driving tumour progression. Loss of *POLD3* expression may therefore be a potential biomarker for increased genomic instability in CRC. However, targeting of *POLD3* may also serve as a therapeutic target to reduce cellular proliferation and enhance chemo sensitization. Previous reports have shown the potential use of immune checkpoint inhibitor immunotherapies in the targeting of cancers with *POLE/POLD1* proofreading mutations, which exhibit high genetic instability [13]. Therefore, the use of immune checkpoint inhibitors may prove some use in the treating of *POLD3* deficient cancers.

This project aimed to look at the effects of potential biomarkers *POLD3* and *CHRD2* in CRC. GWAS data found that the region containing *POLD3* and *CHRD2* contained a high number of SNPs which may increase the risk of CRC development. Therefore, we wished to look at these genes together in order to determine if there was an apparent relationship between the two. However, we found no correlation between *POLD3* and *CHRD2* expression, suggesting that SNPs in the genomic region would not contribute to joint expression patterns. One theory suggests that this area may be more accessible to mutagens, causing an increase in genetic alterations and the emergence of the significant risk SNPs. However preliminary analysis using CBio portal found no copy number changes or mutational hotspots in either gene. Therefore, it appears these genes and their potential

affects in the development of CRC are unrelated, and it is probably by chance they lie together on chromosome 11.

Conclusion

In this thesis we have shown the important roles that *CHRD2* overexpression and *POLD3* knockdown may play in the development of CRC. Overexpression of *CHRD2* leads to an increase in the cancer-stem-cell phenotype, leading to increased chemotherapy and radiation resistance through upregulation of DNA repair pathways. *POLD3* knockdown has been shown to increase genomic instability, through manifestation of DNA breaks and shortened telomeres. However, *POLD3* loss also results in lower proliferative ability and apoptosis, potentially causing both anti and tumorigenic effects in CRC. These two genes, although apparently un-related, are potential novel biomarkers in the risk of CRC development and offer up potential targets in CRC treatment.

References

- 1 Bray F, Ferlay J, Soerjomataram I, *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424. doi: 10.3322/caac.21492
- 2 Nagtegaal ID, Odze RD, Klimstra D, *et al.* The 2019 WHO classification of tumours of the digestive system. *Histopathology.* 2020;76:182–8. doi: 10.1111/his.13975
- 3 Center MM, Jemal A, Smith RA, *et al.* Worldwide Variations in Colorectal Cancer. *CA Cancer J Clin.* 2009;59:366–78. doi: 10.3322/caac.20038
- 4 Hampel H, Kalady MF, Pearlman R, *et al.* Hereditary Colorectal Cancer. *Hematol Oncol Clin North Am.* 2022;36:429–47. doi: 10.1016/j.hoc.2022.02.002
- 5 Henrikson NB, Webber EM, Goddard KA, *et al.* Family history and the natural history of colorectal cancer: systematic review. *Genetics in Medicine.* 2015;17:702–12. doi: 10.1038/gim.2014.188
- 6 Clemmensen SB, Harris JR, Mengel-From J, *et al.* Familial Risk and Heritability of Hematologic Malignancies in the Nordic Twin Study of Cancer. *Cancers (Basel).* 2021;13:3023. doi: 10.3390/cancers13123023
- 7 Glebov OK, Rodriguez LM, Nakahara K, *et al.* Distinguishing right from left colon by the pattern of gene expression. *Cancer Epidemiol Biomarkers Prev.* 2003;12:755–62.
- 8 Marzouk O, Schofield J. Review of Histopathological and Molecular Prognostic Features in Colorectal Cancer. *Cancers (Basel).* 2011;3:2767–810. doi: 10.3390/cancers3022767
- 9 Gualco G, Reissenweber N, Cliché I, *et al.* Flat elevated lesions of the colon and rectum: a spectrum of neoplastic and nonneoplastic entities. *Ann Diagn Pathol.* 2006;10:333–8. doi: 10.1016/j.anndiagnpath.2006.03.003
- 10 Dekker E, Tanis PJ, Vleugels JLA, *et al.* Colorectal cancer. *The Lancet.* 2019;394:1467–80. doi: 10.1016/S0140-6736(19)32319-0
- 11 Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer. *JAMA.* 2021;325:669. doi: 10.1001/jama.2021.0106
- 12 Cunningham D, Atkin W, Lenz H-J, *et al.* Colorectal cancer. *The Lancet.* 2010;375:1030–47. doi: 10.1016/S0140-6736(10)60353-4
- 13 Ambrosini M, Rousseau B, Manca P, *et al.* Immune checkpoint inhibitors for POLE or POLD1 proofreading-deficient metastatic colorectal cancer. *Annals of Oncology.* 2024;35:643–55. doi: 10.1016/j.annonc.2024.03.009
- 14 Boland P, Ma W. Immunotherapy for Colorectal Cancer. *Cancers (Basel).* 2017;9:50. doi: 10.3390/cancers9050050
- 15 Edin S, Gylling B, Li X, *et al.* Opposing roles by KRAS and BRAF mutation on immune cell infiltration in colorectal cancer – possible implications for immunotherapy. *Br J Cancer.* 2024;130:143–50. doi: 10.1038/s41416-023-02483-9

- 16 Tria SM, Burge ME, Whitehall VLJ. The Therapeutic Landscape for KRAS-Mutated Colorectal Cancers. *Cancers (Basel)*. 2023;15:2375. doi: 10.3390/cancers15082375
- 17 Riihimäki M, Hemminki A, Sundquist J, *et al*. Patterns of metastasis in colon and rectal cancer. *Sci Rep*. 2016;6:29765. doi: 10.1038/srep29765
- 18 Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer. *JAMA*. 2021;325:669. doi: 10.1001/jama.2021.0106
- 19 Bakir B, Chiarella AM, Pitarresi JR, *et al*. EMT, MET, Plasticity, and Tumor Metastasis. *Trends Cell Biol*. 2020;30:764–76. doi: 10.1016/j.tcb.2020.07.003
- 20 Aiello NM, Maddipati R, Norgard RJ, *et al*. EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Dev Cell*. 2018;45:681–695.e4. doi: 10.1016/j.devcel.2018.05.027
- 21 Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation*. 2009;119:1420–8. doi: 10.1172/JCI39104
- 22 Tse JC, Kalluri R. Mechanisms of metastasis: Epithelial-to-mesenchymal transition and contribution of tumor microenvironment. *J Cell Biochem*. 2007;101:816–29. doi: 10.1002/jcb.21215
- 23 Neophytou CM, Panagi M, Stylianopoulos T, *et al*. The Role of Tumor Microenvironment in Cancer Metastasis: Molecular Mechanisms and Therapeutic Opportunities. *Cancers (Basel)*. 2021;13:2053. doi: 10.3390/cancers13092053
- 24 Gagnière J. Gut microbiota imbalance and colorectal cancer. *World J Gastroenterol*. 2016;22:501. doi: 10.3748/wjg.v22.i2.501
- 25 Huycke MM, Abrams V, Moore DR. Enterococcus faecalis produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. *Carcinogenesis*. 2002;23:529–36. doi: 10.1093/carcin/23.3.529
- 26 Aymeric L, Donnadiou F, Mulet C, *et al*. Colorectal cancer specific conditions promote *Streptococcus gallolyticus* gut colonization. *Proceedings of the National Academy of Sciences*. 2018;115. doi: 10.1073/pnas.1715112115
- 27 Wang S, Sun J, Chen K, *et al*. Perspectives of tumor-infiltrating lymphocyte treatment in solid tumors. *BMC Med*. 2021;19:140. doi: 10.1186/s12916-021-02006-4
- 28 Cheng Y, Zhu Y, Xu W, *et al*. PKC α in colon cancer cells promotes M1 macrophage polarization via MKK3/6-P38 MAPK pathway. *Mol Carcinog*. 2018;57:1017–29. doi: 10.1002/mc.22822
- 29 Goswami KK, Ghosh T, Ghosh S, *et al*. Tumor promoting role of anti-tumor macrophages in tumor microenvironment. *Cell Immunol*. 2017;316:1–10. doi: 10.1016/j.cellimm.2017.04.005
- 30 Triner D, Devenport SN, Ramakrishnan SK, *et al*. Neutrophils Restrict Tumor-Associated Microbiota to Reduce Growth and Invasion of Colon Tumors in Mice. *Gastroenterology*. 2019;156:1467–82. doi: 10.1053/j.gastro.2018.12.003
- 31 Hirai H, Fujishita T, Kurimoto K, *et al*. CCR1-mediated accumulation of myeloid cells in the liver microenvironment promoting mouse colon cancer metastasis. *Clin Exp Metastasis*. 2014;31:977–89. doi: 10.1007/s10585-014-9684-z

- 32 Li J, Chen D, Shen M. Tumor Microenvironment Shapes Colorectal Cancer Progression, Metastasis, and Treatment Responses. *Front Med (Lausanne)*. 2022;9. doi: 10.3389/fmed.2022.869010
- 33 Zhang Y, Wang S, Lai Q, *et al.* Cancer-associated fibroblasts-derived exosomal miR-17-5p promotes colorectal cancer aggressive phenotype by initiating a RUNX3/MYC/TGF- β 1 positive feedback loop. *Cancer Lett*. 2020;491:22–35. doi: 10.1016/j.canlet.2020.07.023
- 34 Zhang R, Qi F, Zhao F, *et al.* Cancer-associated fibroblasts enhance tumor-associated macrophages enrichment and suppress NK cells function in colorectal cancer. *Cell Death Dis*. 2019;10:273. doi: 10.1038/s41419-019-1435-2
- 35 Nagasaki T, Hara M, Nakanishi H, *et al.* Interleukin-6 released by colon cancer-associated fibroblasts is critical for tumour angiogenesis: anti-interleukin-6 receptor antibody suppressed angiogenesis and inhibited tumour–stroma interaction. *Br J Cancer*. 2014;110:469–78. doi: 10.1038/bjc.2013.748
- 36 Nübel T, Dippold W, Kleinert H, *et al.* Lovastatin inhibits Rho-regulated expression of E-selectin by TNF- α and attenuates tumor cell adhesion. *The FASEB Journal*. 2004;18:140–2. doi: 10.1096/fj.03-0261fje
- 37 Motz GT, Santoro SP, Wang L-P, *et al.* Tumor endothelium FasL establishes a selective immune barrier promoting tolerance in tumors. *Nat Med*. 2014;20:607–15. doi: 10.1038/nm.3541
- 38 Wang J, Uddin MdN, Akter R, *et al.* Contribution of endothelial cell-derived transcriptomes to the colon cancer based on bioinformatics analysis. *Mathematical Biosciences and Engineering*. 2021;18:7280–300. doi: 10.3934/mbe.2021360
- 39 Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61:759–67. doi: 10.1016/0092-8674(90)90186-l
- 40 Pino MS, Chung DC. The Chromosomal Instability Pathway in Colon Cancer. *Gastroenterology*. 2010;138:2059–72. doi: 10.1053/j.gastro.2009.12.065
- 41 Wong K, Xie G. Updates on the Molecular Genetics of Colorectal Cancer. *Colorectal Cancer: Open Access*. 2017;03. doi: 10.21767/2471-9943.100032
- 42 Leary RJ, Lin JC, Cummins J, *et al.* Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proceedings of the National Academy of Sciences*. 2008;105:16224–9. doi: 10.1073/pnas.0808041105
- 43 Law PJ, Timofeeva M, Fernandez-Rozadilla C, *et al.* Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun*. 2019;10:2154. doi: 10.1038/s41467-019-09775-w
- 44 Guinney J, Dienstmann R, Wang X, *et al.* The consensus molecular subtypes of colorectal cancer. *Nat Med*. 2015;21:1350–6. doi: 10.1038/nm.3967
- 45 Dienstmann R, Vermeulen L, Guinney J, *et al.* Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat Rev Cancer*. 2017;17:79–92. doi: 10.1038/nrc.2016.126
- 46 Bauer KM, Hummon AB, Buechler S. Right-side and left-side colon cancer follow different pathways to relapse. *Mol Carcinog*. 2012;51:411–21. doi: 10.1002/mc.20804

- 47 Colussi D, Brandi G, Bazzoli F, *et al.* Molecular Pathways Involved in Colorectal Cancer: Implications for Disease Behavior and Prevention. *Int J Mol Sci.* 2013;14:16365–85. doi: 10.3390/ijms140816365
- 48 Hsieh P. Molecular mechanisms of DNA mismatch repair. *Mutation Research/DNA Repair.* 2001;486:71–87. doi: 10.1016/S0921-8777(01)00088-X
- 49 Yoon YS, Yu CS, Kim TW, *et al.* Mismatch repair status in sporadic colorectal cancer: Immunohistochemistry and microsatellite instability analyses. *J Gastroenterol Hepatol.* 2011;26:1733–9. doi: 10.1111/j.1440-1746.2011.06784.x
- 50 Lanza G, Gafà R, Maestri I, *et al.* Immunohistochemical Pattern of MLH1/MSH2 Expression Is Related to Clinical and Pathological Features in Colorectal Adenocarcinomas with Microsatellite Instability. *Modern Pathology.* 2002;15:741–9. doi: 10.1097/01.MP.0000018979.68686.B2
- 51 Herman JG, Umar A, Polyak K, *et al.* Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *Proceedings of the National Academy of Sciences.* 1998;95:6870–5. doi: 10.1073/pnas.95.12.6870
- 52 Esteller M, Levine R, Baylin SB, *et al.* MLH1 promoter hypermethylation is associated with the microsatellite instability phenotype in sporadic endometrial carcinomas. *Oncogene.* 1998;17:2413–7. doi: 10.1038/sj.onc.1202178
- 53 Boland CR, Goel A. Microsatellite Instability in Colorectal Cancer. *Gastroenterology.* 2010;138:2073-2087.e3. doi: 10.1053/j.gastro.2009.12.064
- 54 Veigl ML, Kasturi L, Olechnowicz J, *et al.* Biallelic inactivation of *hMLH1* by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proceedings of the National Academy of Sciences.* 1998;95:8698–702. doi: 10.1073/pnas.95.15.8698
- 55 Ogino S. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. *Gut.* 2006;55:1000–6. doi: 10.1136/gut.2005.082933
- 56 Jones PA, Laird PW. Cancer-epigenetics comes of age. *Nat Genet.* 1999;21:163–7. doi: 10.1038/5947
- 57 Grady WM, Carethers JM. Genomic and Epigenetic Instability in Colorectal Cancer Pathogenesis. *Gastroenterology.* 2008;135:1079–99. doi: 10.1053/j.gastro.2008.07.076
- 58 Herman JG, Umar A, Polyak K, *et al.* Incidence and functional consequences of *hMLH1* promoter hypermethylation in colorectal carcinoma. *Proceedings of the National Academy of Sciences.* 1998;95:6870–5. doi: 10.1073/pnas.95.12.6870
- 59 Ogino S, Nosho K, Kirkner GJ, *et al.* CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut.* 2009;58:90–6. doi: 10.1136/gut.2008.155473
- 60 Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487:330–7. doi: 10.1038/nature11252
- 61 Schatoff EM, Leach BI, Dow LE. WNT Signaling and Colorectal Cancer. *Curr Colorectal Cancer Rep.* 2017;13:101–10. doi: 10.1007/s11888-017-0354-9

- 62 Barker N. Adult intestinal stem cells: critical drivers of epithelial homeostasis and regeneration. *Nat Rev Mol Cell Biol.* 2014;15:19–33. doi: 10.1038/nrm3721
- 63 Sieber OM, Lamlum H, Crabtree MD, *et al.* Whole-gene APC deletions cause classical familial adenomatous polyposis, but not attenuated polyposis or “multiple” colorectal adenomas. *Proceedings of the National Academy of Sciences.* 2002;99:2954–8. doi: 10.1073/pnas.042699199
- 64 Sparks AB, Morin PJ, Vogelstein B, *et al.* Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. *Cancer Res.* 1998;58:1130–4.
- 65 Fodde R, Kuipers J, Rosenberg C, *et al.* Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nat Cell Biol.* 2001;3:433–8. doi: 10.1038/35070129
- 66 Firestein R, Shima K, Nosho K, *et al.* CDK8 expression in 470 colorectal cancers in relation to β -catenin activation, other molecular alterations and patient survival. *Int J Cancer.* 2010;126:2863–73. doi: 10.1002/ijc.24908
- 67 Horst D, Chen J, Morikawa T, *et al.* Differential WNT Activity in Colorectal Cancer Confers Limited Tumorigenic Potential and Is Regulated by MAPK Signaling. *Cancer Res.* 2012;72:1547–56. doi: 10.1158/0008-5472.CAN-11-3222
- 68 Zhang L, Shay JW. Multiple Roles of APC and its Therapeutic Implications in Colorectal Cancer. *JNCI: Journal of the National Cancer Institute.* 2017;109. doi: 10.1093/jnci/djw332
- 69 Kleeman SO, Leedham SJ. Not All Wnt Activation Is Equal: Ligand-Dependent versus Ligand-Independent Wnt Activation in Colorectal Cancer. *Cancers (Basel).* 2020;12:3355. doi: 10.3390/cancers12113355
- 70 Maruvka YE, Mouw KW, Karlic R, *et al.* Analysis of somatic microsatellite indels identifies driver events in human tumors. *Nat Biotechnol.* 2017;35:951–9. doi: 10.1038/nbt.3966
- 71 Giannakis M, Hodis E, Jasmine Mu X, *et al.* RNF43 is frequently mutated in colorectal and endometrial cancers. *Nat Genet.* 2014;46:1264–6. doi: 10.1038/ng.3127
- 72 Dinu D, Dobre M, Panaitescu E, *et al.* Prognostic significance of KRAS gene mutations in colorectal cancer--preliminary study. *J Med Life.* 2014;7:581–7.
- 73 Nose H, Imazeki F, Ohto M, *et al.* p53 gene mutations and 17p allelic deletions in hepatocellular Carcinoma from Japan. *Cancer.* 1993;72:355–60. doi: 10.1002/1097-0142(19930715)72:2<355::AID-CNCR2820720208>3.0.CO;2-W
- 74 El-Deiry WS, Tokino T, Velculescu VE, *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell.* 1993;75:817–25. doi: 10.1016/0092-8674(93)90500-P
- 75 Jones RG, Plas DR, Kubek S, *et al.* AMP-Activated Protein Kinase Induces a p53-Dependent Metabolic Checkpoint. *Mol Cell.* 2005;18:283–93. doi: 10.1016/j.molcel.2005.03.027
- 76 Swamy M V, Herzog CR, Rao C V. Inhibition of COX-2 in colon cancer cell lines by celecoxib increases the nuclear localization of active p53. *Cancer Res.* 2003;63:5239–42.
- 77 Samuels Y, Waldman T. Oncogenic Mutations of PIK3CA in Human Cancers. 2010:21–41.

- 78 Gabay M, Li Y, Felsner DW. MYC Activation Is a Hallmark of Cancer Initiation and Maintenance. *Cold Spring Harb Perspect Med*. 2014;4:a014241–a014241. doi: 10.1101/cshperspect.a014241
- 79 Hao Y-H, Lafita-Navarro MC, Zacharias L, *et al*. Induction of LEF1 by MYC activates the WNT pathway and maintains cell proliferation. *Cell Communication and Signaling*. 2019;17:129. doi: 10.1186/s12964-019-0444-1
- 80 Rennoll S. Regulation of MYC gene expression by aberrant Wnt/ β -catenin signaling in colorectal cancer. *World J Biol Chem*. 2015;6:290. doi: 10.4331/wjbc.v6.i4.290
- 81 Azkanaz M, Corominas-Murtra B, Ellenbroek SIJ, *et al*. Retrograde movements determine effective stem cell numbers in the intestine. *Nature*. 2022;607:548–54. doi: 10.1038/s41586-022-04962-0
- 82 Barker N, van Es JH, Kuipers J, *et al*. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*. 2007;449:1003–7. doi: 10.1038/nature06196
- 83 Lopez-Garcia C, Klein AM, Simons BD, *et al*. Intestinal Stem Cell Replacement Follows a Pattern of Neutral Drift. *Science (1979)*. 2010;330:822–5. doi: 10.1126/science.1196236
- 84 Cheng X, Xu X, Chen D, *et al*. Therapeutic potential of targeting the Wnt/ β -catenin signaling pathway in colorectal cancer. *Biomedicine & Pharmacotherapy*. 2019;110:473–81. doi: 10.1016/j.biopha.2018.11.082
- 85 Ong BA. Intestinal stem cells and the colorectal cancer microenvironment. *World J Gastroenterol*. 2014;20:1898. doi: 10.3748/wjg.v20.i8.1898
- 86 Zhong Z, Virshup DM. Wnt Signaling and Drug Resistance in Cancer. *Mol Pharmacol*. 2020;97:72–89. doi: 10.1124/mol.119.117978
- 87 Wiese KE, Nusse R, van Amerongen R. Wnt signalling: conquering complexity. *Development*. 2018;145. doi: 10.1242/dev.165902
- 88 Oguma K, Oshima H, Aoki M, *et al*. Activated macrophages promote Wnt signalling through tumour necrosis factor- α in gastric tumour cells. *EMBO J*. 2008;27:1671–81. doi: 10.1038/emboj.2008.105
- 89 GENG Y, LU X, WU X, *et al*. MicroRNA-27b suppresses Helicobacter pylori-induced gastric tumorigenesis through negatively regulating Frizzled7. *Oncol Rep*. 2016;35:2441–50. doi: 10.3892/or.2016.4572
- 90 Flanagan DJ, Barker N, Nowell C, *et al*. Loss of the Wnt receptor Frizzled7 in the gastric epithelium is deleterious and triggers rapid repopulation in vivo. *Dis Model Mech*. Published Online First: 1 January 2017. doi: 10.1242/dmm.029876
- 91 Albuquerque C. The ‘just-right’ signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. *Hum Mol Genet*. 2002;11:1549–60. doi: 10.1093/hmg/11.13.1549
- 92 Voloshanenko O, Erdmann G, Dubash TD, *et al*. Wnt secretion is required to maintain high levels of Wnt activity in colon cancer cells. *Nat Commun*. 2013;4:2610. doi: 10.1038/ncomms3610

- 93 Vincan E, Darcy PK, Farrelly CA, *et al.* Frizzled-7 dictates three-dimensional organization of colorectal cancer cell carcinoids. *Oncogene*. 2007;26:2340–52. doi: 10.1038/sj.onc.1210026
- 94 Voorneveld PW, Kodach LL, Jacobs RJ, *et al.* The BMP pathway either enhances or inhibits the Wnt pathway depending on the SMAD4 and p53 status in CRC. *Br J Cancer*. 2015;112:122–30. doi: 10.1038/bjc.2014.560
- 95 He XC, Zhang J, Tong W-G, *et al.* BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt– β -catenin signaling. *Nat Genet*. 2004;36:1117–21. doi: 10.1038/ng1430
- 96 Ouahoud S, Hardwick JCH, Hawinkels LJAC. Extracellular BMP Antagonists, Multifaceted Orchestrators in the Tumor and Its Microenvironment. *Int J Mol Sci*. 2020;21:3888. doi: 10.3390/ijms21113888
- 97 Beumer J, Puschhof J, Yengej FY, *et al.* BMP gradient along the intestinal villus axis controls zonated enterocyte and goblet cell states. *Cell Rep*. 2022;38:110438. doi: 10.1016/j.celrep.2022.110438
- 98 Wozney JM, Rosen V, Celeste AJ, *et al.* Novel Regulators of Bone Formation: Molecular Clones and Activities. *Science (1979)*. 1988;242:1528–34. doi: 10.1126/science.3201241
- 99 Peng J, Yoshioka Y, Mandai M, *et al.* The BMP signaling pathway leads to enhanced proliferation in serous ovarian cancer-A potential therapeutic target. *Mol Carcinog*. 2016;55:335–45. doi: 10.1002/mc.22283
- 100 Fukuda T, Fukuda R, Tanabe R, *et al.* BMP signaling is a therapeutic target in ovarian cancer. *Cell Death Discov*. 2020;6:139. doi: 10.1038/s41420-020-00377-w
- 101 Wang L, Park P, Zhang H, *et al.* BMP-2 inhibits the tumorigenicity of cancer stem cells in human osteosarcoma OS99-1 cell line. *Cancer Biol Ther*. 2011;11:457–63. doi: 10.4161/cbt.11.5.14372
- 102 Ye L, Kynaston H, Jiang WG. Bone Morphogenetic Protein-9 Induces Apoptosis in Prostate Cancer Cells, the Role of Prostate Apoptosis Response-4. *Molecular Cancer Research*. 2008;6:1594–606. doi: 10.1158/1541-7786.MCR-08-0171
- 103 Hardwick JC, Kodach LL, Offerhaus GJ, *et al.* Bone morphogenetic protein signalling in colorectal cancer. *Nat Rev Cancer*. 2008;8:806–12. doi: 10.1038/nrc2467
- 104 Massagué J. TGF- β SIGNAL TRANSDUCTION. *Annu Rev Biochem*. 1998;67:753–91. doi: 10.1146/annurev.biochem.67.1.753
- 105 Kosinski C, Li VSW, Chan ASY, *et al.* Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. *Proceedings of the National Academy of Sciences*. 2007;104:15418–23. doi: 10.1073/pnas.0707210104
- 106 Ma Y, Yan F, Li L, *et al.* Deletion and down-regulation of SMAD4 gene in colorectal cancers in a Chinese population. *Chin J Cancer Res*. 2014;26:525–31. doi: 10.3978/j.issn.1000-9604.2014.09.02
- 107 Papadopulos ME, Plazzer JP, Macrae FA. Genotype–phenotype correlation of BMPR1a disease causing variants in juvenile polyposis syndrome. *Hered Cancer Clin Pract*. 2023;21:12. doi: 10.1186/s13053-023-00255-3

- 108 Luo M, Brooks M, Wicha M. Epithelial-Mesenchymal Plasticity of Breast Cancer Stem Cells: Implications for Metastasis and Therapeutic Resistance. *Curr Pharm Des.* 2015;21:1301–10. doi: 10.2174/1381612821666141211120604
- 109 Zhou H-M, Zhang J-G, Zhang X, *et al.* Targeting cancer stem cells for reversing therapy resistance: mechanism, signaling, and prospective agents. *Signal Transduct Target Ther.* 2021;6:62. doi: 10.1038/s41392-020-00430-1
- 110 Lombardo Y, Scopelliti A, Cammareri P, *et al.* Bone Morphogenetic Protein 4 Induces Differentiation of Colorectal Cancer Stem Cells and Increases Their Response to Chemotherapy in Mice. *Gastroenterology.* 2011;140:297-309.e6. doi: 10.1053/j.gastro.2010.10.005
- 111 Piccirillo SGM, Reynolds BA, Zanetti N, *et al.* Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature.* 2006;444:761–5. doi: 10.1038/nature05349
- 112 Whissell G, Montagni E, Martinelli P, *et al.* The transcription factor GATA6 enables self-renewal of colon adenoma stem cells by repressing BMP gene expression. *Nat Cell Biol.* 2014;16:695–707. doi: 10.1038/ncb2992
- 113 Walsh DW, Godson C, Brazil DP, *et al.* Extracellular BMP-antagonist regulation in development and disease: tied up in knots. *Trends Cell Biol.* 2010;20:244–56. doi: 10.1016/j.tcb.2010.01.008
- 114 Berglar I, Hehlhans S, Wehle A, *et al.* CHRDL1 Regulates Stemness in Glioma Stem-like Cells. *Cells.* 2022;11:3917. doi: 10.3390/cells11233917
- 115 Kobayashi H, Gieniec KA, Wright JA, *et al.* The Balance of Stromal BMP Signaling Mediated by GREM1 and ISLR Drives Colorectal Carcinogenesis. *Gastroenterology.* 2021;160:1224-1239.e30. doi: 10.1053/j.gastro.2020.11.011
- 116 Ouahoud S, Hardwick JCH, Hawinkels LJAC. Extracellular BMP Antagonists, Multifaceted Orchestrators in the Tumor and Its Microenvironment. *Int J Mol Sci.* 2020;21:3888. doi: 10.3390/ijms21113888
- 117 Sharov AA, Mardaryev AN, Sharova TY, *et al.* Bone Morphogenetic Protein Antagonist Noggin Promotes Skin Tumorigenesis via Stimulation of the Wnt and Shh Signaling Pathways. *Am J Pathol.* 2009;175:1303–14. doi: 10.2353/ajpath.2009.090163
- 118 Valenzuela D, Economides A, Rojas E, *et al.* Identification of mammalian noggin and its expression in the adult nervous system. *The Journal of Neuroscience.* 1995;15:6077–84. doi: 10.1523/JNEUROSCI.15-09-06077.1995
- 119 Krause C, Guzman A, Knaus P. Noggin. *Int J Biochem Cell Biol.* 2011;43:478–81. doi: 10.1016/j.biocel.2011.01.007
- 120 van Hattem WA, Langeveld D, de Leng WWJ, *et al.* Histologic Variations in Juvenile Polyp Phenotype Correlate With Genetic Defect Underlying Juvenile Polyposis. *American Journal of Surgical Pathology.* 2011;35:530–6. doi: 10.1097/PAS.0b013e318211cae1
- 121 Sharov AA, Mardaryev AN, Sharova TY, *et al.* Bone Morphogenetic Protein Antagonist Noggin Promotes Skin Tumorigenesis via Stimulation of the Wnt and Shh Signaling Pathways. *Am J Pathol.* 2009;175:1303–14. doi: 10.2353/ajpath.2009.090163

- 122 Sun Z, Gao X, Zabkiewicz C, *et al.* Noggin is associated with a poor prognosis of gastric cancer by promoting the proliferation of gastric cancer cells via the upregulation of EGFR. *Int J Oncol.* 2020;57:813–24. doi: 10.3892/ijo.2020.5081
- 123 Cai C, Itzel T, Gaitantzi H, *et al.* Identification of liver-derived bone morphogenetic protein (BMP)-9 as a potential new candidate for treatment of colorectal cancer. *J Cell Mol Med.* 2022;26:343–53. doi: 10.1111/jcmm.17084
- 124 Haramis A-PG, Begthel H, van den Born M, *et al.* De Novo Crypt Formation and Juvenile Polyposis on BMP Inhibition in Mouse Intestine. *Science (1979).* 2004;303:1684–6. doi: 10.1126/science.1093587
- 125 Gomez-Puerto MC, Iyengar PV, García de Vinuesa A, *et al.* Bone morphogenetic protein receptor signal transduction in human disease. *J Pathol.* 2019;247:9–20. doi: 10.1002/path.5170
- 126 Gao Z, Houthuijzen JM, ten Dijke P, *et al.* GREM1 signaling in cancer: tumor promotor and suppressor? *J Cell Commun Signal.* 2023;17:1517–26. doi: 10.1007/s12079-023-00777-4
- 127 Dolan V, Murphy M, Sadlier D, *et al.* Expression of Gremlin, a Bone Morphogenetic Protein Antagonist, in Human Diabetic Nephropathy. *American Journal of Kidney Diseases.* 2005;45:1034–9. doi: 10.1053/j.ajkd.2005.03.014
- 128 Zúñiga A, Haramis A-PG, McMahon AP, *et al.* Signal relay by BMP antagonism controls the SHH/FGF4 feedback loop in vertebrate limb buds. *Nature.* 1999;401:598–602. doi: 10.1038/44157
- 129 Gao Z, Houthuijzen JM, ten Dijke P, *et al.* GREM1 signaling in cancer: tumor promotor and suppressor? *J Cell Commun Signal.* 2023;17:1517–26. doi: 10.1007/s12079-023-00777-4
- 130 Kobayashi H, Gieniec KA, Wright JA, *et al.* The Balance of Stromal BMP Signaling Mediated by GREM1 and ISLR Drives Colorectal Carcinogenesis. *Gastroenterology.* 2021;160:1224-1239.e30. doi: 10.1053/j.gastro.2020.11.011
- 131 Koppens MAJ, Davis H, Valbuena GN, *et al.* Bone Morphogenetic Protein Pathway Antagonism by Grem1 Regulates Epithelial Cell Fate in Intestinal Regeneration. *Gastroenterology.* 2021;161:239-254.e9. doi: 10.1053/j.gastro.2021.03.052
- 132 Li R, Zhou H, Li M, *et al.* Gremlin-1 Promotes Colorectal Cancer Cell Metastasis by Activating ATF6 and Inhibiting ATF4 Pathways. *Cells.* 2022;11:2136. doi: 10.3390/cells11142136
- 133 Urra H, Dufey E, Avril T, *et al.* Endoplasmic Reticulum Stress and the Hallmarks of Cancer. *Trends Cancer.* 2016;2:252–62. doi: 10.1016/j.trecan.2016.03.007
- 134 Lee B-R, Chang S-Y, Hong E-H, *et al.* Elevated endoplasmic reticulum stress reinforced immunosuppression in the tumor microenvironment *via* myeloid-derived suppressor cells. *Oncotarget.* 2014;5:12331–45. doi: 10.18632/oncotarget.2589
- 135 Huang J, Pan H, Wang J, *et al.* Unfolded protein response in colorectal cancer. *Cell Biosci.* 2021;11:26. doi: 10.1186/s13578-021-00538-z
- 136 Jaeger E, Leedham S, Lewis A, *et al.* Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1. *Nat Genet.* 2012;44:699–703. doi: 10.1038/ng.2263

- 137 Rohlin A, Eiengård F, Lundstam U, *et al.* *GREM 1* and *POLE* variants in hereditary colorectal cancer syndromes. *Genes Chromosomes Cancer*. 2016;55:95–106. doi: 10.1002/gcc.22314
- 138 Davis H, Irshad S, Bansal M, *et al.* Aberrant epithelial *GREM1* expression initiates colonic tumorigenesis from cells outside the stem cell niche. *Nat Med*. 2015;21:62–70. doi: 10.1038/nm.3750
- 139 Wu I, Moses MA. *BNF-1*, a novel gene encoding a putative extracellular matrix protein, is overexpressed in tumor tissues. *Gene*. 2003;311:105–10. doi: 10.1016/S0378-1119(03)00563-8
- 140 Oren A, Toporik A, Biton S, *et al.* *hCHL2*, a novel chordin-related gene, displays differential expression and complex alternative splicing in human tissues and during myoblast and osteoblast maturation. *Gene*. 2004;331:17–31. doi: 10.1016/j.gene.2004.01.029
- 141 Li D, Xie X-Y, Shen H, *et al.* Chordin-like 2 influences the differentiation fate of retinal pigment epithelium cells by dynamically regulating BMP pathway. *Int J Ophthalmol*. 2022;15:711–20. doi: 10.18240/ijo.2022.05.04
- 142 Oren A, Toporik A, Biton S, *et al.* *hCHL2*, a novel chordin-related gene, displays differential expression and complex alternative splicing in human tissues and during myoblast and osteoblast maturation. *Gene*. 2004;331:17–31. doi: 10.1016/j.gene.2004.01.029
- 143 Fujisawa T, Huang Y, Sebald W, *et al.* The binding of von Willebrand factor type C domains of Chordin family proteins to BMP-2 and Tsg is mediated by their SD1 subdomain. *Biochem Biophys Res Commun*. 2009;385:215–9. doi: 10.1016/j.bbrc.2009.05.041
- 144 Sun J, Liu X, Gao H, *et al.* Overexpression of colorectal cancer oncogene *CHRD2* predicts a poor prognosis. *Oncotarget*. 2017;8:11489–506. doi: 10.18632/oncotarget.14039
- 145 Sun J, Zhao J, Jiang F, *et al.* Identification of novel protein biomarkers and drug targets for colorectal cancer by integrating human plasma proteome with genome. *Genome Med*. 2023;15:75. doi: 10.1186/s13073-023-01229-9
- 146 Chen H, Pan R, Li H, *et al.* *CHRD2* promotes osteosarcoma cell proliferation and metastasis through the BMP-9/PI3K/AKT pathway. *Cell Biol Int*. 2021;45:623–32. doi: 10.1002/cbin.11507
- 147 Wang L, Xu W, Mei Y, *et al.* *CHRD2* promotes cell proliferation by activating the YAP/TAZ signaling pathway in gastric cancer. *Free Radic Biol Med*. 2022;193:158–70. doi: 10.1016/j.freeradbiomed.2022.09.006
- 148 Law PJ, Timofeeva M, Fernandez-Rozadilla C, *et al.* Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun*. 2019;10:2154. doi: 10.1038/s41467-019-09775-w
- 149 Garcia-Diaz M, Bebenek K. Multiple Functions of DNA Polymerases. *CRC Crit Rev Plant Sci*. 2007;26:105–22. doi: 10.1080/07352680701252817
- 150 Rayner E, van Gool IC, Palles C, *et al.* A panoply of errors: polymerase proofreading domain mutations in cancer. *Nat Rev Cancer*. 2016;16:71–81. doi: 10.1038/nrc.2015.12
- 151 Poynter JN, Siegmund KD, Weisenberger DJ, *et al.* Molecular Characterization of MSI-H Colorectal Cancer by *MLH1* Promoter Methylation, Immunohistochemistry, and Mismatch

- Repair Germline Mutation Screening. *Cancer Epidemiology, Biomarkers & Prevention*. 2008;17:3208–15. doi: 10.1158/1055-9965.EPI-08-0512
- 152 Fishel R, Lescoe MK, Rao MRS, *et al*. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. *Cell*. 1993;75:1027–38. doi: 10.1016/0092-8674(93)90546-3
- 153 Yokoyama T, Takehara K, Sugimoto N, *et al*. Lynch syndrome-associated endometrial carcinoma with MLH1 germline mutation and MLH1 promoter hypermethylation: a case report and literature review. *BMC Cancer*. 2018;18:576. doi: 10.1186/s12885-018-4489-0
- 154 Russell H, Kedzierska K, Buchanan DD, *et al*. The MLH1 polymorphism rs1800734 and risk of endometrial cancer with microsatellite instability. *Clin Epigenetics*. 2020;12:102. doi: 10.1186/s13148-020-00889-3
- 155 Palles C, Cazier J-B, Howarth KM, *et al*. Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas. *Nat Genet*. 2013;45:136–44. doi: 10.1038/ng.2503
- 156 Goldsby RE, Lawrence NA, Hays LE, *et al*. Defective DNA polymerase- δ proofreading causes cancer susceptibility in mice. *Nat Med*. 2001;7:638–9. doi: 10.1038/88963
- 157 Goldsby RE, Hays LE, Chen X, *et al*. High incidence of epithelial cancers in mice deficient for DNA polymerase δ proofreading. *Proceedings of the National Academy of Sciences*. 2002;99:15560–5. doi: 10.1073/pnas.232340999
- 158 Albertson TM, Ogawa M, Bugni JM, *et al*. DNA polymerase ϵ and δ proofreading suppress discrete mutator and cancer phenotypes in mice. *Proceedings of the National Academy of Sciences*. 2009;106:17101–4. doi: 10.1073/pnas.0907147106
- 159 Church DN, Briggs SEW, Palles C, *et al*. DNA polymerase ϵ and δ exonuclease domain mutations in endometrial cancer. *Hum Mol Genet*. 2013;22:2820–8. doi: 10.1093/hmg/ddt131
- 160 Erson-Omay EZ, Çağlayan AO, Schultz N, *et al*. Somatic *POLE* mutations cause an ultramutated giant cell high-grade glioma subtype with better prognosis. *Neuro Oncol*. 2015;17:1356–64. doi: 10.1093/neuonc/nov027
- 161 Alexandrov LB, Nik-Zainal S, Wedge DC, *et al*. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415–21. doi: 10.1038/nature12477
- 162 Shinbrot E, Henninger EE, Weinhold N, *et al*. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res*. 2014;24:1740–50. doi: 10.1101/gr.174789.114
- 163 Goldsby RE, Lawrence NA, Hays LE, *et al*. Defective DNA polymerase- δ proofreading causes cancer susceptibility in mice. *Nat Med*. 2001;7:638–9. doi: 10.1038/88963
- 164 Shinbrot E, Henninger EE, Weinhold N, *et al*. Exonuclease mutations in DNA polymerase epsilon reveal replication strand specific mutation patterns and human origins of replication. *Genome Res*. 2014;24:1740–50. doi: 10.1101/gr.174789.114

- 165 Podust VN, Chang L-S, Ott R, *et al.* Reconstitution of Human DNA Polymerase δ Using Recombinant Baculoviruses. *Journal of Biological Chemistry*. 2002;277:3894–901. doi: 10.1074/jbc.M109684200
- 166 Murga M, Lecona E, Kamileri I, *et al.* POLD3 Is Haploinsufficient for DNA Replication in Mice. *Mol Cell*. 2016;63:877–83. doi: 10.1016/j.molcel.2016.07.007
- 167 Jiang C, Fan F, Xu W, *et al.* POLD4 Promotes Glioma Cell Proliferation and Suppressive Immune Microenvironment: A Pan-Cancer Analysis Integrated with Experimental Validation. *Int J Mol Sci*. 2023;24:13919. doi: 10.3390/ijms241813919
- 168 Figueroa JD, Malats N, Real FX, *et al.* Genetic variation in the base excision repair pathway and bladder cancer risk. *Hum Genet*. 2007;121:233–42. doi: 10.1007/s00439-006-0294-y
- 169 Sliwinski T, Ziemba P, Morawiec Z, *et al.* Polymorphisms of the DNA polymerase β gene in breast cancer. *Breast Cancer Res Treat*. 2007;103:161–6. doi: 10.1007/s10549-006-9357-y
- 170 Khanra K, Bhattacharya C, Bhattacharyya N. Association of a Newly Identified Variant of DNA Polymerase Beta (pol β Δ 63-123, 208-304) with the Risk Factor of Ovarian Carcinoma in India. *Asian Pacific Journal of Cancer Prevention*. 2012;13:1999–2002. doi: 10.7314/APJCP.2012.13.5.1999
- 171 Canitrot Y, Cazaux C, Fréchet M, *et al.* Overexpression of DNA polymerase β in cell results in a mutator phenotype and a decreased sensitivity to anticancer drugs. *Proceedings of the National Academy of Sciences*. 1998;95:12586–90. doi: 10.1073/pnas.95.21.12586
- 172 Chou K. DNA Polymerase Eta and Chemotherapeutic Agents. *Antioxid Redox Signal*. 2011;14:2521–9. doi: 10.1089/ars.2010.3673
- 173 Silvestri R, Landi S. DNA polymerases in the risk and prognosis of colorectal and pancreatic cancers. *Mutagenesis*. Published Online First: 24 October 2019. doi: 10.1093/mutage/gez031
- 174 Hughes P, Tratner I, Ducoux M, *et al.* Isolation and identification of the third subunit of mammalian DNA polymerase by PCNA-affinity chromatography of mouse FM3A cell extracts. *Nucleic Acids Res*. 1999;27:2108–14. doi: 10.1093/nar/27.10.2108
- 175 Shikata K, Ohta S, Yamada K, *et al.* The Human Homologue of Fission Yeast cdc27, p66, Is a Component of Active Human DNA Polymerase δ . *J Biochem*. 2001;129:699–708. doi: 10.1093/oxfordjournals.jbchem.a002909
- 176 Fuchs J, Cheblal A, Gasser SM. Underappreciated Roles of DNA Polymerase δ in Replication Stress Survival. *Trends in Genetics*. 2021;37:476–87. doi: 10.1016/j.tig.2020.12.003
- 177 Tumini E, Barroso S, Calero CP, *et al.* Roles of human POLD1 and POLD3 in genome stability. *Sci Rep*. 2016;6:38873. doi: 10.1038/srep38873
- 178 Jaeger E, Webb E, Howarth K, *et al.* Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. *Nat Genet*. 2008;40:26–8. doi: 10.1038/ng.2007.41
- 179 Meta-analysis of genome-wide association data identifies four new susceptibility loci for colorectal cancer. *Nat Genet*. 2008;40:1426–35. doi: 10.1038/ng.262

- 180 Tomlinson IP, Webb E, Carvajal-Carmona L, *et al.* A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. *Nat Genet.* 2008;40:623–30. doi: 10.1038/ng.111
- 181 Tomlinson IPM, Carvajal-Carmona LG, Dobbins SE, *et al.* Multiple Common Susceptibility Variants near BMP Pathway Loci GREM1, BMP4, and BMP2 Explain Part of the Missing Heritability of Colorectal Cancer. *PLoS Genet.* 2011;7:e1002105. doi: 10.1371/journal.pgen.1002105
- 182 Schmit SL, Edlund CK, Schumacher FR, *et al.* Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *JNCI: Journal of the National Cancer Institute.* 2019;111:146–57. doi: 10.1093/jnci/djy099
- 183 Schumacher FR, Schmit SL, Jiao S, *et al.* Genome-wide association study of colorectal cancer identifies six new susceptibility loci. *Nat Commun.* 2015;6:7138. doi: 10.1038/ncomms8138
- 184 Dunlop MG, Dobbins SE, Farrington SM, *et al.* Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet.* 2012;44:770–6. doi: 10.1038/ng.2293
- 185 Esteban-Jurado C, Giménez-Zaragoza D, Muñoz J, *et al.* POLE and POLD1 screening in 155 patients with multiple polyps and early-onset colorectal cancer. *Oncotarget.* 2017;8:26732–43. doi: 10.18632/oncotarget.15810
- 186 Tanikawa C, Kamatani Y, Takahashi A, *et al.* GWAS identifies two novel colorectal cancer loci at 16q24.1 and 20q13.12. *Carcinogenesis.* 2018;39:652–60. doi: 10.1093/carcin/bgy026
- 187 Orlando G, Law PJ, Palin K, *et al.* Variation at 2q35 (*PNKD* and *TMBIM1*) influences colorectal cancer risk and identifies a pleiotropic effect with inflammatory bowel disease. *Hum Mol Genet.* 2016;25:2349–59. doi: 10.1093/hmg/ddw087
- 188 Zeng C, Matsuda K, Jia W-H, *et al.* Identification of Susceptibility Loci and Genes for Colorectal Cancer Risk. *Gastroenterology.* 2016;150:1633–45. doi: 10.1053/j.gastro.2016.02.076
- 189 Sun J, Liu X, Gao H, *et al.* Overexpression of colorectal cancer oncogene CHRDL2 predicts a poor prognosis. *Oncotarget.* 2017;8:11489–506. doi: 10.18632/oncotarget.14039
- 190 Chen H, Pan R, Li H, *et al.* CHRDL2 promotes osteosarcoma cell proliferation and metastasis through the BMP-9/PI3K/AKT pathway. *Cell Biol Int.* 2021;45:623–32. doi: 10.1002/cbin.11507
- 191 Leary RJ, Lin JC, Cummins J, *et al.* Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proceedings of the National Academy of Sciences.* 2008;105:16224–9. doi: 10.1073/pnas.0808041105
- 192 Henson JD, Lau LM, Koch S, *et al.* The C-Circle Assay for alternative-lengthening-of-telomeres activity. *Methods.* 2017;114:74–84. doi: 10.1016/j.ymeth.2016.08.016
- 193 Maloum F, Allaire JM, Gagné-Sansfaçon J, *et al.* Epithelial BMP signaling is required for proper specification of epithelial cell lineages and gastric endocrine cells. *American Journal of Physiology-Gastrointestinal and Liver Physiology.* 2011;300:G1065–79. doi: 10.1152/ajpgi.00176.2010
- 194 Reynolds A, Wharton N, Parris A, *et al.* Canonical Wnt signals combined with suppressed TGF β /BMP pathways promote renewal of the native human colonic epithelium. *Gut.* 2014;63:610–21. doi: 10.1136/gutjnl-2012-304067

- 195 T. Das A, Tenenbaum L, Berkhout B. Tet-On Systems For Doxycycline-inducible Gene Expression. *Curr Gene Ther.* 2016;16:156–67. doi: 10.2174/1566523216666160524144041
- 196 Varnat F, Duquet A, Malerba M, *et al.* Human colon cancer epithelial cells harbour active HEDGEHOG-GLI signalling that is essential for tumour growth, recurrence, metastasis and stem cell survival and expansion. *EMBO Mol Med.* 2009;1:338–51. doi: 10.1002/emmm.200900039
- 197 He XC, Zhang J, Tong W-G, *et al.* BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt- β -catenin signaling. *Nat Genet.* 2004;36:1117–21. doi: 10.1038/ng1430
- 198 Qi Z, Li Y, Zhao B, *et al.* BMP restricts stemness of intestinal Lgr5+ stem cells by directly suppressing their signature genes. *Nat Commun.* 2017;8:13824. doi: 10.1038/ncomms13824
- 199 Kwong LN, Dove WF. APC and Its Modifiers in Colon Cancer. 2009:85–106.
- 200 Liberti M V., Locasale JW. The Warburg Effect: How Does it Benefit Cancer Cells? *Trends Biochem Sci.* 2016;41:211–8. doi: 10.1016/j.tibs.2015.12.001
- 201 Navrátilová J, Hankeová T, Beneš P, *et al.* Low-Glucose Conditions of Tumor Microenvironment Enhance Cytotoxicity of Tetrathiomolybdate to Neuroblastoma Cells. *Nutr Cancer.* 2013;65:702–10. doi: 10.1080/01635581.2013.789118
- 202 Peiris-Pagès M, Martínez-Outschoorn UE, Pestell RG, *et al.* Cancer stem cell metabolism. *Breast Cancer Research.* 2016;18:55. doi: 10.1186/s13058-016-0712-6
- 203 Nandy SB, Orozco A, Lopez-Valdez R, *et al.* Glucose insult elicits hyperactivation of cancer stem cells through miR-424–cdc42–prdm14 signalling axis. *Br J Cancer.* 2017;117:1665–75. doi: 10.1038/bjc.2017.335
- 204 Yoshikawa N, Saito Y, Manabe H, *et al.* Glucose Depletion Enhances the Stem Cell Phenotype and Gemcitabine Resistance of Cholangiocarcinoma Organoids through AKT Phosphorylation and Reactive Oxygen Species. *Cancers (Basel).* 2019;11:1993. doi: 10.3390/cancers11121993
- 205 Tauriello DVF, Calon A, Lonardo E, *et al.* Determinants of metastatic competency in colorectal cancer. *Mol Oncol.* 2017;11:97–119. doi: 10.1002/1878-0261.12018
- 206 Fares J, Fares MY, Khachfe HH, *et al.* Molecular principles of metastasis: a hallmark of cancer revisited. *Signal Transduct Target Ther.* 2020;5:28. doi: 10.1038/s41392-020-0134-x
- 207 Hebert JD, Tian C, Lamar JM, *et al.* The scaffold protein IQGAP1 is crucial for extravasation and metastasis. *Sci Rep.* 2020;10:2439. doi: 10.1038/s41598-020-59438-w
- 208 Nabeshima K, Shimaō Y, Inoue T, *et al.* Immunohistochemical analysis of IQGAP1 expression in human colorectal carcinomas: its overexpression in carcinomas and association with invasion fronts. *Cancer Lett.* 2002;176:101–9. doi: 10.1016/S0304-3835(01)00742-X
- 209 Jiang X, Wang S, Liang Q, *et al.* Unraveling the multifaceted role of EpCAM in colorectal cancer: an integrated review of its function and interplay with non-coding RNAs. *Medical Oncology.* 2023;41:35. doi: 10.1007/s12032-023-02273-6
- 210 Mohtar M, Syafruddin S, Nasir S, *et al.* Revisiting the Roles of Pro-Metastatic EpCAM in Cancer. *Biomolecules.* 2020;10:255. doi: 10.3390/biom10020255

- 211 Duval K, Grover H, Han L-H, *et al.* Modeling Physiological Events in 2D vs. 3D Cell Culture. *Physiology*. 2017;32:266–77. doi: 10.1152/physiol.00036.2016
- 212 Petersen OW, Rønnov-Jessen L, Howlett AR, *et al.* Interaction with basement membrane serves to rapidly distinguish growth and differentiation pattern of normal and malignant human breast epithelial cells. *Proceedings of the National Academy of Sciences*. 1992;89:9064–8. doi: 10.1073/pnas.89.19.9064
- 213 Nelson CM, Bissell MJ. Of Extracellular Matrix, Scaffolds, and Signaling: Tissue Architecture Regulates Development, Homeostasis, and Cancer. *Annu Rev Cell Dev Biol*. 2006;22:287–309. doi: 10.1146/annurev.cellbio.22.010305.104315
- 214 Mseka T, Bamberg JR, Cramer LP. ADF/cofilin family proteins control formation of oriented actin-filament bundles in the cell body to trigger fibroblast polarization. *J Cell Sci*. 2007;120:4332–44. doi: 10.1242/jcs.017640
- 215 Sato T, Vries RG, Snippert HJ, *et al.* Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*. 2009;459:262–5. doi: 10.1038/nature07935
- 216 Navrátilová J, Hankeová T, Beneš P, *et al.* Low-Glucose Conditions of Tumor Microenvironment Enhance Cytotoxicity of Tetrathiomolybdate to Neuroblastoma Cells. *Nutr Cancer*. 2013;65:702–10. doi: 10.1080/01635581.2013.789118
- 217 Ying H, Kimmelman AC, Lyssiotis CA, *et al.* Oncogenic Kras Maintains Pancreatic Tumors through Regulation of Anabolic Glucose Metabolism. *Cell*. 2012;149:656–70. doi: 10.1016/j.cell.2012.01.058
- 218 Rajendran V, Jain MV. In Vitro Tumorigenic Assay: Colony Forming Assay for Cancer Stem Cells. 2018:89–95.
- 219 Danielyan L, Schwab M, Siegel G, *et al.* Cell motility and migration as determinants of stem cell efficacy. *EBioMedicine*. 2020;60:102989. doi: 10.1016/j.ebiom.2020.102989
- 220 Moore N, Lyle S. Quiescent, Slow-Cycling Stem Cell Populations in Cancer: A Review of the Evidence and Discussion of Significance. *J Oncol*. 2011;2011:1–11. doi: 10.1155/2011/396076
- 221 Sangiorgi E, Capecchi MR. Bmi1 is expressed in vivo in intestinal stem cells. *Nat Genet*. 2008;40:915–20. doi: 10.1038/ng.165
- 222 Tao S, Tang D, Morita Y, *et al.* Wnt activity and basal niche position sensitize intestinal stem and progenitor cells to DNA damage. *EMBO J*. 2015;34:624–40. doi: 10.15252/embj.201490700
- 223 Sheng X, Lin Z, Lv C, *et al.* Cycling Stem Cells Are Radioresistant and Regenerate the Intestine. *Cell Rep*. 2020;32:107952. doi: 10.1016/j.celrep.2020.107952
- 224 Montgomery RK, Carlone DL, Richmond CA, *et al.* Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proceedings of the National Academy of Sciences*. 2011;108:179–84. doi: 10.1073/pnas.1013004108
- 225 Rees WD, Tandun R, Yau E, *et al.* Regenerative Intestinal Stem Cells Induced by Acute and Chronic Injury: The Saving Grace of the Epithelium? *Front Cell Dev Biol*. 2020;8. doi: 10.3389/fcell.2020.583919

- 226 Malla SB, Byrne RM, Lafarge MW, *et al.* Pathway level subtyping identifies a slow-cycling biological phenotype associated with poor clinical outcomes in colorectal cancer. *Nat Genet.* 2024;56:458–72. doi: 10.1038/s41588-024-01654-5
- 227 Abdullah LN, Chow EK. Mechanisms of chemoresistance in cancer stem cells. *Clin Transl Med.* 2013;2. doi: 10.1186/2001-1326-2-3
- 228 Leowattana W, Leowattana P, Leowattana T. Systemic treatment for metastatic colorectal cancer. *World J Gastroenterol.* 2023;29:1569–88. doi: 10.3748/wjg.v29.i10.1569
- 229 Sobrero A, Guglielmi A, Grossi F, *et al.* Mechanism of action of fluoropyrimidines: relevance to the new developments in colorectal cancer chemotherapy. *Semin Oncol.* 2000;27:72–7.
- 230 Gill S, Thomas RR, Goldberg RM. Colorectal cancer chemotherapy. *Aliment Pharmacol Ther.* 2003;18:683–92. doi: 10.1046/j.1365-2036.2003.01735.x
- 231 Pasetto LM, Jirillo A, Iadicicco G, *et al.* FOLFOX versus FOLFIRI: a comparison of regimens in the treatment of colorectal cancer metastases. *Anticancer Res.* 2005;25:563–76.
- 232 Lieber MR. The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End-Joining Pathway. *Annu Rev Biochem.* 2010;79:181–211. doi: 10.1146/annurev.biochem.052308.093131
- 233 Chiu S-J, Lee Y-J, Hsu T-S, *et al.* Oxaliplatin-induced gamma-H2AX activation via both p53-dependent and -independent pathways but is not associated with cell cycle arrest in human colorectal cancer cells. *Chem Biol Interact.* 2009;182:173–82. doi: 10.1016/j.cbi.2009.08.019
- 234 Srinivas US, Dyczkowski J, Reißbarth T, *et al.* 5-Fluorouracil sensitizes colorectal tumor cells towards double stranded DNA breaks by interfering with homologous recombination repair. *Oncotarget.* 2015;6:12574–86. doi: 10.18632/oncotarget.3728
- 235 Mah L-J, El-Osta A, Karagiannis TC. γ H2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia.* 2010;24:679–86. doi: 10.1038/leu.2010.6
- 236 Marechal A, Zou L. DNA Damage Sensing by the ATM and ATR Kinases. *Cold Spring Harb Perspect Biol.* 2013;5:a012716–a012716. doi: 10.1101/cshperspect.a012716
- 237 Häfner MF, Debus J. Radiotherapy for Colorectal Cancer: Current Standards and Future Perspectives. *Visc Med.* 2016;32:172–7. doi: 10.1159/000446486
- 238 Nassar D, Blanpain C. Cancer Stem Cells: Basic Concepts and Therapeutic Implications. *Annual Review of Pathology: Mechanisms of Disease.* 2016;11:47–76. doi: 10.1146/annurev-pathol-012615-044438
- 239 Gillespie MS, Ward CM, Davies CC. DNA Repair and Therapeutic Strategies in Cancer Stem Cells. *Cancers (Basel).* 2023;15:1897. doi: 10.3390/cancers15061897
- 240 Sedelnikova OA, Horikawa I, Zimonjic DB, *et al.* Senescing human cells and ageing mice accumulate DNA lesions with unreparable double-strand breaks. *Nat Cell Biol.* 2004;6:168–70. doi: 10.1038/ncb1095
- 241 Longley DB, Harkin DP, Johnston PG. 5-Fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer.* 2003;3:330–8. doi: 10.1038/nrc1074

- 242 Phi LTH, Sari IN, Yang Y-G, *et al.* Cancer Stem Cells (CSCs) in Drug Resistance and their Therapeutic Implications in Cancer Treatment. *Stem Cells Int.* 2018;2018:1–16. doi: 10.1155/2018/5416923
- 243 Ng KP, Manjeri A, Lee KL, *et al.* Physiologic hypoxia promotes maintenance of CML stem cells despite effective BCR-ABL1 inhibition. *Blood.* 2014;123:3316–26. doi: 10.1182/blood-2013-07-511907
- 244 Wang Q-E. DNA damage responses in cancer stem cells: Implications for cancer therapeutic strategies. *World J Biol Chem.* 2015;6:57. doi: 10.4331/wjbc.v6.i3.57
- 245 Bao S, Wu Q, McLendon RE, *et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature.* 2006;444:756–60. doi: 10.1038/nature05236
- 246 Liu J, Xiao Q, Xiao J, *et al.* Wnt/ β -catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduct Target Ther.* 2022;7:3. doi: 10.1038/s41392-021-00762-6
- 247 Wright JB, Brown SJ, Cole MD. Upregulation of c- MYC in cis through a Large Chromatin Loop Linked to a Cancer Risk-Associated Single-Nucleotide Polymorphism in Colorectal Cancer Cells. *Mol Cell Biol.* 2010;30:1411–20. doi: 10.1128/MCB.01384-09
- 248 Kaldis P, Pagano M. Wnt Signaling in Mitosis. *Dev Cell.* 2009;17:749–50. doi: 10.1016/j.devcel.2009.12.001
- 249 Zhan T, Rindtorff N, Boutros M. Wnt signaling in cancer. *Oncogene.* 2017;36:1461–73. doi: 10.1038/onc.2016.304
- 250 Park J-I, Venteicher AS, Hong JY, *et al.* Telomerase modulates Wnt signalling by association with target gene chromatin. *Nature.* 2009;460:66–72. doi: 10.1038/nature08137
- 251 Schwitalla S, Fingerle AA, Cammareri P, *et al.* Intestinal Tumorigenesis Initiated by Dedifferentiation and Acquisition of Stem-Cell-like Properties. *Cell.* 2013;152:25–38. doi: 10.1016/j.cell.2012.12.012
- 252 Wu Z-Q, Li X-Y, Hu CY, *et al.* Canonical Wnt signaling regulates Slug activity and links epithelial–mesenchymal transition with epigenetic Breast Cancer 1, Early Onset (BRCA1) repression. *Proceedings of the National Academy of Sciences.* 2012;109:16654–9. doi: 10.1073/pnas.1205822109
- 253 Steinhardt AA, Gayyed MF, Klein AP, *et al.* Expression of Yes-associated protein in common solid tumors. *Hum Pathol.* 2008;39:1582–9. doi: 10.1016/j.humpath.2008.04.012
- 254 Azzolin L, Zanconato F, Bresolin S, *et al.* Role of TAZ as Mediator of Wnt Signaling. *Cell.* 2012;151:1443–56. doi: 10.1016/j.cell.2012.11.027
- 255 Kovács SA, Fekete JT, Győrffy B. Predictive biomarkers of immunotherapy response with pharmacological applications in solid tumors. *Acta Pharmacol Sin.* 2023;44:1879–89. doi: 10.1038/s41401-023-01079-6
- 256 Geng L, Rachakonda G, Morré DJ, *et al.* Indolyl-quinuclidinols inhibit ENOX activity and endothelial cell morphogenesis while enhancing radiation-mediated control of tumor vasculature. *The FASEB Journal.* 2009;23:2986–95. doi: 10.1096/fj.09-130005

- 257 Smith CA, Mont S, Traver G, *et al.* Targeting Enox1 in tumor stroma increases the efficacy of fractionated radiotherapy. *Oncotarget*. 2016;7:77926–36. doi: 10.18632/oncotarget.12845
- 258 Venkateswaran A, Sekhar KR, Levic DS, *et al.* The NADH Oxidase ENOX1, a Critical Mediator of Endothelial Cell Radiosensitization, Is Crucial for Vascular Development. *Cancer Res*. 2014;74:38–43. doi: 10.1158/0008-5472.CAN-13-1981
- 259 Song Y, Bi Z, Liu Y, *et al.* Targeting RAS–RAF–MEK–ERK signaling pathway in human cancer: Current status in clinical trials. *Genes Dis*. 2023;10:76–88. doi: 10.1016/j.gendis.2022.05.006
- 260 Klebanoff CA, Finkelstein SE, Surman DR, *et al.* IL-15 enhances the *in vivo* antitumor activity of tumor-reactive CD8 + T Cells. *Proceedings of the National Academy of Sciences*. 2004;101:1969–74. doi: 10.1073/pnas.0307298101
- 261 Lange SS, Takata K, Wood RD. DNA polymerases and cancer. *Nat Rev Cancer*. 2011;11:96–110. doi: 10.1038/nrc2998
- 262 Prindle MJ, Loeb LA. DNA polymerase delta in dna replication and genome maintenance. *Environ Mol Mutagen*. 2012;53:666–82. doi: 10.1002/em.21745
- 263 Cohen S, Guenolé A, Lazar I, *et al.* A POLD3/BLM dependent pathway handles DSBs in transcribed chromatin upon excessive RNA:DNA hybrid accumulation. *Nat Commun*. 2022;13:2012. doi: 10.1038/s41467-022-29629-2
- 264 Sung H, Ferlay J, Siegel RL, *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin*. 2021;71:209–49. doi: 10.3322/caac.21660
- 265 Barresi V, Cinnirella G, Valenti G, *et al.* Gene expression profiles in genome instability-based classes of colorectal cancer. *BMC Cancer*. 2018;18:1265. doi: 10.1186/s12885-018-5174-z
- 266 Dunlop MG, Dobbins SE, Farrington SM, *et al.* Common variation near CDKN1A, POLD3 and SHROOM2 influences colorectal cancer risk. *Nat Genet*. 2012;44:770–6. doi: 10.1038/ng.2293
- 267 Spier I, Holzapfel S, Altmüller J, *et al.* Frequency and phenotypic spectrum of germline mutations in *POLE* and seven other polymerase genes in 266 patients with colorectal adenomas and carcinomas. *Int J Cancer*. 2015;137:320–31. doi: 10.1002/ijc.29396
- 268 Beroukhi R, Mermel CH, Porter D, *et al.* The landscape of somatic copy-number alteration across human cancers. *Nature*. 2010;463:899–905. doi: 10.1038/nature08822
- 269 Bryan TM, Englezou A, Dalla-Pozza L, *et al.* Evidence for an alternative mechanism for maintaining telomere length in human tumors and tumor-derived cell lines. *Nat Med*. 1997;3:1271–4. doi: 10.1038/nm1197-1271
- 270 Gaspar TB, Sá A, Lopes JM, *et al.* Telomere Maintenance Mechanisms in Cancer. *Genes (Basel)*. 2018;9:241. doi: 10.3390/genes9050241
- 271 Kim S, Chowdhury T, Yu HJ, *et al.* The telomere maintenance mechanism spectrum and its dynamics in gliomas. *Genome Med*. 2022;14:88. doi: 10.1186/s13073-022-01095-x
- 272 Recagni M, Bidzinska J, Zaffaroni N, *et al.* The Role of Alternative Lengthening of Telomeres Mechanism in Cancer: Translational and Therapeutic Implications. *Cancers (Basel)*. 2020;12:949. doi: 10.3390/cancers12040949

- 273 Yeager TR, Neumann AA, Englezou A, *et al.* Telomerase-negative immortalized human cells contain a novel type of promyelocytic leukemia (PML) body. *Cancer Res.* 1999;59:4175–9.
- 274 Zhang J-M, Zou L. Alternative lengthening of telomeres: from molecular mechanisms to therapeutic outlooks. *Cell Biosci.* 2020;10:30. doi: 10.1186/s13578-020-00391-6
- 275 Zhou Z, Wang L, Ge F, *et al.* *Pold3* is required for genomic stability and telomere integrity in embryonic stem cells and meiosis. *Nucleic Acids Res.* 2018;46:3468–86. doi: 10.1093/nar/gky098
- 276 Costantino L, Sotiriou SK, Rantala JK, *et al.* Break-Induced Replication Repair of Damaged Forks Induces Genomic Duplications in Human Cells. *Science (1979).* 2014;343:88–91. doi: 10.1126/science.1243211
- 277 Dilley RL, Verma P, Cho NW, *et al.* Break-induced telomere synthesis underlies alternative telomere maintenance. *Nature.* 2016;539:54–8. doi: 10.1038/nature20099
- 278 Fuchs J, Cheblal A, Gasser SM. Underappreciated Roles of DNA Polymerase δ in Replication Stress Survival. *Trends in Genetics.* 2021;37:476–87. doi: 10.1016/j.tig.2020.12.003
- 279 Kim W, Ludlow AT, Minutes J, *et al.* Regulation of the Human Telomerase Gene TERT by Telomere Position Effect—Over Long Distances (TPE-OLD): Implications for Aging and Cancer. *PLoS Biol.* 2016;14:e2000016. doi: 10.1371/journal.pbio.2000016
- 280 Rosso I, Jones-Weinert C, Rossiello F, *et al.* Alternative lengthening of telomeres (ALT) cells viability is dependent on C-rich telomeric RNAs. *Nat Commun.* 2023;14:7086. doi: 10.1038/s41467-023-42831-0
- 281 Li X, Heyer W-D. Homologous recombination in DNA repair and DNA damage tolerance. *Cell Res.* 2008;18:99–113. doi: 10.1038/cr.2008.1
- 282 Gong C, Yang H, Wang S, *et al.* hTERT Promotes CRC Proliferation and Migration by Recruiting YBX1 to Increase NRF2 Expression. *Front Cell Dev Biol.* 2021;9:658101. doi: 10.3389/fcell.2021.658101
- 283 Toaldo C, Pizzimenti S, Cerbone A, *et al.* PPAR γ ligands inhibit telomerase activity and hTERT expression through modulation of the Myc/Mad/Max network in colon cancer cells. *J Cell Mol Med.* 2010;14:1347–57. doi: 10.1111/j.1582-4934.2009.00966.x
- 284 Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. *Nature.* 1997;386:623–7. doi: 10.1038/386623a0
- 285 Orsetti B, Selves J, Bascoul-Mollevi C, *et al.* Impact of chromosomal instability on colorectal cancer progression and outcome. *BMC Cancer.* 2014;14:121. doi: 10.1186/1471-2407-14-121
- 286 Guterres AN, Villanueva J. Targeting telomerase for cancer therapy. *Oncogene.* 2020;39:5811–24. doi: 10.1038/s41388-020-01405-w
- 287 Law PJ, Timofeeva M, Fernandez-Rozadilla C, *et al.* Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun.* 2019;10:2154. doi: 10.1038/s41467-019-09775-w

- 288 Orlando G, Law PJ, Palin K, *et al.* Variation at 2q35 (*PNKD* and *TMBIM1*) influences colorectal cancer risk and identifies a pleiotropic effect with inflammatory bowel disease. *Hum Mol Genet.* 2016;25:2349–59. doi: 10.1093/hmg/ddw087
- 289 Schmit SL, Edlund CK, Schumacher FR, *et al.* Novel Common Genetic Susceptibility Loci for Colorectal Cancer. *JNCI: Journal of the National Cancer Institute.* 2019;111:146–57. doi: 10.1093/jnci/djy099
- 290 Bajaj J, Diaz E, Reya T. Stem cells in cancer initiation and progression. *Journal of Cell Biology.* 2020;219. doi: 10.1083/jcb.201911053
- 291 Barker N, Ridgway RA, van Es JH, *et al.* Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature.* 2009;457:608–11. doi: 10.1038/nature07602
- 292 Preston SL, Wong W-M, Chan AO-O, *et al.* Bottom-up histogenesis of colorectal adenomas: origin in the monocryptal adenoma and initial expansion by crypt fission. *Cancer Res.* 2003;63:3819–25.
- 293 Shih I-M, Wang T-L, Traverso G, *et al.* Top-down morphogenesis of colorectal tumors. *Proceedings of the National Academy of Sciences.* 2001;98:2640–5. doi: 10.1073/pnas.051629398
- 294 Schwitalla S, Fingerle AA, Cammareri P, *et al.* Intestinal Tumorigenesis Initiated by Dedifferentiation and Acquisition of Stem-Cell-like Properties. *Cell.* 2013;152:25–38. doi: 10.1016/j.cell.2012.12.012
- 295 Verhagen MP, Joosten R, Schmitt M, *et al.* Non-stem cell lineages as an alternative origin of intestinal tumorigenesis in the context of inflammation. *Nat Genet.* 2024;56:1456–67. doi: 10.1038/s41588-024-01801-y
- 296 Litman T, Brangi M, Hudson E, *et al.* The multidrug-resistant phenotype associated with overexpression of the new ABC half-transporter, MXR (ABCG2). *J Cell Sci.* 2000;113:2011–21. doi: 10.1242/jcs.113.11.2011
- 297 Ikawa M, Impraime CC, Wang G, *et al.* Isolation and characterization of aldehyde dehydrogenase isozymes from usual and atypical human livers. *J Biol Chem.* 1983;258:6282–7.
- 298 Kim R, Emi M, Tanabe K. Role of mitochondria as the gardens of cell death. *Cancer Chemother Pharmacol.* 2006;57:545–53. doi: 10.1007/s00280-005-0111-7
- 299 Noda T, Nagano H, Takemasa I, *et al.* Activation of Wnt/ β -catenin signalling pathway induces chemoresistance to interferon- α /5-fluorouracil combination therapy for hepatocellular carcinoma. *Br J Cancer.* 2009;100:1647–58. doi: 10.1038/sj.bjc.6605064
- 300 Yang W, Yan H-X, Chen L, *et al.* Wnt/ β -Catenin Signaling Contributes to Activation of Normal and Tumorigenic Liver Progenitor Cells. *Cancer Res.* 2008;68:4287–95. doi: 10.1158/0008-5472.CAN-07-6691
- 301 Arango D, Wilson AJ, Shi Q, *et al.* Molecular mechanisms of action and prediction of response to oxaliplatin in colorectal cancer cells. *Br J Cancer.* 2004;91:1931–46. doi: 10.1038/sj.bjc.6602215

- 302 DU M, SU X-M, ZHANG T, *et al.* Aberrant promoter DNA methylation inhibits bone morphogenetic protein 2 expression and contributes to drug resistance in breast cancer. *Mol Med Rep.* 2014;10:1051–5. doi: 10.3892/mmr.2014.2276
- 303 LIU G, LIU Y-J, LIAN W-J, *et al.* Reduced BMP6 expression by DNA methylation contributes to EMT and drug resistance in breast cancer cells. *Oncol Rep.* 2014;32:581–8. doi: 10.3892/or.2014.3224
- 304 Eramo A, Ricci-Vitiani L, Zeuner A, *et al.* Chemotherapy resistance of glioblastoma stem cells. *Cell Death Differ.* 2006;13:1238–41. doi: 10.1038/sj.cdd.4401872
- 305 Tate CM, Pallini R, Ricci-Vitiani L, *et al.* A BMP7 variant inhibits the tumorigenic potential of glioblastoma stem-like cells. *Cell Death Differ.* 2012;19:1644–54. doi: 10.1038/cdd.2012.44
- 306 LIAN W-J, LIU G, LIU Y-J, *et al.* Downregulation of BMP6 enhances cell proliferation and chemoresistance via activation of the ERK signaling pathway in breast cancer. *Oncol Rep.* 2013;30:193–200. doi: 10.3892/or.2013.2462
- 307 Ciuculan L, Sheppard K, Dong L, *et al.* Treatment with Anti-Gremlin 1 Antibody Ameliorates Chronic Hypoxia/SU5416-Induced Pulmonary Arterial Hypertension in Mice. *Am J Pathol.* 2013;183:1461–73. doi: 10.1016/j.ajpath.2013.07.017
- 308 Kim M, Yoon S, Lee S, *et al.* Gremlin-1 Induces BMP-Independent Tumor Cell Proliferation, Migration, and Invasion. *PLoS One.* 2012;7:e35100. doi: 10.1371/journal.pone.0035100
- 309 Kobayashi H, Gieniec KA, Wright JA, *et al.* The Balance of Stromal BMP Signaling Mediated by GREM1 and ISLR Drives Colorectal Carcinogenesis. *Gastroenterology.* 2021;160:1224-1239.e30. doi: 10.1053/j.gastro.2020.11.011
- 310 Cheng C, Wang J, Xu P, *et al.* Gremlin1 is a therapeutically targetable FGFR1 ligand that regulates lineage plasticity and castration resistance in prostate cancer. *Nat Cancer.* 2022;3:565–80. doi: 10.1038/s43018-022-00380-3
- 311 Malkova A, Ira G. Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev.* 2013;23:271–9. doi: 10.1016/j.gde.2013.05.007
- 312 Pardo B, Aguilera A. Complex Chromosomal Rearrangements Mediated by Break-Induced Replication Involve Structure-Selective Endonucleases. *PLoS Genet.* 2012;8:e1002979. doi: 10.1371/journal.pgen.1002979
- 313 Minocherhomji S, Ying S, Bjerregaard VA, *et al.* Replication stress activates DNA repair synthesis in mitosis. *Nature.* 2015;528:286–90. doi: 10.1038/nature16139
- 314 Valle L, Hernández-Illán E, Bellido F, *et al.* New insights into POLE and POLD1 germline mutations in familial colorectal cancer and polyposis. *Hum Mol Genet.* 2014;23:3506–12. doi: 10.1093/hmg/ddu058

Appendices

Appendix 1:

Linux scripts for RNAseq analysis and data processing

Step 1. Mapping

```
#!/bin/bash
#SBATCH -n 1
#SBATCH -c 1
#SBATCH -J JOB NAME
#SBATCH -t 48:00:00
#SBATCH --mem-per-cpu=10G
#SBATCH -o SBATCH_XXX_out.txt
#SBATCH -e SBATCH_XXX_err.txt

source /home/2151225/.bashrc; cd /home/2151225/scratch/RNAseqEloise/TOPHAT/XXX ;
/home/2151225/scratch/TOOLS/tophat-2.1.1.Linux_x86_64/tophat2 -p 2 -o
/home/2151225/scratch/RNAseqEloise/TOPHAT/XXX
/home/2151225/scratch/RNAseq/DNA/Homo_sapiens.GRCh38.dna
/home/2151225/scratch/RNAseqEloise/XXX.fastq.gz
/home/2151225/scratch/RNAseqEloise/XXX.fastq.gz
```

Step 2. Merge/Filter

```
#!/bin/bash
#SBATCH -n 1
#SBATCH -c 1
#SBATCH -J Merge_filter
#SBATCH -t 12:00:00
#SBATCH --mem-per-cpu=10G
#SBATCH -o SBATCH_Rachel_filter_out.txt
#SBATCH -e SBATCH_Rachel_filter_err.txt

module load samtools/1.12

cd /home/ 2151225 /scratch/ RNAseqEloise /TOPHAT/XXX ; samtools view -bq 30 accepted_hits.bam
> XXX.accepted_hits.hg38.q30.bam
```

Step 3. Cufflinks

```
#!/bin/bash
#SBATCH -n 1
#SBATCH -c 1
#SBATCH -J Cufflinks_XXX
```

```

#SBATCH -t 12:00:00

#SBATCH --mem-per-cpu=10G

#SBATCH -o SBATCH_cuffXXX_out.txt

#SBATCH -e SBATCH_cuffXXX_err.txt

source /home/2151225/.bashrc; cd /home/2151225/scratch/ RNAseqEloise /TOPHAT/XXX ; cufflinks -G
/home/2151225/scratch/RNAseq/GENCODE/gencode.v38.chr_patch_hapl_scaff.annotation.exon.gtf -o

/home/2151225/scratch/ RNAseqEloise CUFFLINKS_DF/XXX

/home/2151225/scratch/ RNAseqEloise /TOPHAT_DF/XXX/XXX.accepted_hits.hg38.q30.bam

```

Step 4. Cuffdiff

```

#!/bin/bash

#SBATCH -n 1

#SBATCH -c 1

#SBATCH -J DMSOvLOW

#SBATCH -t 48:00:00

#SBATCH --mem-per-cpu=10G

#SBATCH -o SBATCH_DMSOvLOW_out.txt

#SBATCH -e SBATCH_DMSOvLOW_err.txt

source /home/2151225/.bashrc; cd /home/2151225/scratch/RNAseqEloise/TOPHAT/RESULTS ;
cuffdiff -L DMSO,LOW

/home/2151225/scratch/OOC/EXAMPLE/INPUT/GENCODE/hg38/gencode.v29.annotation.chr22.exon.g
tf -o

/home/2151225/scratch/RNAseqEloise/CUFFDIFF/DMSOvLOW DMSO.accepted_hits.hg38.q30.bam
LOW.accepted_hits.hg38.q30.bam

```

R studio analysis for data visualization

#Load data and create matrix

```

low <-
read.csv("/Users/User/Documents/PhD/Data/RNAseq/CUFFDIFF/ChrDMSOvChrLOW/gene_exp.diff",
header=TRUE,sep="\t")

med <-
read.csv("/Users/User/Documents/PhD/Data/RNAseq/CUFFDIFF/ChrDMSOvChrMED/gene_exp.diff",
header=TRUE,sep="\t")

high <-
read.csv("/Users/User/Documents/PhD/Data/RNAseq/CUFFDIFF/ChrDMSOvChrHIGH/gene_exp.diff",
header=TRUE,sep="\t")

as.data.frame(low)

as.data.frame(med)

as.data.frame(high)

low <- low[,c(2,3,8,9,10,12,14)]

med <- med[,c(2,3,8,9,10,12,14)]

```

```

high <- high[,c(2,8,9,10,12,14)]
head(low)
head(med)
head(high)
temp2<- merge(low, med, by="gene_id")
countdata<- merge(temp2, high, by= "gene_id")
head(countdata)
dim(countdata)
countdata<-countdata[,c(1,2,3,4,5,6,7,9,10,11,12,13,14,15,16,17,18)]
head(countdata)
colnames(countdata)[1]<-"Gene_ID"
colnames(countdata)[2]<-"Gene"
colnames(countdata)[3]<-"DMSO"
colnames(countdata)[4]<-"LOW"
colnames(countdata)[5]<-"Fold_changeLOW"
colnames(countdata)[6]<-"PVALLOW"
colnames(countdata)[7]<-"SignificantLOW"
colnames(countdata)[8]<-"DMSO2"
colnames(countdata)[9]<-"MED"
colnames(countdata)[10]<-"Fold_changeMED"
colnames(countdata)[11]<-"PVALMED"
colnames(countdata)[12]<-"SignificantMED"
colnames(countdata)[13]<-"DMSO3"
colnames(countdata)[14]<-"HIGH"
colnames(countdata)[15]<-"Fold_changeHIGH"
colnames(countdata)[16]<-"PVALHIGH"
colnames(countdata)[17]<-"SignificantHIGH"

```

#Adding “up or down” annotation

```

countdata$RegulationLOW <- "No Change"
countdata<-within(countdata, RegulationLOW[Fold_changeLOW> (1)]<-"Up")
countdata<-within(countdata, RegulationLOW[Fold_changeLOW< (-1)]<-"Down")
countdata$RegulationMED <- "No Change"
countdata<-within(countdata, RegulationMED[Fold_changeMED> (1)]<-"Up")
countdata<-within(countdata, RegulationMED[Fold_changeMED< (-1)]<-"Down")
countdata$RegulationHIGH <- "No Change"
countdata<-within(countdata, RegulationHIGH[Fold_changeHIGH> (1)]<-"Up")
countdata<-within(countdata, RegulationHIGH[Fold_changeHIGH< (-1)]<-"Down")

```

#Plot volcano graph

```
countdata[1:3,c("DMSO","LOW","Fold_changeLOW","RegulationLOW")]
countdata[1:3,c("DMSO2","MED","Fold_changeMED","RegulationMED")]
countdata[1:3,c("DMSO3","HIGH","Fold_changeHIGH","RegulationHIGH")]
head(countdata)
countdata$SignificanceLOW <- "NS"
countdata<-within(countdata, SignificanceLOW[PVALLOW<0.05]<-"**")
countdata<-within(countdata, SignificanceLOW[PVALLOW<0.01]<-"****")
countdata<-within(countdata, SignificanceLOW[PVALLOW<0.001]<-"*****")
countdata$SignificanceMED <- "NS"
countdata<-within(countdata, SignificanceMED[PVALMED<0.05]<-"**")
countdata<-within(countdata, SignificanceMED[PVALMED<0.01]<-"****")
countdata<-within(countdata, SignificanceMED[PVALMED<0.001]<-"*****")
countdata$SignificanceHIGH <- "NS"
countdata<-within(countdata, SignificanceHIGH[PVALHIGH<0.05]<-"**")
countdata<-within(countdata, SignificanceHIGH[PVALHIGH<0.01]<-"****")
countdata<-within(countdata, SignificanceHIGH[PVALHIGH<0.001]<-"*****")
dim(countdata)
view(countdata)
as.matrix(countdata)
ggplot(countdata,aes(x=Fold_changeLOW,y=-log10(PVALLOW),col=SignificanceLOW, label=Gene))+
geom_point(size=1) +
geom_text_repel(aes(label=ifelse(PVALLOW<0.01,as.character(Gene,"")),hjust=0, vjust=1, size=2)
+scale_color_manual(values=c("Blue", "Red", "Purple", "Gray"))+ labs(x="Log2Fold Change", y="-
log10(P-Value)")
ggplot(countdata,aes(x=Fold_changeMED,y=-log10(PVALMED),col=SignificanceMED, label=Gene))+
geom_point(size=1) +
geom_text_repel(aes(label=ifelse(PVALMED<0.01,as.character(Gene,"")),hjust=0, vjust=1, size=2)
+scale_color_manual(values=c("Blue", "Red", "Purple", "Gray"))+ labs(x="Log2Fold Change", y="-
log10(P-Value)")
ggplot(countdata,aes(x=Fold_changeHIGH,y=-log10(PVALHIGH),col=SignificanceHIGH, label=Gene))+
geom_point(size=1) +
geom_text_repel(aes(label=ifelse(PVALHIGH<0.01,as.character(Gene,"")),hjust=0, vjust=1, size=2)
+scale_color_manual(values=c("Blue", "Red", "Purple", "Gray"))+ labs(x="Log2Fold Change", y="-
log10(P-Value)")
library(ggrepel)
```

#Obtain a list of significant genes

```
medsig <- countdata[countdata$PVALMED<0.01,]
highsig <- countdata[countdata$PVALHIGH<0.01,]
```

```

medsig1<- medsig[,c(2,6,13)]
write_xlsx(medsig, "/Users/User/Documents/PhD/Data/medsig.xlsx")
highsig1<- highsig[,c(2,11,14)]
write_xlsx(highsig, "/Users/User/Documents/PhD/Data/RNAseq data/highsig.xlsx")
as.data.frame(medandhighsig)
medandhighsig<- Reduce(intersect, list(medsig$Gene,highsig$Gene))
write_xlsx(medandhighsig, "/Users/User/Documents/PhD/Data/RNAseq data/medandhighsig1.xlsx")

```

#Heat map preparation

```

sig1<-countdata[countdata$SignificanceLOW=="*"]
sig2<-countdata[countdata$SignificanceMED=="*"]
sig3<-countdata[countdata$SignificanceHIGH=="*"]
temp1<-merge(sig2, sig3, by="Gene_ID")
temp2<-merge(temp1, sig3, by="Gene_ID")
sig<-temp2
write_xlsx(temp2,"C:/Users/User/Documents/PhD/Data/Ptk6RNAseq\\rko.only.significant.genes.xlsx")
view(rko.significant.genes)

```

#Heat map graph plot

```

heatmap.1<- sig[,c("Gene", "DMSO", "DMSO2", "DMSO3", "LOW", "MED", "HIGH")]
rownames(heatmap.1)<-heatmap.1$Gene
heatmap.1$Gene<- NULL #remove this column
as.matrix(heatmap.1)
heatmap(heatmap.1)
heatmap(as.matrix(heatmap.1, cexROW = 100, margins =
c(100,100,ylab=2,fontsize_row=fontsize_row)))
heatmap

```