Elucidating the role of the TGF beta superfamily in metastatic spread of colorectal cancer

Natthamon Sonthi

BSc, MSc

Thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

School of Cancer Sciences
College of Medical, Veterinary and Life Sciences
University of Glasgow

Sep 2025

Abstract

Colorectal cancer (CRC) is an aggressive disease and the leading cause of cancer death, characterised by high heterogeneity and various risk factors related to its etiology (1-3). During carcinogenesis, adenomatous polyps, which represent most premalignant lesions (85-90% of sporadic CRC), can develop into CRC (4-6). About 20-25% of CRC patients are diagnosed with metastatic disease, which is associated with poorer survival rates. CRC can spread to various tissues, including lymph nodes, liver, lungs, peritoneum, bones, and the central nervous system (7). The liver is the most common site for detecting metastatic CRC and is involved in 25% to 50% of cases. However, a rare form of CRC with bone metastasis occurs in 3% to 7% of patients; these individuals often have worse survival outcomes, with shorter survival times and limited treatment options (8-10). TGF-β and BMP signalling pathways are crucial mechanisms in tissue homeostasis, promoting cell proliferation and differentiation during crypt formation in the intestine (11-15). Dysregulation of these pathways in intestinal cells can impair their tumour-suppressing functions and facilitate tumour development (16-18). In particular, mutations in *TGFBR2*, *BMPR1A*, and *SMAD4* have been identified as contributing to CRC carcinogenesis (2, 3, 19-23). Patients with consensus molecular subtype (CMS) 4, categorised based on gene expression signature, are associated with prominent activation of TGF-β, along with stromal infiltration, epithelial to mesenchymal transition (EMT), and angiogenesis, leading to poorer survival outcomes (19).

In silico analysis of gene expression levels of TGF-β and BMP signalling components across CRC cell lines using the DepMap database, along with examination of key proteins and their phosphorylation statuses through Western blots, revealed that ligand activation occurs at various levels, with notable activation of upstream receptors such as TGFBR2, ACVR1B, and BMPR1A. Increased expression of signal transduction genes, including SMAD1, SMAD2, SMAD3, and SMAD5, was also observed. Interestingly, although phosphorylation of SMAD1/5/8 was detected in SW620 cell lines (a metastatic CRC cell line), SMAD4, a central element in signal transduction, was found to be downregulated. Apart from the differential expression of TGF-β and BMP signalling in CRC, the investigation of gene expression in these pathways within disease-free bone marrow (BM) cells using the Stemformatics database was also investigated to provide some insight into whether this morphogenic pathway is involved in metastatic spread to the bone.

Here we demonstrated high levels of TGF-ß and BMP signalling pathways in the bone marrow associated cells, indicating maintenance of homeostasis and serving as a baseline reference for further research.

Assessment of epithelial SMAD4 expression in colorectal polyps using immunohistochemistry (IHC) and digital weighted histoscoring with QuPath revealed that low SMAD4 levels in adenomatous polyps correlated with higher grades of dysplasia, different histological subtypes, the presence of metachronous polyps, and served as a prognostic marker. This marker indicated an increased risk of developing metachronous polyps, particularly in the tubulovillous polyp subtype. Additionally, transcriptomic analysis of tubulovillous polyps showed that upregulation of genes involved in protein deubiquitination occurs in polyps with low SMAD4 levels, along with a likely enrichment in tyrosine metabolism, PPAR signalling, arginine and proline metabolism, leukocyte transendothelial migration, and basal cell carcinoma. In CRC, lower epithelial SMAD4 expression was strongly associated with higher tumour stages and increased tumour stroma. Moreover, lower SMAD4 expression in CRC tumours had prognostic significance, predicting decreased cancer-specific survival in CRC patients, especially in right-sided tumours. A combination of tumour SMAD4 levels and stroma percentages suggested that the worst survival outcome was in patients with low SMAD4 expression in the tumour and high stroma content. Transcriptomic analysis also identified downregulation of *SOD3* and enrichment of aminopeptidase activity in this group.

Differential expression patterns at gene and protein levels, along with phosphorylation activity, were observed across the complex crosstalk of MAPK/ERK, WNT/ß-catenin, TGF-ß/BMP, and PI3K signalling pathways in CRC cell lines. However, the mechanisms regulating SMAD4 activity in CRC remain unclear. Testing the TGF-ß and BMP signalling inhibitor (LDN-212854) in combination with the standard chemotherapy (Fluoropyrimidine; 5-FU) showed synergistic effects on CRC cell viability, cell cycle arrest, cell proliferation, and cell recovery in an *in vitro* 2D study. Conditioned media from BM and hepatic cells influence changes in CRC behaviour. The 3D bioprinted SW620 spheroids in 2% alginate and 8% gelatin hydrogel supported physiological interactions, spheroid survival and growth, and were used for drug screening, demonstrating efficacy of the LDN-212854/5-FU combination in a more mechanophysical 3D system.

Further development of this *in vitro* 3D model to become multicell by incorporating metastatic CRC culture with BM and hepatic niches would enhance future CRC research and drug discovery.

Table of Contents

Abstract 2

Table of Contents 5

List of Tables 11

List of Figures 12

Publications in Preparation 15

Presentations 15

Acknowledgments 16

Author's Declaration 18

List of Abbreviations 19

Chapter 1 Introduction 25

1.1 Normal gut 25

1.1.1 Gut anatomy and function 25

1.1.2 Gut structure, cellular component and microenvironment 27

1.2 Colorectal polyps – Premalignant lesions of CRC 29

1.3 Colorectal cancer (CRC) 31

1.3.1 CRC incidence and mortality 31

1.3.2 CRC risk factors 32

1.3.3 Colorectal cancer screening 33

1.3.4 CRC staging system and classification 38

1.3.5 CRC pathogenesis and key mutations 53

1.4 Cell signalling pathways in the gut 56

1.4.1 Regulatory cell signalling pathways in healthy intestinal crypts 56

1.4.2 Dysregulation of cell signalling pathways in CRC development and progression 71

1.5 Metastasis of CRC 73

1.5.1 Epidemiology of Metastatic CRC 73

1.5.2 Sites of metastatic CRC 75

1.6 Therapeutic approaches for CRC 81
1.6.1 Surgery 81
1.6.2 Radiotherapy 82
1.6.3 Chemotherapy 82
1.6.4 Targeted therapy 83
1.6.5 Immunotherapy 84
1.7 TGF-β superfamily signalling in normal and cancer cells 85
1.7.1 TGF-β signalling members and their function 85
1.7.2 Dysregulation of TGF-β/BMP signalling in cancer 87
1.7.3 Inhibition of TGF-β/BMP signalling 89
1.8 Modelling in cancer research 91
1.8.1 *In vitro* models 92
1.8.2 *In vivo* models 93
1.8.3 Other models 94
1.8.4 Clinical trials and Precision medicine 95
1.9 The research aims and hypotheses 97
1.9.1 Research aims 97
1.9.2 Research hypotheses 97
Chapter 2 Materials and Methods 98
2.1 Materials 98
2.1.1 Open access gene expression databases 98
2.1.2 Patient cohorts 102
2.1.3 Cell lineage and tissue culture 103
2.1.4 Immunohistochemistry (IHC) 107
2.1.5 Flow cytometry (FACS) 108
2.1.6 Molecular biology 109
2.1.7 Western blotting 111
2.1.8 Spheroids culture 117
2.2 Methods 118
2.2.1 *in silico* analysis - Accessing online gene expression databases and Visualization 118
2.2.2 Patient tissue studies - Establishing prognostic significance 120
2.2.3 *In vitro* 2D cell line studies 126
2.2.4 Flow cytometry (FACS) 136

2.2.5 Molecular biology - RNA expression analysis 144
2.2.6 Western blotting - Protein expression analysis 145
2.2.7 *In vitro* 3D Spheroids studies 149
2.2.8 Statistical analysis.....155

Chapter 3 Evaluation of intrinsic expression patterns of the TGF-β and BMP signalling pathways in CRC cell lines and cells within the bone marrow. 157

3.1 Introduction 157

3.2 *In silico* analysis of TGF-β and BMP pathways in CRC cell lines using the gene expression database from the Cancer Dependency Map (DepMap) portal 158

3.2.1 Differential expression of TGF-β and BMP signalling was observed across CRC cell lines. 158

3.3 *In silico* analysis of TGF-β and BMP signalling pathways in disease-free bone marrow using Stemformatics 161

3.3.1 High expression of TGF-β and BMP signalling was observed across disease-free bone marrow 161

3.4 RNA and protein expression analysis of TGF-β and BMP signalling in CRC cell lines. 165

3.4.1 Downregulation of SMAD4 in metastatic CRC cell lines.....165

3.5 Discussion.....167

Chapter 4 SMAD4 expression in colorectal adenomatous polyps and its correlation with clinical outcomes 170

4.1 Introduction 170

4.2 SMAD4 antibody specificity in CRC cell lines.....171

4.3 Clinicopathological characteristics of the INCISE cohort.....173

4.4 Expression of SMAD4 in colorectal adenomatous polyps and its association with clinical outcomes 175

4.5 SMAD4 expression in tubulovillous polyps predicts the development of metachronous polyps and contributes to transcriptomic alterations. 183

4.5.1 Expression of SMAD4 in tubulovillous polyps and its association with clinical outcomes 183

4.5.2 Transcriptomic analysis of low and high SMAD4 expression in tubulovillous polyps 185

4.6 Discussion.....188

Chapter 5 SMAD4 expression in colorectal cancer and its correlation with clinical outcomes 191

5.1 Introduction.....191

5.2 Clinicopathological characteristics of the cancer cohorts.....192

5.2.1 Clinicopathological characteristics of the early CRC stages cohort .192

5.2.2 Clinicopathological characteristics of the GRI cohort 193

5.3 Expression of SMAD4 in CRC and its association with clinical outcomes and characteristics.....197

5.3.1 Association between SMAD4 expression, clinical characteristics, and clinical outcomes in the early CRC stages cohort 197

5.3.2 Association between SMAD4 expression, clinical characteristics, and clinical outcomes in the GRI cohort 200

5.4 SMAD4 expression in CRC and sites of tumour development predicts survival outcomes 208

5.5 SMAD4 expression in CRC tumours with high tumour stroma predicts survival outcomes and contributes to transcriptomic analysis.....210

5.5.1 Combining tumour SMAD4 expression levels with tumour stroma percentages in CRC and their association with survival prognosis.....210

5.5.2 Transcriptomic analysis of tumours with low and high SMAD4 expression involving high stroma in CRC.....213

5.6 Discussion.....216

Chapter 6 Signalling mechanism alterations related to SMAD4 levels in CRC cells, synergistic effects of TGF-β and BMP signalling inhibition in combination with standard chemotherapy in a metastatic CRC cell, and developing a model of metastasis 220

6.1 Introduction.....220

6.2 Alterations in signalling mechanisms involved in regulating SMAD4 activity.....221

6.2.1 Differential expression patterns of pathways which crosstalk with SMAD4 activity regulation were observed across CRC cell lines. 221
6.2.2 Investigation of crosstalk pathways in regulating SMAD4 expression through stimulation with growth factors such as BMP4 and EGF, as well as treatment with Bortezomib. 224
6.3 Inhibitory effects of BMPR1 inhibitors (Dor, K02288, and LDN-212854) on the TGF-β/BMP signalling pathways and a standard chemotherapeutic agent (5-FU) in CRC cell lines. 227
6.4 Targeting TGF-β/BMP signalling pathways both alone and in combination with chemotherapy. 229
6.4.1 Synergistic effects of the drug combination between LDN-212854 and 5-FU on cell viability. 229
6.4.2 Inhibition effects of the drug combination between LDN-212854 and 5-FU on protein levels of signalling mechanisms 235
6.4.3 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on apoptosis 238
6.4.4 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell cycle analysis 242
6.4.5 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell proliferation using CTV analysis. 245
6.4.6 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell recovery using clonogenic assay 248
6.5 Developing an *in vitro 3D* culture model of CRC metastasis 252
6.5.1 Establishing fluorescent CRC cell lines for 3D culture models 252
6.5.2 Cell viability of CRC cell lines in the conditioned media of BM and HCC cells 255
6.5.3 Inhibition effects of drug treatment of LDN212854 and 5-FU on the proliferation of CRC cells cultured in conditioned media 259
6.5.4 Formation of CRC spheroids on the Aggrewell™ plate 265
6.5.5 Cell viability of culturing SW620 spheroids in hydrogels over 7 days 272
6.5.6 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell death of SW620 spheroids in hydrogels on days 4 and 7 277
6.6 Discussion 287

Chapter 7 Discussion 292

7.1 Understanding the different expression patterns of the TGF-ß and BMP signalling pathways in CRC cell lines and cells within the bone marrow to identify dysregulation associated with CRC prognosis.....292

7.2 Low SMAD4 expression in colorectal adenomatous polyps predicted the development of metachronous polyps, particularly in tubulovillous polyps ..294

7.3 Low SMAD4 expression in colorectal cancer predicted worse survival outcomes, involving significant predictors in the right-sided colon and tumours with high stroma 296

7.4 Studying signalling mechanism alterations related to SMAD4 levels in CRC cells, revealing synergistic effects of TGF-ß and BMP signalling inhibition in combination with standard chemotherapy in a metastatic CRC cell, and developing a model of metastasis 298

7.5 Future direction 303

7.6 Conclusion.....304

List of References.....305

List of Tables

Table 2-1 Gene members in transforming growth factor-beta (TGF-β) superfamily 98
Table 2-2 Eligible microarray datasets consisting of bone marrow cells obtained from Stemformatics 99
Table 2-3 CRC cell line information 103
Table 2-4 Other cell line information 104
Table 2-5 Tissue culture plastics 104
Table 2-6 Tissue culture reagents 105
Table 2-7 Inhibitors, antibiotics and compounds 106
Table 2-8 IHC consumables 107
Table 2-9 FACS consumables 108
Table 2-10 1X wash/perm solution 108
Table 2-11 Molecular biology materials and reagents 109
Table 2-12 Primers 109
Table 2-13 Western blotting consumables 111
Table 2-14 Preparation of western blotting stock buffer solutions 112
Table 2-15 Preparation of protein lysis buffer (RIPA buffer) 113
Table 2-16 Preparation of 5X sample buffer 113
Table 2-17 Preparation of a 0.75 mm acrylamide gel for hand casting 113
Table 2-18 Preparation of Western blot working buffer solutions 114
Table 2-19 Antibodies 115
Table 2-20 Spheroids culture plastics 117
Table 2-21 Spheroids culture reagents 117
Table 2-22 QuPath thresholds for classifying epithelial SMAD4 staining 123
Table 2-23 FACS parameters for Annexin V apoptosis analysis 137
Table 2-24 FACS parameters for Ki-67 cell cycle analysis 139
Table 2-25 FACS parameters for CTV analysis 141
Table 2-26 BSA standard dilution 146
Table 4-1 The relationship of SMAD4 expression in colorectal adenomatous polyps and clinicopathological characteristics 182
Table 5-1 The relationship of SMAD4 expression in early-stage CRC and clinicopathological characteristics 200
Table 5-2 The relationship of SMAD4 expression in CRC stage I-III and clinicopathological characteristics from the GRI cohort 206

List of Figures

Figure 1-1 The large intestine in the gastrointestinal (GI) tract..... 27
Figure 1-2 Intestinal crypt in intestinal tissue..... 28
Figure 1-3 Classical adenoma and serrated lesions in the detection of colorectal polyps 30
Figure 1-4 Tumour, lymph node, metastasis (TNM) system for staging colorectal cancer..... 39
Figure 1-5 Clinical and molecular differences between right- and left-sided colorectal cancers 41
Figure 1-6 Pathogenesis and key mutations in colorectal cancer 56
Figure 1-7 Overview of cell fate and signalling pathways in the regulation of intestinal crypts 58
Figure 1-8 Classical Wnt signalling pathway 59
Figure 1-9 Classical Notch signalling pathway 60
Figure 1-10 Classical EGF signalling pathway 62
Figure 1-11 Classical BMP signalling pathway as a part of TGF-β signalling pathway 63
Figure 1-12 Classical Hedgehog signalling pathway..... 65
Figure 1-13 Classical EPH/Ephrin signalling pathway..... 66
Figure 1-14 Other signalling pathways, including Hippo, MAPKs, PI3K-AKT, JAK-STAT, and NF-κB signalling. 71
Figure 1-15 Metastatic cascade from the primary tumour to distant sites..... 74
Figure 1-16 Hepatic microenvironment as a target for metastatic CRC..... 76
Figure 1-17 Peritoneal microenvironment as a target for metastatic CRC..... 78
Figure 1-18 Lung microenvironment as a target for metastatic CRC..... 79
Figure 1-19 Bone microenvironment as a target for metastatic CRC 81
Figure 1-20 TGF-β and BMP signalling pathways and targeted inhibition 85
Figure 1-21 Overview of modelling in cancer research 92
Figure 2-1 Workflow for bioinformatic analysis of TGF-β and BMP signalling pathways in colorectal cancer (CRC) and bone marrow (BM) niche.....118
Figure 2-2 Workflow for determining the prognostic significance of SMAD4 expression in pre-malignant and CRC patient tissues 120
Figure 2-3 Image processing for digital pathology scoring of IHC staining on tissue microarrays (TMAs) using QuPath.....123
Figure 2-4 Assessment of cell viability using the Trypan blue exclusion assay ..128
Figure 2-5 Preparation of drug serial dilutions 129
Figure 2-6 Assessment of cell proliferation using the WST-1 assay 130
Figure 2-7 Conditions for the drug combination study applying the Chou-Talalay method 131
Figure 2-8 Assessment of cell proliferation and recovery through the clonogenic assay 133
Figure 2-9 Establish fluorescent CRC cell lines expressing EGFP and BFP.....135
Figure 2-10 Conditions for studying conditioned media from different cell types 136
Figure 2-11 Gating strategy for Annexin V apoptosis analysis.....138
Figure 2-12 Gating strategy for Ki-67 cell cycle analysis 140
Figure 2-13 Gating strategy for CTV analysis 142
Figure 2-14 Gating strategy for sorting EGFP and BFP cells 143
Figure 2-15 Preparation of protein standards and samples for protein quantification.....147

Figure 2-16 Order of assembly for semi-dry and wet transfer in western blotting 148
Figure 2-17 Workflow for developing an *in vitro* 3D culture model for drug screening in metastatic colorectal cancer 149
Figure 2-18 Image processing for spheroid analysis using Fiji.....155
Figure 3-1 Heatmaps of the TGF-ß superfamily in colorectal cancer cell lines .160
Figure 3-2 Heatmaps of the TGF-ß superfamily in disease-free BM cell types ..164
Figure 3-3 RNA and protein expression of the TGF-ß and BMP signalling pathways in colorectal cancer cell lines.....167
Figure 4-1 Specificity of SMAD4 antibody in CRC cell lines 173
Figure 4-2 Colorectal polyps and clinicopathological features of the INCISE cohort 175
Figure 4-3 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in colorectal adenomatous polyps 178
Figure 4-4 SMAD4 expression in relation to the detection of metachronous polyps or CRC development and clinical outcomes 182
Figure 4-5 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal adenomatous polyps, stratified by histological subtypes 185
Figure 4-6 Transcriptomic analysis comparing low and high SMAD4 expression in tubulovillous polyps.....188
Figure 5-1 Colorectal cancer and clinicopathological features of the early-stage CRC cohort 193
Figure 5-2 Colorectal cancer and clinicopathological features of the GRI cohort 196
Figure 5-3 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in the early-stage CRC Cohort 199
Figure 5-4 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in CRC stage I-III from the GRI cohort 203
Figure 5-5 SMAD4 expression in CRC predicting survival outcomes and its correlation with clinical characteristics.....206
Figure 5-6 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal cancer, stratified by tumour location 210
Figure 5-7 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal cancer, stratified by tumour stroma percentage (TSP) 213
Figure 5-8 Transcriptomic analysis comparing low and high SMAD4 expression in CRC tumours with high stroma.....215
Figure 6-1 Crosstalk pathways of TGF-ß/BMP signalling, Wnt signalling, and MAPK/ERK signalling involving SMAD4 activity regulation in CRC cell lines.....224
Figure 6-2 Effects of BMP4 and EGF stimulation and Bortezomib treatment on protein levels in crosstalk pathways regulating SMAD4 activity in DLD1 and SW620 cell lines.....227
Figure 6-3 Inhibitory effects on cell viability of BMPR1 inhibitors (Dor, K02288, and LDN-212854) and a standard chemotherapeutic agent (5-FU) in DLD1 and SW620 cell lines.....229
Figure 6-4 Inhibitory effects on cell viability of LDN-212854 and 5-FU in DLD1 cell line 230
Figure 6-5 Inhibitory effects on cell viability of LDN-212854 and 5-FU in SW620 cell line 232

Figure 6-6 Synergistic effects of the drug combination between LDN-212854 and 5-FU on cell viability of DLD1 and SW620 cell lines 234
Figure 6-7 Inhibition effects of the drug combination between LDN-212854 and 5-FU on protein levels of signalling mechanisms 237
Figure 6-8 Inhibition effects of the drug combination between LDN-212854 and 5-FU on apoptosis analysis of the SW620 cell line.....242
Figure 6-9 Inhibition effects of the drug combination between LDN-212854 and 5-FU on cell cycle analysis of the SW620 cell line.....244
Figure 6-10 Inhibition effects of the drug combination between LDN-212854 and 5-FU on CTV analysis of the SW620 cell line 247
Figure 6-11 Optimisation of seeding density and culture conditions for SW620 cell line in Clonogenic assay.....250
Figure 6-12 Inhibition effects of the drug combination between LDN-212854 and 5-FU on cell recovery of the SW620 cell line using Clonogenic assay.....251
Figure 6-13 Establishing fluorescent CRC cell lines (SW620 and DLD1) for multicell 3D modelling 255
Figure 6-14 Effects on cell proliferation of culturing CRC cell lines in the conditioned media from BM and HCC cells.....259
Figure 6-15 Inhibitory effects on cell viability of LDN-212854 and 5-FU treatment in DLD1 cells, which were cultured in 20% conditioned media from BM and HCC cells 262
Figure 6-16 Inhibitory effects on cell viability of LDN-212854 and 5-FU treatment in SW620 cells, which were cultured in 20% conditioned media from BM and HCC cells 265
Figure 6-17 DLD1 spheroid formation during culture on an Aggrewell™ plate ..268
Figure 6-18 SW620 spheroid formation during culture on an Aggrewell™ plate.271
Figure 6-19 Cell viability of SW620 spheroids within hydrogels over 7 days 276
Figure 6-20 Inhibitory effects on cell viability of SW620 spheroids in hydrogels with LDN-212854 and 5-FU treatment on day 4 281
Figure 6-21 Inhibitory effects on cell viability of SW620 spheroids in hydrogels with LDN-212854 and 5-FU treatment on day 7 286
Figure 7-1 A schematic overview of the key findings from this thesis in elucidating the role of the TGF beta superfamily in metastatic spread of colorectal cancer 302

Publications in Preparation

Natthamon Sonthi, Sara Samir Foad Al Badran, Kathryn Pennel, Lauren Hope, Gerard Lynch, Aula Ammar, Jean Quinn, Heather Jorgensen, Noori Maka, Stephen McSorley, INCISE collaborative, Joanne Edwards, Helen Wheadon. **"The role of the TGF beta superfamily in colorectal cancer development and progression"**

Presentations

2025 **"The role of the TGF beta superfamily in colorectal cancer development and progression".** Poster at the CRCL 6th International Cancer Symposium, Cancer Research Center of Lyon, Lyon, France

2024 **"Elucidating the role of the TGF beta superfamily in colorectal cancer development and progression"**. Poster at CRUK Scotland-Colorectal Cancer Annual Theme Day Symposium, The Institute of Genetics and Cancer, Lecture Theatre, Edinburgh, United Kingdom

Acknowledgments

Firstly, I would like to sincerely thank my supervisors: Professor Helen Wheadon, Professor Joanne Edwards, and Dr Heather Jørgensen, for the opportunity to undertake this project and for their guidance and support throughout my PhD. I am immensely grateful to have had Helen as my primary supervisor. With her exceptional thoughtfulness, patience, and empathy, she consistently supported and mentored me, greatly helping me to improve my knowledge and integrity over the four years. I am also very thankful to Joanne as my secondary supervisor. She is always kind, knowledgeable, and encourages me to produce excellent work. I am deeply appreciative of Heather, who has always been generous and supportive, and has also created an enjoyable environment at our special events.

I would like to express my sincere gratitude to Dr Jean Quinn, my reviewer and lab manager in Professor Edwards lab at Wolfson Wohl. Her instructions in the lab and support for all my needs have been significant to me. I also wish to thank Dr Gillian Horne, another reviewer, for her valuable feedback and helpful guidance during my Annual Progress Reviews.

Thank you very much to everyone, past and present, in Professor Edwards' lab. There has always been a warm and welcoming environment there. Special thanks to my best friends, Dr Kathryn Pennel, Dr Phimmada Hatthakarnkul, and Sara Samir Foad Al Badran, for their friendship and all support with lab work during the early stages of my PhD. Huge thanks to my kind friends, Dr Amna Ahmed Mohemmd Matly, Dr Emma Parsons, Dr Guang-Yu Lian, and Molly McKenzie, for their helpful assistance. Thank you to Dr Aula Ammar and Dr Gerard Lynch for their insightful advice. I couldn't have done it without Leonor Patricia Schubert Santana and Dr Natalie Fisher for their support with bioinformatic analysis. My thanks also go to the Glasgow Tissue Research Facility (GTRF) team and the patients for providing valuable colorectal tissues for this study.

I am delighted to have been part of the Wheadon lab at the Paul O'Gorman Leukaemia Research Centre (POG) since my second year of study. I am very grateful to everyone, past and present, in the lab for their support during my PhD. I also appreciate the assistance from Professor Mhairi Copland, Professor Alison Michie, Dr Alan Hair, Jennifer Cassels, Angela Newlands, and Diane Verrecchia

while I was working in the lab. Many thanks to my lovely friends Wenjing Huang, Isla Nosratzadeh, Dr Lauren Hope, Xenia Bubnova, Dr Aikaterini Miari, Dr Shaun Patterson, Dr Ahmed Bahaaeldin Nabih Elmarghany, Tom Benson, Sean Cooke, Valdemar Rodrigues Augusto Tchipenhe, Hager Salah Ali Abouelnaga, Jiatian Li, Dr Jodie Hay, Dr Jamie Lees, Dr Hassan Nasser B Almuhanna, Jin Tongsom, Dr Amy Muir, and Dr Rachel Craig for their friendship, help in the lab, and for creating a joyful environment.

I would like to thank Dr George Skalka from the Cancer Research UK Scotland Institute for providing a hepatic cell line, as well as Professor Manuel Salmeron-Sanchez and others for their support during my work on 3D bio-printing at the Advanced Research Centre (ARC).

A heartfelt thank you to the scholarship from Chulabhorn Royal Academy, Thailand, and the Scholar Team from the Office of Educational Affairs at the Royal Thai Embassy for their funding and support throughout my PhD journey. Special thanks to Professor Chirayu Auewarakul and Associate Professor Danupon Nantajit for providing a great opportunity for my PhD study.

Finally, I am truly grateful to my friends and family for their support throughout my journey. My mother and father, Kanlaya Sonthi and Sorayut Sonthi, always support me wholeheartedly, and I hope they feel proud of me. Additionally, I would like to thank my sister, Kewalin Sonthi, for her love and support all along. Thanks to my great friends, Nisabhat Tonwoot, Noppakoaw Raungsombut, Nanthakarn Wannago, Watcharakorn Riaboi, Nanthapong Sornkaew, Dr Warapan Numprasit, Walaiphorn Woraharn, Thanatpornpat Kritalug, Hathaipat Sukjariangporn and Krittaya Aksonnam, who shared many wonderful moments. Lastly, thank you to Zizou, my neighbourhood cat, who fills my heart with happiness while I write this thesis.

Author’s Declaration

I declare that all work included in this thesis was carried out by myself, except for where explicitly stated. No part of this work has been submitted for any other qualification at any educational institution.

Natthamon Sonthi

September 2025

List of Abbreviations

5-FU	5-Fluorouracil
7-AAD	7-Aminoactinomycin D
ACTB	Actin Beta
Alg	Alginate
ALK	Activin Receptor-Like Kinase
AMH	Anti-Müllerian Hormone
APC	Adenomatous Polyposis Coli
APN	Aminopeptidase N
APS	Ammonium Persulfate
B2M	Beta-2-Microglobulin
BCs	Band Cells
BFP	Blue Fluorescent Protein
BM	Bone Marrow
BME	Basement Membrane Extract
BMP	Bone Morphogenetic Protein
BMPR	BMP Receptor
BRAF	B-Raf Proto-Oncogene, Serine/Threonine Kinase
BSA	Bovine Serum Albumin
BSG	British Society of Gastroenterology
c-Myc	Cellular Myelocytomatosis Oncogene
CAFs	Cancer-Associated Fibroblasts
Calcein AM	Calcein Acetoxymethyl
CBCs	Crypt Base Columnar Cells
CI	Combination Index
CIMP	CpG Island Methylator Phenotype
CIN	Chromosomal Instability
CMP	Common Myeloid Progenitor
CMS	Consensus Molecular Subtypes
co-SMAD	Common-Mediator SMAD
CRC	Colorectal Cancer
CRIS	Colorectal Cancer Intrinsic Subtypes
CRLM	Colorectal Liver Metastasis
CSS	Cancer-Specific Survival
CTCF	Corrected Total Cell Fluorescence

CTNNB1	Catenin, Beta 1
CTV	CellTrace™ Violet
DAPI	4′,6-Diamidino-2-Phenylindole
DepMap	Dependency Map
DFS	Disease-Free Survival
dMMR	Deficient DNA Mismatch Repair
DMSO	Dimethyl Sulfoxide
Dor	Dorsomorphin
ECM	Extracellular Matrix
EGF	Epidermal Growth Factor
EGFP	Enhanced Green Fluorescent Protein
EGFR	EGF receptor
EMT	Epithelial-to-Mesenchymal Transition
EPCs	Erythropoietic Cells
ERK1/2	Extracellular signal-regulated kinase 1 and 2
EthD-1	Ethidium Homodimer-1
FACS	Fluorescence-Activated Cell Sorting
FAP	Familial Adenomatous Polyposis Coli
FBS	Fetal Bovine Serum
FDR	False Discovery Rate
FFPE	Formalin Fixed Paraffin-Embedded
FITC	Fluorescein Isothiocyanate
FSC	Forward Scatter
GC	Granulocyte
GDF	Growth and Differentiation Factor
Gel	Gelatin
GF	Growth Factors
GI	Gastrointestinal
GMPs	Granulocyte-Macrophage Progenitor Cells
GMS	Glasgow Microenvironment Score
GO	Gene Ontology
GRI	Glasgow Royal Infirmary
GSEA	Gene Set Enrichment Analysis
GSK-3β	Glycogen Synthase Kinase 3 beta
GTRF	Glasgow Tissue Research Facility
H&E	Haematoxylin and Eosin
H2AX	Histone H2A, Variant, Family Member X

hBMP4	Human Bone Morphogenetic Protein 4
HBSS	Hanks' Balanced Salt Solution
HCC	Hepatocellular Carcinoma
HER2	Human EGFR2
Hh	Hedgehog
HNPCC	Hereditary Nonpolyposis Colorectal Cancer
HP	Hyperplastic Polyps
HPCs	Haematopoietic Precursor Cells
HR	Hazard Ratio
HRP	Horseradish Peroxidase
HSCs	Haematopoietic Stem Cells
I-SMADs	Inhibitory SMADs
IBD	Inflammatory Bowel Disease
IC50	Inhibitory Concentration of 50%
ICC	Interclass Correlation Coefficient
iCMS	Intrinsic CMS
IFN-γ	Interferon-γ
IGF2	Insulin-Like Growth Factor 2
IHC	Immunohistochemistry
IKK	I-KappaB Kinase
IL	Interleukin
IMF	Intrinsic Epithelial Subtype, MSI Status, and Fibrosis
INCISE	Improved Polyp Surveillance
ISCs	Intestinal Stem Cells
ITS-G	Insulin-Transferrin-Selenium
JAG1/2	Jagged-1 and 2
JAK/STAT	Janus Kinase/Signal Transducer and Activator of Transcription
JNK	c-Jun NH2-Terminal Kinase
JPS	Juvenile Polyposis Syndrome
KEGG	Kyoto Encyclopedia of Genes and Genomes
KM	Kaplan–Meier
KMG	Klintrup–Mäkinen Grade
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LAP	Latency-Associated Peptide
LC	Leukocyte
LCCs	Left-Sided Colon Carcinomas
LGR5	Leucine-Rich Repeat-Containing G Protein-Coupled Receptors 5

LOH	Loss of Heterozygosity
LPS	Lipopolysaccharide
LTBP	Latent TGF-ß Binding Protein
mAb	Monoclonal Antibody
MAPCs	Multipotent Adult Progenitor Cells
MAPK	Mitogen-Activated Protein Kinase
MEPs	Megakaryocyte-Erythroid Progenitor Cells
MFI	Mean Fluorescence Intensity
mGPS	Modified Glasgow Prognostic Score
MM	Metamyelocyte
MMR	Mismatch Repair
MO	Macrophage
Mono	Monocyte
MPPs	Haematopoietic Multipotent Progenitor Cells
MSCDI	MSC Differentiation Intermediate
MSCs	Mesenchymal Stromal Cells
MSI	Microsatellite Instability
MSS	Microsatellite Stability
mTOR	Mechanistic Target of Rapamycin
MY	Myelocyte
NES	Normalised Enrichment Score
NF-κB	Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cells
NK	Natural Killer
OB	Osteoblast
OD	Optical Density
ODI	Osteogenic Differentiation Intermediate
OS	Overall Survival
PARP-1	Poly(ADP-ribose) Polymerase-1
PB	Polybrene
PBS	Phosphate Buffered Saline
PC	Pericyte
PDOs	Patient-Derived Organoids
PDS	Pathway-Derived Subtypes
PFS	Progression-Free Survival
PI	Petersen Index
PI	Prognostic Index
PI	Propidium Iodide

PI3K	Phosphatidylinositol 3-Kinase
PIAS	Protein Inhibitor of Activated STAT
PIK3CA	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase Catalytic Subunit Alpha
PM	Promyelocyte
pMMR	MMR-Proficient
PMNs	Polymorphonuclear Cells
PNI	Perineural Invasion
POLD1	Polymerase (DNA) Delta 1, Catalytic Subunit
POLE	Polymerase (DNA) Epsilon, Catalytic Subunit
PS	Phosphatidylserine
qPCR	Quantitative Real-time Polymerase Chain Reaction
QuPath	Quantitative Pathology & Bioimage Analysis
R-SMADs	Receptor-Regulated SMADs
RCCs	Right-Sided Colon Carcinomas
RFS	Relapse-Free Survival
RTK	Receptor Tyrosine Kinase
SBEs	SMAD-Binding Elements
SBoSP	Scottish Bowel Screening Programme
SCNA	Somatic Copy Number Alterations
SCs	Stromal Cells
SDS	Sodium Dodecyl Sulfate
SH-PTP2	Src Homology Region 2 Domain-Containing Phosphatase-2
SMAD	Small Mother Against Decapentaplegic
SMURF1/2	SMAD Ubiquitination Regulatory Factors 1 and 2
SOD3	Superoxide Dismutase 3
SREs	Skeletal-Related Events
SSC	Side Scatter
SSL	Sessile Serrated Lesions
TB	Tumour Budding
TBS	Tris-Buffered Saline
TBSN	TBS with Nonidet P-40
TBST	TBS with Tween 20
TCGA	The Cancer Genome Atlas
TGF-β	Transforming Growth Factor Beta
TMA	Tissue MicroArray
TME	Tumour Microenvironment

TMS	Tumour Microenvironment Score
TNF	Tumour Necrosis Factor
TNM	Tumour, Lymph Node, Metastasis
TP53	Tumour Protein p53
TPM	Transcription Per Million
TSA	Traditional Serrated Adenoma
TSP	Tumour Stromal Percentages
TxRed	Texas Red
WHS	Weighted Histoscore
WNT1	Wingless and INT1
WST	Water-Soluble Tetrazolium
WT	Wild Type
ZEB1/2	Zinc Finger E-Box Binding Homeobox 1 and 2
β-TrCP	Beta-Transducin Repeat-Containing Protein

Chapter 1 Introduction

1.1 Normal gut

1.1.1 Gut anatomy and function

The gut, often referred to as the gastrointestinal (GI) tract, is an essential part of the human body's system for obtaining nutrition and supplying substrates for cellular energy metabolism. This system begins with the ingestion and mastication of food by the mouth and teeth, followed by the transport of the ingested food (bolus) through peristalsis in the oesophagus. The bolus passes through the upper oesophageal sphincter (UES) and the lower oesophageal sphincter (LES), which prevent gastric acid reflux, before entering the stomach. The stomach is composed of the cardia (oesophagogastric junction), fundus (proximal part), corpus (middle part), antrum (distal part), and pylorus (duodenal sphincter). In the stomach, the ingested food is temporarily stored along with gastric acid to prepare for digestion and minimise bacterial contamination. The acidic gastric content is then delivered to the small intestine (24).

The duodenum, jejunum, and ileum are three sections of the small bowel. In the duodenum, the papilla of Vater is a rounded projection of the bile duct (biliary) joining the pancreatic duct, which releases an alkaline bile/pancreatic juice mixture to neutralise the acidic gastric content. Subsequently, the preprocessed intestinal content is digested in the jejunum and ileum. All nutritional components, including amino acids, carbohydrates, lipids, and vitamins, are absorbed and transported via the portal vein to the liver. The liver receives a dual blood supply from the hepatic artery and portal vein, delivering oxygenated blood from systemic circulation and nutrient-rich, deoxygenated blood from digestive organs. The blood from both vessels mixes in sinusoids, circulates through hepatocytes, and flows out through the hepatic veins. The liver not only produces bile and involves metabolism but also synthesises and stores glycogen, protein, and vitamins. The pancreas is an auxiliary organ that produces pancreatic juices containing various enzymes involved in carbohydrate, fat, and protein digestion, as well as secreting insulin for glucose metabolism (24, 25).

The last part of the alimentary tract is the large bowel. The indigestible mass passes through the ileocolic valve, caecum, and vermiform appendix, followed by the colon (ascending, transverse, descending, and sigmoid colon), rectum, and anal canal with the anal sphincter. This process enables the resorption of remaining nutrients and vitamins processed by colonic bacteria, as well as the removal of water to decrease the mass of the faeces before defecation. The length of the colon is approximately 1.5 metres shorter than that of the small intestine (6-8 metres). Three distinct longitudinal muscular bands (taenia coli) extend from the caecum to the rectum, distinguishing it from the small intestine. Typical pouches (haustra) are sacculated between the longitudinal bands. The hepatic flexure and splenic flexure are angulated between the ascending colon, transverse colon, and descending colon. Colorectal sidedness is defined by the right-sided colon (caecum, ascending colon, and proximal two-thirds of the transverse colon), left-sided colon (one-third of the transverse colon to sigmoid colon), and rectum, Figure 1-1. The superior mesenteric artery and vein supply the right-sided colon, while the inferior mesenteric artery and vein supply the remaining left-sided colon and rectum. Similar to the blood supply pattern, the superior and inferior mesenteric plexuses of the nerve and lymphatic system supply the large intestine (24-27).

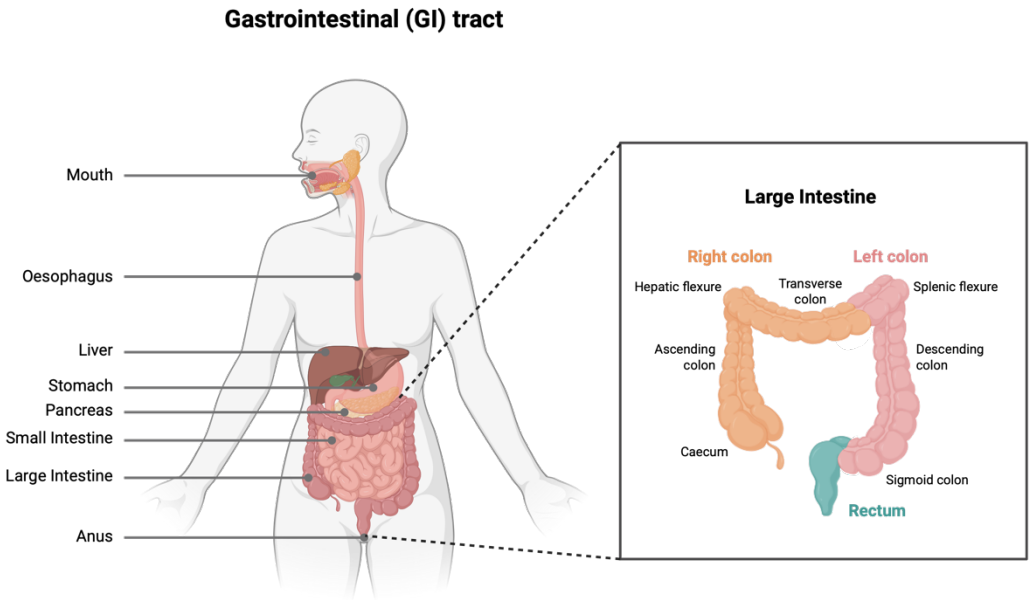


Figure 1-1 The large intestine in the gastrointestinal (GI) tract

The diagram of the GI tract illustrates the organs of the gut, including the mouth, oesophagus, liver, stomach, pancreas, small intestine, large intestine, and anus. The large intestine is divided into three sections: the right colon (from the caecum, ascending colon, hepatic flexure, and the proximal two-thirds of the transverse colon), the left colon (one-third of the transverse colon, splenic flexure, descending colon, and sigmoid colon), and the rectum. This figure was created using BioRender.com and adapted from Schneider A. and Feussner H., 2017 and Dekker E. et al., 2019.

1.1.2 Gut structure, cellular component and microenvironment

The large intestinal tissue consists of four layers: mucosa, submucosa, muscularis mucosa (thick longitudinal muscle layers), and serosa (adventitia). The mucosa is divided into epithelium, connective tissue (lamina propria), and muscularis mucosa (thin circular muscle layer) (25). The intestinal crypt is a niche characterised by a population of stem cells located at the base of a finger-like structure (villus), formed by a single epithelial layer with underlying connective tissue. Millions of villi and crypts line the intestinal mucosa and extend toward the intestinal lumen. Intestinal stem cells (ISCs) at the basal crypts exhibit multipotency, allowing them to differentiate into all tissue-specific epithelial cells and self-renew during tissue homeostasis, Figure 1-2. These cells can be categorised into slow-cycling stem cells (quiescent and reserve stem cells), which express markers of the polycomb complex protein Bmi1, homeodomain-only protein X (Hopx), and mouse telomerase reverse transcriptase (Tert), and actively cycling stem cells (crypt base columnar cells, CBCs), which express markers of leucine-rich repeat-containing G protein-coupled receptor 5 (Lgr5) and proliferate into progenitor cells. Early progenitor precursors include absorptive progenitors, which differentiate into enterocytes, and secretory progenitors, which differentiate into Paneth, goblet, enteroendocrine, and tuft cells. The differentiated cells migrate from the crypt-villus junction toward the apical villi and eventually shed into the intestinal lumen. However, Paneth cells remain in the stem cell region to secrete essential signals such as EGF and Wnt for ISC maintenance. Additionally, the mesenchyme surrounding the crypt base, also known as pericryptal myofibroblasts, supports the ISC state by producing important factors such as Wnt ligands, Wnt agonists R-spondins (Rspos), and BMP antagonists (12, 14, 28).

The other relevant component of the intestinal microenvironment is the microbiota. With numerous advantages for the host, intestinal bacteria are essential for maintaining gut homeostasis, regulating immune responses and metabolism, and protecting against pathogens. The density and composition of microbiota vary significantly, influenced by various factors, including nationality, dietary choices, medical treatments, ageing, and the specific regions of the gastrointestinal tract. Dense bacterial populations are predominantly found in the colon. However, variations among the three core enterotypes (*Bacteroides*, *Prevotella*, and *Ruminococcus*) are observed across different sample populations. Alterations in microbial composition, known as dysbiosis, may lead to pathogenesis, such as inflammatory bowel disease and cancer (29).

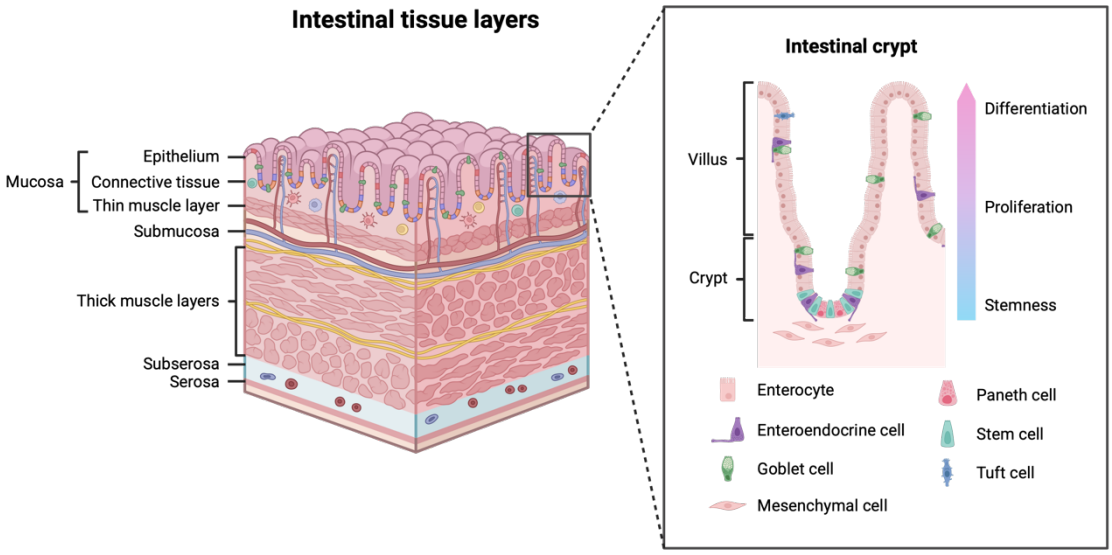


Figure 1-2 Intestinal crypt in intestinal tissue

The diagram of intestinal tissue illustrates four layers: mucosa (including epithelium, connective tissue, and a thin muscle layer), submucosa, muscularis mucosa (thick longitudinal muscle layers), and serosa (adventitia). Intestinal crypts are structures composed of a single epithelial layer with underlying connective tissue, lined by the intestinal mucosa, and form finger-like projections (villi) that extend towards the intestinal lumen. The niche of the intestinal crypts contains a population of stem cells located at the base of the villi, surrounded by mesenchymal cells. Intestinal stem cells (ISCs) have the ability to proliferate and differentiate into all tissue-specific cells along the villi, including enterocytes, enteroendocrine cells, goblet cells, Paneth cells, and tuft cells. This figure was created using BioRender.com and adapted from Spit M. et al., 2018 and Malijauskaite S. et al., 2021.

1.2 Colorectal polyps – Premalignant lesions of CRC

Colorectal polyps are significant precursors to tumourigenesis, as these benign dysplastic lesions can progress to CRC (5, 6). Neoplastic polyps in the intestinal mucosa are believed to develop through two primary premalignant pathways: the adenoma-carcinoma pathway (85-90% of sporadic CRC) and the serrated neoplasia pathway (10-15% of sporadic CRC), both of which involve several molecular alterations (4). Histopathological features describe the characteristics of polyps. Classical adenomas, the most common premalignant lesions, can exist in either polypoid or non-polypoid forms. Visible polypoid adenomas are detected by colonoscopy (Figure 1-3), while non-polypoid adenomas are rarely identified due to their small size and resemblance to flat mucosal discolouration. The levels of glandular epithelium proliferation in adenomas can be classified into two dysplasia grades: low and high. Low-grade dysplasia is characterised by cytological atypia of the colonic mucosa, while architectural changes indicate high-grade dysplasia. Furthermore, adenoma classification based on appearance is divided into tubular, tubulovillous, and villous adenomas. The tubular adenoma is the most common type (65–80% of all polyps removed), featuring a smooth surface resembling normal tissue with parallel crypts. A tubulovillous adenoma displays transitional features of both tubular and villous adenomas (accounting for 10–25%). A villous adenoma (5–10%) is characterised by finger-like projections and severe atypia or dysplasia, with each branch or peduncle containing a fibrovascular core or stalk extending from the mucosa (4, 30).

Serrated lesions are a key characteristic of polyps. A group of heterogeneous polyps, featuring a sawtooth-like pattern and a flat, broad base of sessile morphology, complicates detection. Hyperplastic polyps (HP), typically smaller than 5 mm, are often found in the distal colon and rectum. Sessile serrated lesions (SSL), moderately sized between 5 and 10 mm with minor cytological atypia and complex abnormal architecture, are usually located in the proximal colon. Traditional serrated adenoma (TSA), a larger dysplastic lesion greater than 10 mm with a stalk, is commonly found in the distal colon (4).

Additionally, hereditary syndromes can potentially lead to colorectal cancer. Familial adenomatous polyposis coli (FAP) is associated with the formation of hundreds of adenomas in the colon and rectum. While hamartomas can develop

from both sporadic and inherited polyposis syndromes- such as juvenile polyposis syndrome (JPS), Peutz-Jeghers syndrome (PJS), and Cowden syndrome- they are characterised by disorganised tumour-like tissue growths. These polyposis syndromes may pose a risk of CRC as well as extraintestinal manifestations. Prolonged inflammatory processes related to chronic Inflammatory bowel disease (IBD), such as ulcerative colitis (UC), may create polypoid lesions that can transform into elevated dysplasia, thereby increasing the risk of adenocarcinoma (4).

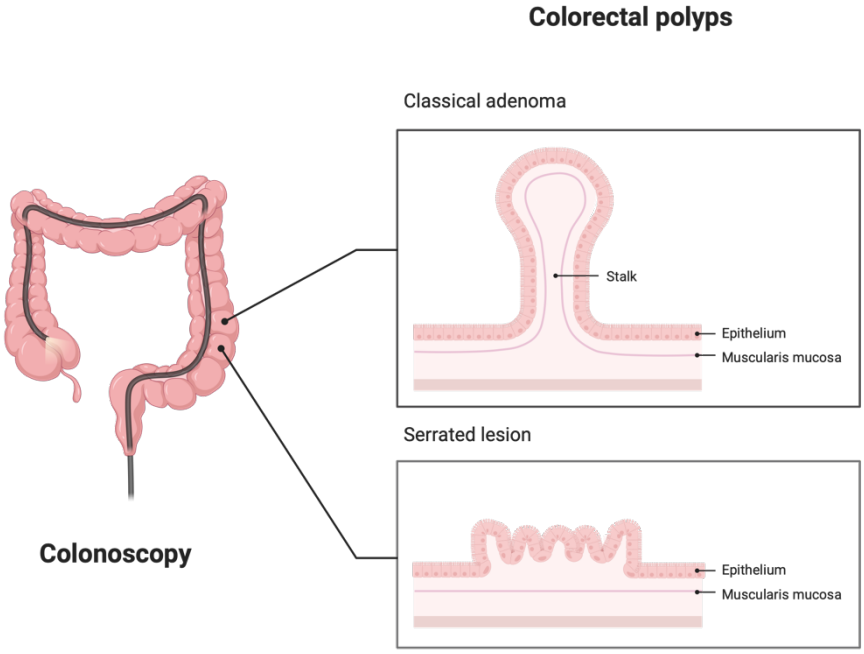


Figure 1-3 Classical adenoma and serrated lesions in the detection of colorectal polyps

The diagram illustrates the use of colonoscopy to detect colorectal polyps, which are characterised as classical adenomas (e.g., a finger-like projection with a stalk) and serrated lesions (e.g., a sawtooth-like pattern). This figure was created using BioRender.com and adapted from Haggitt RC. et al., 1985.

1.3 Colorectal cancer (CRC)

1.3.1 CRC incidence and mortality

According to the Global Cancer Statistics 2022, colorectal cancer (CRC) was the third-leading cancer globally, accounting for 9.6% of cases (over 1.9 million new cases/year), following lung and breast cancer, as estimated by the International Agency for Research on Cancer (IARC). CRC was recognised as the second leading cause of cancer death, at 9.3% (904,000 deaths/year), behind lung cancer. CRC is diagnosed more frequently in men than in women; however, the mortality rate is similar for both sexes. The incidence and mortality rates of CRC is 3 to 4 times higher in very high/high Human Development Index (HDI) populations compared to medium/low HDI populations. In Europe, Australia/New Zealand, and Northern America, there is a high incidence rate among both men and women. Although there has been a steady or declining trend in incidence for all ages combined over the last few decades due to the introduction of screening, better awareness and individual health monitoring, a rising incidence in younger adults aged under 50 years has been reported (31).

In the United Kingdom, bowel cancer was the fourth most commonly diagnosed cancer and the second leading cause of cancer deaths between 2017 and 2019, with an average of 44,100 new cases annually and 16,800 fatalities. Between 2013 and 2017, more than 50% of bowel cancer patients, both men and women in England, were predicted to survive for at least ten years. Patients diagnosed with CRC between the ages of 15 and 44 survive longer with approximately 63% surviving at least ten years, In contrast, only 39.4% of CRC patients diagnosed at age 75 and older are predicted to have a ten-year survival rate (32). In Scotland, data reported from the Scottish Cancer Registry (SCR) between 1997 and 2017 highlighted a declining incidence of CRC in the screening group aged over 50 years, from 156.5 cases per 100,000 in 2010 to 123.9 per 100,000 in 2017. Meanwhile, there was an increase in diagnoses among younger people, rising from 5.3 to 6.8 cases per 100,000 between 2000 and 2017 (33).

1.3.2 CRC risk factors

Colorectal cancer (CRC) is a disease that arises in the epithelial cells of the large intestine and is influenced by several risk factors, including previous family predisposition of developing CRC and personal medical history, lifestyle, and non-modifiable factors. First, family and personal medical history significantly correlate with an increased risk of colorectal carcinogenesis, contributing to a 2-8% incidence of CRC in individuals with an inherited genetic predisposition. The hereditary syndromes frequently identified include hereditary nonpolyposis colorectal cancer (HNPCC), also known as Lynch syndrome, and familial adenomatous polyposis coli. Both syndromes are autosomal dominant disorders caused by mismatch repair (MMR) mutations, primarily in DNA MMR genes, including *MLH1*, *MSH2*, *MSH6*, and *PMS2* for HNPCC, as well as defects in the adenomatous polyposis coli (APC) gene, leading to the dysregulation of DNA replication and cell division, which results in the development of multiple colon polyps. Additionally, Peutz-Jeghers syndrome (PJS), juvenile polyposis syndrome (JPS), PTEN hamartoma tumour syndrome (PHTS), and MUTYH-associated polyposis (MAP), which are rare inherited conditions, also increase the risk of CRC development. Inflammatory bowel disease (IBD), commonly known as Crohn's disease or ulcerative colitis, causes chronic inflammation, contributing to colorectal carcinogenesis and potentially promoting tumour growth and progression. Precancerous neoplastic lesions, known as colon polyps, are key precursors in common tumourigenesis. Moreover, diabetes mellitus (chronic hyperglycemia) and cholecystectomy (gallbladder removal surgery) may elevate the risk of CRC development related to chronic inflammation and altered physiological conditions (34, 35).

Individual lifestyle is another concern regarding the accumulation of risk factors for colorectal carcinogenesis. Dietary patterns high in red and processed meat, while lacking fibre, fruits, vegetables, calcium, vitamin D, and dairy products, can lead to various pathogen induction, dysbiosis, and disrupted calcium homeostasis, potentially contributing to this process. Being overweight, along with obesity and physical inactivity, may promote oxidative stress, suppress the immune response, and alter hormone levels. Additionally, cigarette smoking and alcohol consumption have been suggested to induce carcinogenesis through a combination

of carcinogens and products formed during the oxidative metabolism of ethanol (34, 35).

Non-modifiable factors such as gut microbiota, age, sex, race, and socioeconomic influences that potentially impact CRC occurrence are highlighted. As noted, gut microbiota plays a crucial role in maintaining and protecting the gastrointestinal (GI) tract. Any alteration in normal microbiota function can create a microenvironment conducive to cancer development and progression. The majority of new CRC cases are diagnosed in patients over 50 years old, representing 90% of cases. Moreover, the incidence and mortality rates for male patients with CRC are higher than for female patients, with a 30% higher incidence in men and a 40% higher risk of dying from the disease. Non-Hispanic Black individuals exhibit a higher incidence than Asians/Pacific Islanders by 50% and non-Hispanic Whites by 20%. Lastly, low socioeconomic status (SES) may lead to limited access to medical services and resources, as well as a higher risk diet and lifestyle. Therefore, the development of CRC will eventually progress into more advanced aggression stages if prevention, diagnosis and curative care are insufficient, especially in communities with low SES (34, 35).

1.3.3 Colorectal cancer screening

During colorectal carcinogenesis, many patients with colorectal polyps or early-stage cancer often remain asymptomatic. However, common symptoms observed in clinics include abdominal pain, changes in bowel habits and movements, weight loss, nausea, vomiting, malaise, anorexia, and abdominal distension. These symptoms vary depending on the tumour's location, size, aggressiveness, and the presence or absence of metastases. Early stages are normally curative through surgery and good responses to chemotherapy treatment, whereas more aggressive tumours require advanced treatment approaches. Consequently, screening tests have been widely implemented for the early detection of colorectal polyps and cancer; these could reduce CRC incidence and mortality (30, 36).

1.3.3.1 Stool-based tests

Several screening tests are available for CRC screening. The guaiac faecal occult blood test (gFOBT) is a non-invasive procedure used to detect blood in stool samples. However, this test is known to potentially result in false positives due to dietary changes and medication intake (37). Consequently, since 2017, it has been replaced by the quantitative Faecal Immunochemical Test (qFIT). The qFIT improves adherence by utilising an antibody to detect human haemoglobin through automated analysis and quantification of faecal haemoglobin concentration (f-Hb). In Scotland, the Scottish Bowel Screening Programme (SBoSP) recommends screening eligible people aged 50 years and older every two years with the qFIT (38, 39). If the qFIT result is ≥80 µg haemoglobin/g faeces (µg Hb/g F), a colonoscopy will be undertaken (39, 40).

1.3.3.2 Blood-based tests

Blood-based tests, also known as liquid biopsies, have been developed as a less invasive and more reliable method for detecting CRC at an early stage and for monitoring cancer recurrence, progression, and therapeutic responses. However, these tests are not included in routine care. Key genetic and epigenetic mutations can be identified through circulating tumour cells (CTCs), circulating tumour DNA (ctDNA), exosomes, and tumour-educated platelets (TEPS). For instance, promoter hypermethylation of septin 9 (SEPT9) serves as a biomarker in CRC, detectable in cell-free DNA from patient plasma (37, 41). Additionally, carcinoembryonic antigen (CEA) and carbohydrate antigen 19-9 (CA19-9) are measured to monitor disease progression (42).

1.3.3.3 Modified Glasgow Prognostic Score (mGPS)

Cancer patients, particularly those in advanced stages, often experience an imbalance in metabolic rate, leading to weight loss or cachexia, reduced performance status, decreased quality of life, and increased morbidity and mortality. This imbalance is associated with the activation of a systemic inflammatory response due to tumour hypoxia, necrosis, or local tissue damage. An inflammation-based score has been developed to assess the systemic inflammatory response in clinical settings. For example, the modified Glasgow Prognostic Score (mGPS) combines C-reactive protein and albumin into three

scores (0-2): score 0 for patients with normal levels of both C-reactive protein (CRP ≤ 10 mg/L) and albumin (≥ 35 g/L); score 1 for patients showing one of the biochemical abnormalities; and score 2 for patients with both elevated C-reactive protein (> 10 mg/L) and hypoalbuminaemia (< 35 g/L). C-reactive protein and albumin are recognised as typical acute-phase proteins that are synthesised and released from hepatocytes into the systemic circulation to support the immune system. The mGPS is an independent prognostic indicator of tumour stage, performance status, and treatment in primary operable colorectal cancer. The combination of elevated C-reactive protein concentration and hypoalbuminaemia is linked to poorer outcomes ($p < 0.0001$) (43-45).

1.3.3.4 Direct visualisation tests

Several procedures have been developed to visualise abnormal masses in the bowel. Flexible sigmoidoscopy is a direct visualisation tool that can detect and biopsy polyps in the distal part of the colorectum, significantly reducing the mortality and incidence of CRC. Since this procedure requires bowel preparation, specifically enemas, it is recommended to perform it every five years. A colonoscopy is indicated as a standard screening method that can identify polyps throughout the colorectum, beginning at age 50 and is also recommended at five-year intervals in the screening programme. Before the colonoscopy, dietary modifications and the administration of a purgative preparation with sedation are performed. It is combined with polypectomy to resect the lesions, which can significantly reduce the incidence rate of CRC. However, it has been suggested that there is a missing rate of approximately 22% for some cases with polyps of any size (30, 37).

Virtual colonoscopy, also known as computed tomographic colonography (CTC), is a procedure used to visualise the entire colorectum through bowel preparation, which includes air distension, sedation, and/or radiographically tagged agents. It is recommended for screening every five years and is employed after an incomplete colonoscopy or in complex cases that require experienced staff and resources, while also considering risks such as radiation exposure, reduced sensitivity in detecting small polyps, and potential pitfalls. Magnetic resonance colonography (MRC) is another method, particularly due to the relatively high radiation exposure during CTC, and it is advantageous for detecting large polyps;

however, it is not often used in clinical practice. With similar benefits and limitations in screening for colonic polyps or cancer, capsule colonoscopy is not routinely used. The camera is designed to visualise lesions throughout the entire bowel (30, 37).

In accordance with the SBoSP guidelines, colonoscopy is generally employed to investigate new symptomatic patients as part of early detection screening. After the detection and removal of polyps that potentially prevent progression to cancer, a surveillance colonoscopy is performed to follow up the patients at higher risk of CRC (40, 46).

1.3.3.5 Polypectomy and surgery

Endoscopic polypectomy is an effective method for biopsying colorectal lesions and is also recognised as a treatment for colorectal polyps. This procedure can be performed using forceps as a simple tool; however, it may not guarantee complete removal at the base of the polyps. The snare technique is more effective for polyp removal, as a wire is threaded through the endoscope channel to encircle the polyp's stalk and employs monopolar electric current to separate it. Nevertheless, its greatest advantage is primarily applicable to pedunculated polyps and is limited to sessile polyps. Therefore, saline injection into the submucosal layer beneath the polyp, followed by dissection from the underlying muscular layer, allows for the complete removal of the entire sessile polyp with normal mucosal borders. This technique is known as Endoscopic Submucosal Dissection (ESD) (30).

Surgery, specifically colectomy, is another option for treating colorectal polyps, as some polyps cannot be fully removed from synchronous lesions through colonoscopic resection. Dysplasia may persist and potentially lead to metachronous lesions, future polyps, or even cancer. Surgery is conducted according to oncological guidelines for anatomical resection, which includes the relevant mesentery containing the lymphatic basin with arterial high ligation, as well as histological features associated with an increased risk of residual disease and lymph node metastases (30). The Haggitt system categorises pedunculated polyps into four levels of invasion of the submucosa: mucosal high-grade dysplasia (level 0), invasion limited to the head of the polyp (level 1), invasion into the neck

of the polyp (level 2), invasion into any part of the stalk (level 3), and invasion beyond the stalk but above the muscularis propria (level 4) (4, 47). The Kikuchi system provides a classification of submucosal infiltration for sessile polyps, divided into three levels: superficial invasion of the upper third from the muscularis propria (SM1), moderate invasion (SM2), and deep invasion of the lower third (SM3) (4, 48). After polyp resection, patients are advised to follow up within a timeframe of five years, depending on the degree of the index lesions (30).

1.3.3.6 Surveillance screening programmes

During screening, individuals found to have premalignant polyps through colonoscopy, constituted 25% to 50% of cases. Additionally, these premalignant polyps appear to contribute to the development of metachronous polyps (future polyps) or CRC. However, it is suggested that only 5% of patients will go on to develop CRC. In 2020, the British Society of Gastroenterology (BSG), the Association of Coloproctology of Great Britain and Ireland (ACPGBI), and Public Health England (PHE) established surveillance guidelines for post-polypectomy (endoscopic follow-up) and post-CRC resection as part of the national bowel cancer screening framework. This initiative aims to maximise the benefits of the procedure for high-risk patients while minimising unnecessary management for low-risk groups. An efficient procedure at initial colonoscopy can remove all identified premalignant polyps, potentially leading to reduced CRC incidence and mortality by preventing these lesions from progressing to CRC. Nevertheless, advanced colorectal polyps (including both advanced serrated and adenomatous polyps sized at least 10 mm or containing any dysplasia grade of serrated polyps or high-grade dysplasia of adenomatous polyps) are cleared, although the higher-risk group may still develop CRC. Thus, high-risk patients with either two or more premalignant polyps, including at least one advanced colorectal polyp or five or more premalignant polyps, are considered to receive a one-off surveillance colonoscopy at 3 years. Post-CRC resection patients will receive a 1-year clearance colonoscopy, followed by a surveillance colonoscopy after three more years (6).

Clinical outcomes from a surveillance colonoscopy involving a total of 2,643 eligible patients in a retrospective multicentre study conducted in Scotland indicated that approximately 45.6% of patients had not presented with

metachronous lesions, 39.8% had non-advanced polyps, 13.7% had advanced polyps, and 0.9% of patients had CRC. Although the BSG 2020 criteria, based on index polyps, significantly predicted a higher rate of overall metachronous polyps in high-risk patients compared to low-risk patients ($p < 0.001$), there was no difference observed between non-advanced and advanced metachronous polyps, nor in the prediction of late metachronous lesions detected after 2 years from the index polypectomy (49).

1.3.4 CRC staging system and classification

As a consequence of surgical resection, colorectal tissue specimens are immediately preserved in formalin, then processed and diagnosed in the pathology laboratory. Precancerous lesions are evaluated as previously described. Tumours are classified according to the following classification systems to determine their staging and stratification subtypes regarding heterogeneity for patient treatment management and prognosis.

1.3.4.1 TNM staging

To report on the pathological stages of colorectal tumours after surgery, the 8th edition of the staging system defined by the American Joint Committee on Cancer (AJCC) is utilised in relation to the tumour, lymph node, metastasis (TNM) system, which is a standardised classification system. A primary tumour (T category) may be described by a lack of information on the tumour (TX), no evidence of a primary tumour (T0), *in situ* cancer, or precancer that is growing in the muscularis propria (Tis), as well as tumour size and/or extent of spread into nearby structures (T1-T4). Lymph nodes (N category), located near the primary tumour, are often evaluated as lacking information regarding regional lymph nodes (NX), showing no infiltrated tumour in lymph nodes (N0), and demonstrating the spread of tumour in lymph nodes (N1-N2). Metastasis (M category) examines cancer spread to other parts of the body, reporting no distant cancer metastasis (M0) and distant metastasis (M1). The AJCC staging is categorised by stage 0, referred to as carcinoma *in situ* (Tis, N0, M0); stage I, where cancer has grown through the muscularis mucosa but not spread to nearby lymph nodes (T1-2, N0, M0); stage II, where cancer has grown into the bowel wall but not spread to nearby lymph nodes (T3-4, N0, M0); stage III, where cancer has grown through the muscularis

mucosa into the bowel wall and spread to nearby lymph nodes but not to distant sites (T1-4, N1-2, M0); and stage IV, where cancer has spread to distant organs or parts of the peritoneum (Any T, Any N, M1) (50, 51), see Figure 1-4.

According to the TNM staging system, Dukes' classification is an older method for tumour staging and predicting survival outcomes. This system evaluates the extent of tumour infiltration through the bowel wall and the presence of lymph node metastases. It is divided into four groups: a tumour that penetrates the muscularis propria (Dukes A), a tumour that invades beyond the muscularis propria (Dukes B), a tumour that shows adenocarcinoma in at least one lymph node (Dukes C), and a tumour that has metastasised to other organs (Dukes D) (4, 52). Dukes A is associated with a longer survival rate compared to the higher stages (53). Adjuvant therapy is considered beneficial for Dukes B and C patients with related high-risk features (54).

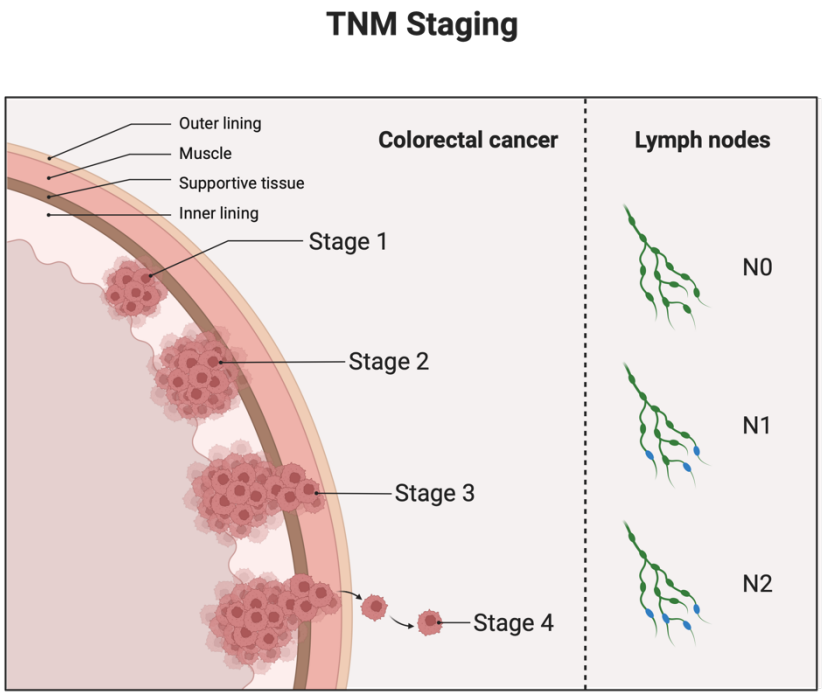


Figure 1-4 Tumour, lymph node, metastasis (TNM) system for staging colorectal cancer

The diagram illustrates the TNM system used for staging colorectal cancer, which is based on features of the primary tumour (T category), lymph nodes (N category), and metastasis (M category), and categorises it into stages I to IV (the AJCC staging). This figure was created using BioRender.com and adapted from Cancer Research UK.

1.3.4.2 Pathological prognostic parameters

Location of colorectal carcinoma

Focusing on the tumour, the location of the primary tumour is emphasised as being associated with various clinical and biological characteristics. The sidedness of colorectal carcinoma is classified as right-sided colon, left-sided, and rectal (55). Notably, rectal cancer (RC) differs from colon cancer (CC) in molecular carcinogenesis, pathology, surgical topography and procedures, and multimodal treatment, despite both types of cancer sharing a similar anatomical structure (56). Right-sided colon carcinomas (RCCs) are more likely to occur with higher incidence in females, among older age groups, with high-grade histology, and at more advanced tumour stages, showing trends of metastasis to the peritoneum. In contrast, liver and lung metastasis are associated with left-sided colon carcinomas (LCCs) and RC (26, 55, 56).

In terms of genetic and epigenetic mutations, RCCs are found to be more involved in hypermutation, CpG island Methylator phenotype (CIMP)-high and Microsatellite instability (MSI)-high phenotypes, as well as increased mutations of *KRAS*, *BRAF*, *PIK3CA*, and *TGFβR2*. Additionally, chromosomal instability (CIN) related to mutations of *APC*, *KRAS*, *TP53*, and *SMAD4*, along with overexpression of several genes, including epidermal growth factor receptor (EGFR) ligands (*EGFs*), epiregulin (EREG), amphiregulin (*AREG*), human epidermal growth factor receptor 2 (*HER2*), vascular endothelial growth factor (*VEGF-1*), and cyclooxygenase-2 (*COX-2*), is indicated to be associated with a higher incidence in left-sided tumours. In relation to the gene expression-based classification of CRC, also known as Consensus Molecular Subtypes (CMS) (57). RCCs are associated with CMS1 (microsatellite instability immune) tumours, which are linked to an increased tumour mutational burden, MSI, CIMP+, *and BRAF* mutation. Conversely, left-sided cancers are associated with the classical subtype CMS2, characterised by activation of WNT and MYC signalling. A better overall clinical outcome is predicted in patients with left-sided cancers. Furthermore, in response to anti-EGFR therapy, better overall survival (OS; duration after surgery until all-cause mortality) and progression-free survival (PFS; duration after treatment until recurrence or all-cause mortality) compared to RCCs. Meanwhile, right-sided RAS

wild-type tumours show a superior response to anti-VEGF therapy (26, 55), see Figure 1-5.

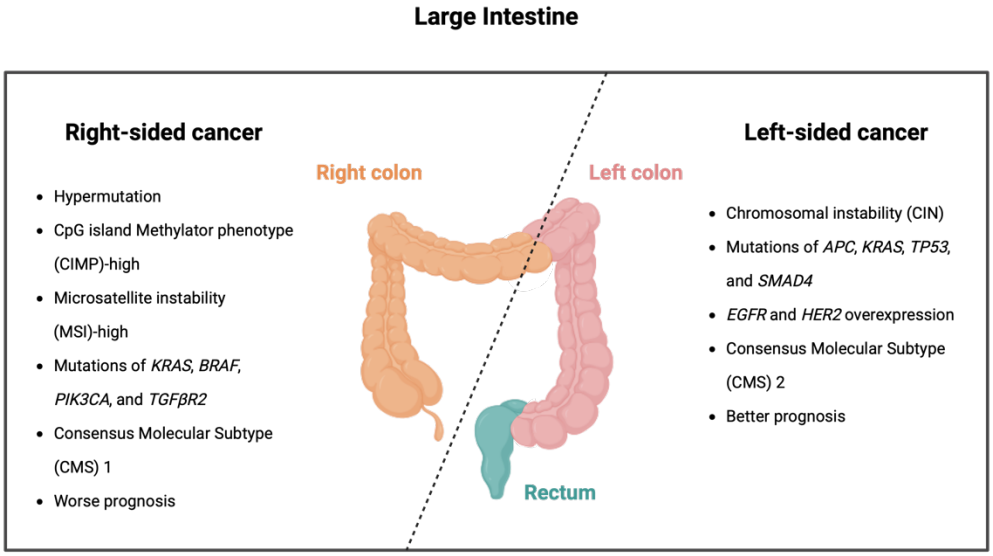


Figure 1-5 Clinical and molecular differences between right- and left-sided colorectal cancers

The diagram illustrates key clinical and molecular differences between right-sided and left-sided colorectal cancers. Right-sided cancers are involved in hypermutation, CIMP-high, and MSI-high phenotypes, with increased mutations in *KRAS*, *BRAF*, *PIK3CA*, and *TGFβR2*, related to CMS 1, and tend to have a worse prognosis. Conversely, left-sided cancers involve CIN, with mutations in *APC*, *KRAS*, *TP53*, and *SMAD4*, overexpression of *EGFR* and *HER2*, and are associated with CMS 2 and a better prognosis. This figure was created using BioRender.com and adapted from Puccini A., Marshall JL., Salem ME., 2018 and Stintzing S. et al., 2017.

Petersen Index (PI)

Nearly half of CRC patients were diagnosed with Dukes' B and had the potential for relapse and recurrence. Since Dukes' C benefits from adjuvant chemotherapy in reducing recurrence and mortality, the pathological prognostic parameters were evaluated for their relationship with survival outcomes in Dukes' B colon cancer. Four main parameters, including peritoneal involvement, venous invasion, margin involvement, and tumour perforation, are significant independent prognostic factors. Petersen et al. combined these factors into the prognostic index (PI) to guide high-risk patients on the necessity of adjuvant chemotherapy for Dukes' B colon cancer. The scoring system (score 0-5) can be calculated by assigning a PI value of one for peritoneal involvement with or without ulceration, adding 1 for extramural or submucosal venous spread, adding 1 if the margin is

involved or inflamed, and adding 1 for tumour perforation. Scores of at least two were grouped as a high-risk group, indicating reduced survival times, with a five-year survival rate of 49.8% (58).

Perineural invasion (PNI)

Similarly, perineural invasion (PNI) serves as another pathological prognostic parameter involved in tumour-stromal interactions for stratifying CRC patients, particularly those with node-negative disease, regarding adjuvant treatment. An aggressive tumour phenotype is characterised by the invasion of tumour cells within the nerve sheath layer or in the perineural space, affecting at least one-third of the nerve circumference. PNI-positive patients are associated with reduced disease-free survival (DFS; duration after surgery or randomisation until recurrence or all-cause mortality) and overall survival (both $p < 0.0001$). Among node-negative patients, those with PNI-positive status exhibit poorer outcomes compared to PNI-negative and CRC stage III patients (DFS at 5 years: 29% with $p = 0.0002$ and OS at 5 years: 43% with $p = 0.002$) (59).

Tumour budding (TB)

Tumour budding is defined as a single tumour cell or a cluster of four or fewer tumour cells, as observed in haematoxylin and eosin (H&E) full sections. It can be examined using the hotspot method at the invasive front, where tumour buds are counted from a single field with the highest density of tumour buds using an X20 objective lens and reported in a field measuring 0.785 mm^2. The budding counts are classified into two levels: low TB (0–9 buds) and high TB (≥ 10 buds) (60, 61). Additionally, tumour buds may be challenging to distinguish from peritumoural inflammation and reactive stromal fibroblasts. An option for immunohistochemistry staining of cytokeratin can differentiate the tumour buds by their cytoplasmic reactivity and the presence of a clearly identifiable nucleus (62). Tumour budding has been suggested to be associated with an EMT and is an independent factor related to worse outcomes for CRC patients, including the detection of tumour budding in pT1 CRC, which identifies a risk factor for lymph node metastasis and requires surgical resection; the detection of tumour budding in stage II CRC, indicating a high risk for metastatic relapse and suggesting adjuvant therapy selection; and the presence of tumour budding in preoperative

biopsies of rectal cancer patients, which identifies a high risk for local recurrence and metastasis, thereby influencing neoadjuvant treatment and risk-adapted surgery (60, 62).

Klintrup–Mäkinen grade (KMG)

The Klintrup–Mäkinen grade (KMG) is a practical and routine assessment for evaluating local inflammatory reactions in CRC. The peritumoural inflammatory reaction is identified at the invasive margin, which is the interface between the host stroma and the invading edge of a tumour in haematoxylin and eosin-stained sections. A four-degree scale defines the inflammatory reaction and the quantity of lymphoid cells, neutrophilic granulocytes, and eosinophilic granulocytes: score 0 indicates no increase in inflammatory cells; score 1 denotes a mild and patchy increase in inflammatory cells; score 2 reflects a band-like infiltrate at the invasive margin with some evidence of destruction of cancer cell islets; and score 3 represents a very prominent inflammatory reaction with frequent destruction of cancer cells. Subsequently, the scores are categorised into low grade (score 0-1) and high grade (score 2-3). High-grade inflammation has stage-independent prognostic value, which is related to better 5-year survival (87.6%, $p < 0.00005$) in CRC, particularly in Dukes' stage A and B cancer patients, and is not likely to benefit from adjuvant therapies compared to the low-grade group (63, 64).

Tumour stromal percentages (TSP)

At the invasive margin, the tumour stroma facilitates cancer progression as part of the tumour microenvironment, which is independently associated with more aggressive tumour growth, invasion, the host systemic inflammatory response (Modified Glasgow Prognostic Score, mGPS), and a weak peritumoural inflammatory infiltrate (Klintrup–Mäkinen grade, KMG). Tumour stromal percentages (TSP) can be assessed from the deepest point of the tumour invasion margin (at X5 magnification) on H&E-stained sections. Average percentages are calculated for the stroma area, showing the nearest 5% of tumour cells on all four sides of a single field image at X10 magnification. Subsequently, it is categorised into low TSP (≤50%) and high TSP (>50%). Despite the association of reduced cancer-specific survival (CSS; duration after surgery until cancer-related

mortality) with high tumour stroma (p = 0.001) in primary operable CRC patients, similar predictive outcomes were observed in patients who underwent adjuvant chemotherapy (p = 0.009), and anti-angiogenic therapies may benefit these high-risk patients (65).

Glasgow Microenvironment Score (GMS)

A combination of tumour-infiltrating inflammatory cells and tumour stroma was developed to assess CRC progression and outcomes based on routine histopathologic evaluation of tumour and microenvironment interactions. The cumulative score was calculated from KMG and TSP following the previously described methods, referred to as the Glasgow Microenvironment Score (GMS), which is categorised into three groups: GMS 0, strong KMG regardless of TSP; GMS 1, weak KMG with low TSP; and GMS 2, weak KMG with high TSP. The GMS is a significant prognostic factor independent of lymph node involvement, venous invasion, and MMR status in primary operable CRC patients. Patients with stage I–III colorectal cancer exhibiting weak KMG and high TSP (GMS 2) were associated with reduced 5-year survival outcomes compared to other groups (GMS 0 = 89%, GMS 1 = 75%, and GMS 2 = 51%, p < 0.001) (66). Additionally, GMS 2 predicted worse survival outcomes in patients with TNM I–III CRC within 5 years, in terms of disease-free survival (GMS 0, 1, and 2 of 71%, 58%, and 46%, respectively; GMS 0 vs GMS 2, p = 0.002) and relapse-free survival (RFS; duration after surgery until recurrence or cancer-related mortality; GMS 0, 1, and 2 of 83%, 70%, and 51%, respectively; GMS 0 vs GMS 2, p < 0.001). Furthermore, TNM II–III CRC patients, all of whom received adjuvant chemotherapy for at least 3 months with GMS 0, were associated with better treatment outcomes (DFS for GMS 0, 1, and 2 of 69%, 63%, and 53%, respectively; GMS 0 vs GMS 2 p < 0.001), especially when receiving FOLFOX (bolus and infused fluorouracil with oxaliplatin) in the low-risk group TNM III (Petersen index) (67).

Tumour Microenvironment Score (TMS)

Recently, the tumour microenvironment score (TMS) was established as a comprehensive and routine prognostic tool based on assessing interactions between the tumour and its microenvironment as identified in H&E-stained sections. The combination of tumour budding (TB), Klintrup–Mäkin grade (KMG),

and tumour stromal percentage (TSP) is used to examine the levels of tumour budding at the invasive front, tumour-infiltrating immune cells, and tumour-associated stroma at the invasive margin, respectively, following previously described methods. The score categorises patients into four groups: TMS 0, high KMG; TMS 1, low for KMG, TSP, and TB; TMS 2, low KMG and high for TB or TSP; and TMS 3, low KMG and high for TSP and TB. TMS has been shown to be an independent prognostic factor in stage I–III CRC patients undergoing surgical resection. Patients with TMS 3 were associated with the shortest CSS compared to the other groups (a 5-year survival; TMS 0 = 86%, TMS 1 = 69%, TMS 2 = 60%, and TMS 3 = 48%; $p < 0.001$), and this subgroup was also linked to higher cancer stage, aggressive tumour features, invasion and recurrence, as well as enriched gene expression for hallmark signalling pathways such as EMT ($p < 0.001$), IL2/STAT5 ($p = 0.007$), and angiogenesis ($p = 0.017$). However, a lower immunological interaction in the TMS 3 tumour may render it less responsive to checkpoint inhibitors (61).

1.3.4.3 Molecular alterations at the genomic level

Regarding the previously described CRC pathogenesis, three significant alterations in genetic and epigenetic mechanisms- chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) - have been widely utilised to classify the heterogeneity of tumour characteristics based on genomic instability in clinical practice.

Chromosomal instability (CIN)

CIN refers to changes in chromosomes or karyotypic abnormalities, including somatic copy number alterations (SCNA) caused by imbalances in chromosome number (aneuploidy), subchromosomal genomic deletions, insertions, amplifications, or loss of heterozygosity (LOH). These alterations are associated with mechanistic defects in chromosomal segregation that govern sister chromatid separation, telomere shortening leading to genomic reorganisation, dysfunctional DNA damage-response machinery, and LOH at tumour suppressor genes, resulting in a relative paucity of base pair mutations in coding sequences. Several approaches, including cytometry, karyotyping, loss of heterozygosity analysis, fluorescent *in situ* hybridisation (FISH), and comparative genomic

hybridisation (CGH), can designate a tumour as CIN-positive. The onset of CIN is identified as the acquisition of mutations in the initiation and progression of CRC tumourigenesis, which can arise from both somatic and familial mutations. These mutations occur in the tumour suppressor genes APC and *TP53*, as well as in the oncogenes *KRAS* and *PIK3CA* (19, 68).

Approximately 75% of CRC is involved in chromosomal instability (CIN^+/MSS/ $CIMP^-$) (69). Tumours that exhibit loss of APC, located on chromosome 5q, activate the Wnt signalling pathway in 80% of CRC cases. Due to Wnt signalling regulation, several genes become constitutively activated, including *AXIN1*, *AXIN2*, *CTNNB1*, *MYC*, the cyclin D1 gene (*CCND1*), the *VEGF* genes, and the peroxisome proliferator-activated receptor delta (*PPARδ*) gene (19, 68). *KRAS* mutation tumours often arise after mutations in *APC*, linked to single-nucleotide point mutations at codons 12, 13, and 61, which are present in nearly 40% of CRC cases. Several growth factor signalling pathways, including the EGFR pathway, Raf-MEK-ERK pathway, and PI3K signalling via MTOR, alongside the transcription factor nuclear factor kB (NF-κB), result from *KRAS* mutation activation, accompanied by dysregulation of glucose metabolism and hypoxia (19, 68). Furthermore, mutations in genes encoding the catalytic subunits of PI3K at exons 9 and 20, accounting for almost 20% of CRC, have been shown to correlate with *KRAS* mutation tumours (19, 68). The *TP53* gene (located on chromosome 17p) controls numerous transcription factors and functions as a tumour suppressor and a central coordinator of cellular responses to stress induced by several oncogenic proteins, such as c-Myc and RAS. Missense mutations represent the most common type of *TP53* mutations, accounting for approximately 80%. The loss of function in TP53 is observed in 4-26% of adenomas and in 50-75% of CRC cases. Furthermore, *TP53* mutations are more frequently identified in the later stages, accounting for 60% of CIN tumours (19, 68).

LOH is found in an average of 25–30% of tumours and is most frequently lost on five chromosomes (1, 5, 8, 17, and 18). Approximately 70% of allelic loss on chromosome 18 has been detected in advanced CRC stages. The mutations of the genes *Deleted in Colorectal Cancer (DCC),* which encode tumour suppressor genes (Netrin-1, a cell surface receptor), occur in about 6% of CRC. Other tumour suppressor genes, *SMAD2* and *SMAD4* genes, which encode intracellular protein mediators in the transforming growth factor beta (TGF-ß) pathway, occur in fewer

than 20% and 10% of CRC, respectively (19, 68). A reduction in survival time is associated with the effects of the loss of 18q in advanced CRC, and CRC stage II patients would benefit from adjuvant therapy (19, 70).

Microsatellite instability (MSI)

MSI is a hypermutable pathway characterised by frequent somatic DNA base pair mutations. This phenomenon involves repeating genomic sequences (short tandem repeats) composed of mononucleotide, dinucleotide, or higher-order nucleotide repeats. These sequences lead to cumulative errors in DNA replication and impair accurate replication during new strand synthesis. Consequently, they result in frameshift mutations and dysfunctional proteins. Due to immunogenicity, the tumour becomes infiltrated by an abundance of cytotoxic T cells and natural killer cells. The mutations in MSI are strongly associated with the inactivation of DNA MMR genes, including *MLH1*, *MSH2*, *MSH6*, and *PMS2*, resulting in instability within microsatellite regions (19, 71). Approximately 15% of CRC cases are associated with MSI (CIN^-/MSI/$CIMP^{+/-}$) (69). The development of MSI tumours is additionally caused by an increased frequency of epigenetic silencing of the *MLH1* gene through promoter hypermethylation by the cytosine/guanine (CpG) island methylator phenotype (CIMP) in sporadic CRC, mutations in *BRAF* causing a V to E substitution at amino acid 600 ($BRAF^{V600E}$), and mutations in genes that encode proteins regulating cell proliferation, cell cycle arrest, and apoptosis, such as the *TGFBR2* gene (detected in 90% of MSI tumours), as well as a germline mutation in a DNA MMR gene (Lynch syndrome) (19).

It is emphasised that the malfunction of the MMR system or deficient DNA mismatch repair (dMMR) represents a significant signature in MSI, contrasting with MMR-proficient (pMMR) or microsatellite stability (MSS). Two different methods are employed to assess the status of dMMR and MSI. Molecular MSI testing is conducted using DNA extracted from frozen or formalin-fixed paraffin-embedded (FFPE) tumour tissue, followed by the simultaneous amplification of five markers in multiplex PCR with reference panels (Bethesda and pentaplex panel). If at least 40% of the panel is unstable, the tumour is classified as having MSI-high status. Another method involves MMR protein testing through IHC staining of MLH1, MSH2, MSH6, and PMS2. Loss of nuclear tumour staining in one of the four MMR proteins defines a dMMR tumour (71). The dMMR/MSI testing in CRC patients has clinical

relevance in identifying adjuvant chemotherapy for stage II CRC. Patients with stage II dMMR/MSI CRCs are associated with a better prognosis than the pMMR/MSS tumours, but they are linked to chemoresistance to fluoropyrimidine. In addition, anti-VEGF and immune checkpoint inhibitors (ICI) have been suggested to be beneficial in chemoresistant dMMR/MSI metastatic tumours. Furthermore, the detection of a monoallelic germline mutation of the MMR genes is applied to screen for Lynch syndrome (19, 71).

CpG island Methylator phenotype (CIMP)

CIMP refers to the increased methylation of genomic regions at promoter sites that are rich in cytosine preceding guanine dinucleotides, known as CpG islands. In the general mechanism of epigenetic regulation for genomic stability, embryonic development, and tissue differentiation, the process of DNA methylation is linked to the addition of a methyl group to the 5-position of cytosine by DNA methyltransferases (DNMT). Aberrant hypermethylation leads to the inappropriate silencing of gene expression, particularly of tumour suppressor genes, which contributes to tumourigenesis through epigenetic instability. Several genes have been identified as associated with CIMP, such as *CDKN2A*, which encodes the tumour suppressor p16; CXC motif chemokine 12 (*CXCL12*), which encodes a chemokine ligand; and *MLH1*, encoding a mismatch repair gene (72). Approximately 20% of CRC is linked to DNA methylation ($CIN^{-}/MSI^{+/-}/CIMP^{+}$) (69).

Nevertheless, accuracy and consistency in classifying the CIMP tumour phenotype remain challenging. Multiple panels for DNA methylation analysis (combined bisulfite restriction analysis and Methylation-specific polymerase chain reaction; MSP) have been developed based on specific methylated loci (72). The classical marker panel consists of *hMLH1*, *p16*, *MINT1*, *MINT2*, and *MINT31*. CIMP tumours can be classified using the five markers into CIMP-negative (no methylated loci), CIMP-low (1-3 methylated loci), and CIMP-high (4-5 methylated loci) (73). Additionally, *CACNA1G*, *CRABP1*, *IGF2*, *NEUROG1*, *RUNX3*, *SOCS1*, *HIC1*, *IGFBP3*, and *WRN* have also been included in other panels (74-78). Furthermore, CIMP tumours have been shown to be associated with *BRAF* and *KRAS* mutations (74, 75, 78), as well as MSI (72). Even though aberrant DNA methylation of genes has been studied for prognostic value in CRC patients, controversy and the unclear

reproducibility of the predictive role of the CIMP phenotype regarding survival outcomes and chemoresistance still remain (19, 72).

1.3.4.4 Molecular pathological classification at the transcriptional levels and biological pathways of tumours and their microenvironment

After introducing DNA-based analysis for CRC tumours through the Cancer Genome Atlas (TCGA) network project in 2012, this project highlights two significant mechanisms of genomic instability that contribute to substantial heterogeneity within tumours: CIN, which accounts for 84%, and MSI, which represents 13 to 16% of sporadic CRC. As previously described, CIN tumours are detected with a high frequency of SCNA, MSS phenotype, and mutations related to WNT signalling, which are emphasised in tumourigenesis through the adenoma-carcinoma sequence. Another tumour phenotype linked to MSI includes hypermutation associated with dMMR, CIMP-H, *BRAF* mutation, and tumour-infiltrating lymphocytes (TILS), as well as ultramutated proofreading exonuclease domains in the two polymerases POLE and POLD1 (79, 80). Following the genomic-level classification, other molecular alterations based on high-throughput gene expression, pathological assessments, and cell signalling pathway analyses of tumour cells and their tumour microenvironment compartments have been studied to develop molecular subclassification and risk stratification methods for revealing the intrinsic biology of CRC in inter-tumoural and intra-tumoural heterogeneity, and informing pathology-led studies on clinical attributes (80).

Consensus molecular subtypes (CMS)

CMS investigated bulk RNA expression profiling data of CRC from six independent studies and subsequently identified four major CMS groups, along with a residual mixed group: CMS1, a MSI group related to hypermutated, microsatellite unstable, and intense immune activation (14%); CMS2, a canonical group related to epithelial characteristics, marked by WNT and MYC signalling activation (37%); CMS3, a metabolic group associated with epithelial cells and exhibiting evident metabolic dysregulation (13%); CMS4, a mesenchymal group linked to prominent TGF-β activation, stromal invasion, and angiogenesis (23%); and mixed features that possibly represent a transition phenotype or intratumoural heterogeneity (13%) (57).

The CMS classification approach has served as a transcription-based reference in alignment with clinical outcomes and potential therapeutic responses in subsequent studies (80). For example, a five-marker immunohistochemistry (IHC) panel (FRMD6, ZEB1, HTR2B, CDX2, and cytokeratin) was developed alongside standard MSI/dMMR testing and pan-cytokeratin IHC to identify molecular subtypes in CRC. However, it still has limitations in clearly distinguishing between classes. Additionally, digital pathology tools such as QuPath and Halo, along with image-based H&E approaches, have emerged as alternative resources to facilitate pathological diagnosis. An image-based CMS (imCMS) deep learning classifier was designed with accurate spatial assessment of H&E images for concordance with transcriptional CMS (81). Furthermore, to understand the underlying diversity of epithelial cells, single-cell intrinsic consensus molecular subtype (iCMS) classification was proposed to refine the CMS subtypes using single-cell transcriptomes of tumour epithelial cells. Two intrinsic subtypes, iCMS2 and iCMS3, are based on intrinsic epithelial subtype, MSI status, and fibrosis (IMF) classification. The iCMS3 tumours were associated with *BRAF*, *KRAS*, and *PIK3CA* mutations and are more similar to MSI-H than others, which may lead to benefits with immunotherapy. In contrast, iCMS2 tumours were associated with mutations in *APC* and *TP53*. The CMS4 tumour can be identified into iCMS2 (better outcome) and iCMS3 (worse outcome) subtypes. Nevertheless, this approach cannot define distinct biology within CMS2 and CMS3 tumours (80, 82).

Colorectal cancer intrinsic subtypes (CRIS)

Regarding the influence of the TME on intratumoural heterogeneity and its clinical implications, fibroblast infiltration and genes specifically originating from cancer-associated fibroblasts (CAFs) have been identified as playing a role in disease relapse (80). Therefore, to focus on cancer-cell intrinsic transcriptional features, profiling tumours from patient-derived xenografts (PDXs), while filtering out signals from stromal and immune components, has been established and is defined as CRC intrinsic subtypes (CRIS). This classification consists of five groups with distinctive molecular, functional, and phenotypic characteristics: CRIS-A, related to mucinous, glycolytic, enriched for MSI or *KRAS* mutations; CRIS-B, linked to TGF-β pathway activity, EMT, and poor prognosis; CRIS-C, associated with elevated EGFR signalling and sensitivity to EGFR inhibitors; CRIS-D, linked to WNT

activation, IGF2 gene overexpression, and amplification; and CRIS-E, related to a Paneth cell-like phenotype and *TP53* mutations. This CRIS subtyping can be classified regardless of primary or metastatic tumours, which has an improved impact on the stability of classification (80, 83).

Cell of origin and plasticity in tumourigenesis

Due to advancements in single-cell technologies that provide greater insight into the cellular biology of tumours, these applications also assist in confirming aberrant molecular associations related to the precancerous pathways of conventional adenomas and serrated polyps (80). Multi-omic analysis, including single-cell RNA sequencing of polyps, highlights that conventional adenomatous polyps are associated with WNT-driven expansion of stem cell populations (LGR5+ cells). In contrast, serrated polyps are related to *BRAF* mutation and MSI-H, deriving from differentiated cells through metaplasia, activating antigen processing through a regenerative mechanism that leads to a cytotoxic immune microenvironment characterised by CD8+ T cells, NKs, and γδT cells (84).

Furthermore, changes in cell composition and state are emphasised during the progression from normal tissue to cancer. In the early precancerous stages, there is an increase in stem-like epithelial cells exhibiting molecular alterations, particularly in adenomatous polyps. Tregs, exhausted T cells, pre-CAFs, and mature CAFs demonstrate a higher proportion in the cancer-associated tumour microenvironment during the transformation state (85). However, multiple distinct pathways may provide alternative routes for progression due to cellular plasticity. LGR5 negativity is associated with stromal/inflammatory lesions and the dominance of a differentiated regenerative-like metaplasia stem population, particularly serrated polyps (84). In addition, LGR5 negativity appears to switch intrinsic capacity between the LGR5-negative ANXA1+ regenerative stem cell (RSC) state and LGR5-positive colonic crypt base columnar cells (CBCS) properties in order to adapt and respond to cell-intrinsic and microenvironmental signalling disruption, which leads to aggressive features related to metastasis and chemoresistance (86, 87).

Pathway-derived subtypes (PDS)

Ongoing inter- and intra-cellular signalling pathways within the heterogeneous tumour mass provide numerous essential features associated with biological processes, cellular components, and molecular functions in tumour development and clinical outcomes. Hallmark characteristics, from altered gene expression aligned to Gene Ontology signatures, are typically identified through the Molecular Signature Database (MSigDB) (80, 88). The elevation of signalling pathways is highlighted in CMS subtypes, such as cell cycle activation and signalling in both WNT and MYC targets, which are associated with CMS2; metabolic signalling pathways related to CMS3; and TGF-β activation and EMT signalling linked to CMS4 (57). Nevertheless, pathway-level data seems to provide deep insight into molecular mechanisms of tumours and the biological signalling related to treatment response (80). An alternative approach integrates information on gene ontology and biological activation states for the initial stages of classification, replacing the focus on gene-level data at the beginning. Pathway-derived subtypes (PDS), using transcriptional data based on mechanistic signalling cascades, can classify CRC tumours with independent of *KRAS* mutational status into three major subtypes, including PDS1, which is linked to high proliferation with increased cell cycle-related pathways and MYC/WNT target activation of CMS2 tumours, fast-cycling canonical stem cells and good prognosis (26%); PDS2, is associated with immune/stroma-rich of CMS1 and CMS4 tumours, regenerative stem cells, and intermediate prognosis (31%); PDS3, is related to a slow-cycling or a lethargic biological phenotype (low Ki67 staining) of CMS2 tumours with neural-like traits, increased numbers of differentiated colonic epithelial lineages (stem-depleted), transcriptionally repressed and the worst prognosis (30%); and a smaller, more heterogeneous residual ‘mixed’ group (13%) (89).

From histological features and fundamental genomic and transcriptomic mechanisms to revealing unique insights into tumour biological traits and intercompartmental crosstalk through signalling pathways that drive tumour-wide responses, these could lead to a better understanding of cancer-relevant biological phenotypes and guide novel biological discovery and risk stratification in future research (80).

1.3.5 CRC pathogenesis and key mutations

CRC is a complex disease associated with various risk factors, both hereditary and sporadic (non-hereditary), that contribute to its etiology. Throughout the stages of CRC tumourigenesis, there are four main phases: initiation, promotion, progression, and metastasis. The initiation phase is linked to irreversible genetic alterations, which may be caused by various carcinogens (1-3). This phase results in the failure of DNA repair mechanisms or cellular apoptosis, leading to the emergence of small benign neoplasms (adenomas) in the colorectal epithelium over a span of 30 to 60 years. In the promotion phase, altered cells accumulate additional genetic and epigenetic disruptions, leading to proliferation. The formation of a clinically detectable large adenoma during the subsequent progression phase occurs in cancer development over 10 to 20 years. Eventually, cancer penetrates through the muscularis mucosa, becomes invasive and spreads from the primary tumour to distant sites via haematogenous or lymphatic routes within the subsequent 5 years, known as the metastasis phase (2, 90), Figure 1-6.

The pathogenesis of CRC can be characterised by cellular and molecular features, particularly as two distinct precursor lesions: the adenoma-carcinoma pathway and the serrated pathway. These pathways exhibit different yet overlapping mutations at the genetic and epigenetic levels, contributing to the heterogeneity of CRC observed in inter- and intra-tumoural settings (3). As the classical model of colorectal carcinogenesis (91), the adenoma-carcinoma sequence highlights the mutation of the *APC* gene, a tumour suppressor gene, linking to the initial adenoma formation and the subsequent accumulation of additional genetic mutations, chromosomal instability (CIN), and microsatellite stable (MSS) status, including key genes such as *KRAS*, *PIK3CA*, *TP53*, and *SMAD4*. As a result, diverse signalling mechanisms of the Wnt signalling pathway (APC/glycogen synthase kinase 3 beta (GSK-3β)/β-catenin/AXIN), RAS signalling pathway via the EGFR-RAS-RAF-MEK-MAPK (ERK) pathway, phosphatidylinositol 3-kinase (PI3K)-AKT-mTOR signalling, the TP53 pathway, and the SMAD pathway as part of TGF-β signalling lead to the promotion of tumour growth (2, 3, 19-21). Adenoma polyps are traditional tubular adenomas with an 85-95% incidence of sporadic CRC (2). Apart from somatic mutations in the APC tumours that present in the majority of CRC, 1% of CRC is associated with germline mutations in the *APC* gene occurring in FAP (3). Adenocarcinomas are classified in most CRCs and subdivided into low-

grade and high-grade tumours. However, CRC also presents rarer histological subtypes: mucinous adenocarcinoma, adenosquamous carcinoma, signet-cell carcinoma, medullary carcinoma, micropapillary, serrated, cribriform comedo-type, adenosquamous, spindle cell, and undifferentiated (3, 92).

Another alternative pathway, the serrated polyps with sawtooth morphology, including hyperplastic polyps (HP), traditional serrated adenomas (TSA), sessile serrated lesions (SSL), and mixed polyps, is defined with a 10-15% incidence of CRC (2, 3, 19). With heterogeneous lesions in the mitogen-activated protein kinase–ERK (MAPK/ERK) signalling pathways observed, mutation of the oncogenic *BRAF* gene (activating V600E mutation) is suggested to be linked to the initial activation of most serrated tumour development. In addition to acquiring activation of MAPK/ERK signalling, microsatellite instability (MSI), and the CpG island methylator phenotype (CIMP) of MMR genes, such as MutL protein homolog 1 (*MLH1*), O-6-methylguanine DNA Methyltransferase (*MGMT*), and the cyclin-dependent kinase inhibitor 2A (*CDKN2A*) encoded p16 protein, these factors can drive two distinct routes of serrated lesion development. The first route involves the MSS pathway. The BRAF mutant tumour acquires additional mutations, such as *APC*, *KRAS*, and *TP53*, as well as activation of several pathways, such as Wnt signalling, TGF-β signalling, and the epithelial-to-mesenchymal transition (EMT), to promote the development of traditional serrated adenomas (MSS/CIMP$^{+/-}$). The other route, through the MSI pathway, involves hypermethylation on the promoter of MMR genes, which leads to the development of MSI-high tumours with or without BRAF and KRAS mutations (MSI-H/CIMP-H), commonly found in sessile serrated lesions (3, 19, 20). Additionally, Lynch syndrome is known as the most common hereditary colon cancer syndrome, accounting for 2-5%. The tumour develops through the MSI pathway, which is related to germline mutation of MMR genes (MSI-H/CIMP^{-}) (2, 3, 19, 20).

Acquiring various genetic and epigenetic mutations in pathway components contributes to the dysregulation of several key signalling pathways during CRC carcinogenesis. The Wnt/β-catenin and MAPK/ERK signalling pathways, associated with *KRAS* and *BRAF* mutations, are frequently implicated in the cellular proliferation and differentiation of emerging mutant cells (2, 93, 94). Gaining a more progressive advantage is linked to a deviant PI3K/AKT/mTOR cascade that facilitates cellular growth, survival, and metabolism (2, 95). This process also

involves a shift in the TGF-β/BMP/SMAD pathway's role, transitioning from a tumour suppressor in early neoplastic events to an enhancer that promotes aggressive features related to EMT and metastasis in later stages (2, 18). Additionally, the disruption of normal stem cell regulation through Notch, Hedgehog, and Hippo signalling pathways may alter cellular processes in CRC progression, including proliferation, differentiation, apoptosis, migration, invasion, EMT, and angiogenesis (2, 96-98). Furthermore, tumourigenesis is also influenced by inflammatory processes in response to cytokines and growth factors, as well as DNA damage from both present and absent carcinogens. The TNF, NF-κB, and JAK/STAT signalling pathways play a critical role in inflammation-associated CRC, contributing to the tumour microenvironment (TME), angiogenesis, and immune escape mechanisms (2, 99).

Moreover, the contribution of epigenetic alterations, which consist of DNA methylation, histone modifications, and non-coding RNAs (ncRNAs), disrupts the balance of cellular homeostasis. These mechanisms potentially regulate the dynamics of gene expression and signalling pathways involved in cell cycle control, apoptosis, DNA repair, and cell adhesion in CRC (2, 100). DNA polymerase proteins POLE and POLD1, specifically single-nucleotide variants, cause errors in DNA repair, resulting in a hypermutated phenotype, which has been suggested as a new molecular pathway (hypermutated-single nucleotide variants; HM-SNV pathway) (19, 20).

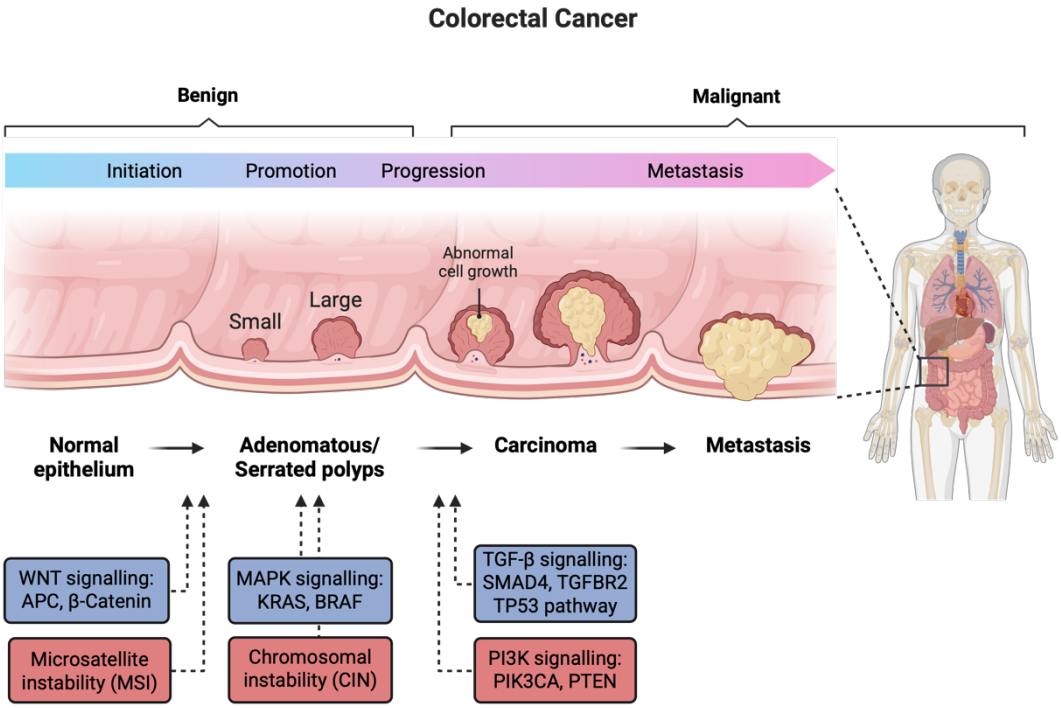


Figure 1-6 Pathogenesis and key mutations in colorectal cancer

The diagram illustrates the stages of colorectal tumour development: initiation, promotion, progression, and metastasis. These stages are associated with various mutations that drive the progression from normal epithelium to polyps, carcinoma, and metastasis. This figure was created using BioRender.com and adapted from Markowitz SD. and Bertagnolli MM., 2009.

1.4 Cell signalling pathways in the gut

1.4.1 Regulatory cell signalling pathways in healthy intestinal crypts

The intestinal crypts have been previously described as a niche for various intestinal cell types, including intestinal stem cells (ISCs) and differentiated cells (intestinal epithelial cells), supported by connective tissue that forms the villus structure. In addition to the multipotency for differentiation and self-renewal of stem cells, there exists a high level of plasticity among the cells, and numerous sources of growth factors are available from both the epithelium and the surrounding stroma within their microenvironment. These factors control cell fate specification and differentiation for tissue homeostasis and the response to tissue damage. Several essential factors and signals are highlighted in the intestinal crypts (12, 14, 101), see Figure 1-7.

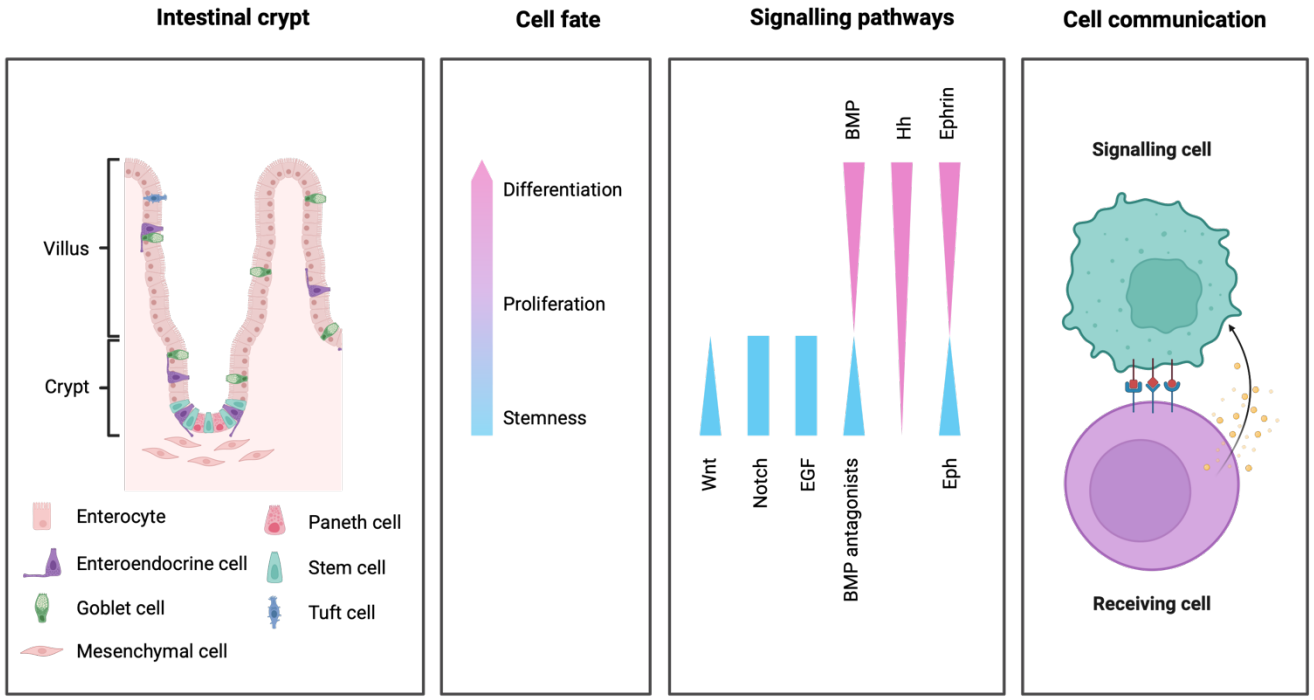

Intestinal crypt
Villus
Crypt
Enterocyte
Enteroendocrine cell
Goblet cell
Mesenchymal cell
Paneth cell
Stem cell
Tuft cell
Cell fate
Differentiation
Proliferation
Stemness
Signalling pathways
BMP
Hh
Ephrin
Wnt
Notch
EGF
BMP antagonists
Eph
Cell communication
Signalling cell
Receiving cell
WNT
Notch
EGF
BMP
Hh
EPH/Ephrin
Neighbor cell
Wnt
LRP
Frizzled
Cytoplasm
Dishevelled
GSK-3β
CKIα
Axin
APC
β-Catenin
Nucleus
β-Catenin
TCF/LEF
Gene transcription
Delta
Notch
NICD
Co-A
MAM
CSL
Target genes
EGFR receptor
GRB2
SOS
p85
p110
Ras
Raf
MEK
ERK
BMP
SMAD1/5/8
SMAD4
SMAD6
Hedgehog
PTCH
SMO
SUFU
GLI1
Ephrin
EPH
Kinase
TF

Figure 1-7 Overview of cell fate and signalling pathways in the regulation of intestinal crypts

The diagrams illustrate the intestinal crypt and its connection to cell fate, signalling pathways, and cell communication, such as WNT, Notch, EGF, BMP, Hh, EPH/Ephrin pathways, which are crucial for regulation within the intestinal crypt. This figure was created using BioRender.com and adapted from Spit M. et al. 2018; Malijauskaite S. et al., 2021; and Li Q. et al., 2024.

Wnt signalling

Wnt signalling is a crucial pathway in the classical signalling mechanisms of homeostatic stem cells that regulate the balance between self-renewal and lineage commitment of crypt-based columnar cells (CBCs). This signalling is activated in CBCS, and the surrounding epithelial and stromal cells provide WNT ligands to the stem cells in a gradient toward the villi. The canonical pathway, referred to as Wnt/ß-catenin, involves WNT ligand proteins binding to Frizzled (FZD) and low-density lipoprotein receptor-related protein 5 or 6 (LRP5/LRP6), which induces the dimerisation of a heterodimeric complex. The dimeric receptors recruit Dishevelled (DVL) to disrupt the formation of the complex comprising glycogen synthase kinase 3ß (GSK-3ß), Axin, adenomatous polyposis coli (APC), and casein kinase I α (CKIα), along with ß-catenin. Subsequently, preventing the phosphorylation of ß-catenin and inhibiting its degradation causing ß-catenin to accumulate in the cytoplasm. The active form of ß-catenin (dephosphorylated) is then translocated to the nucleus to associate with the transcription factor T cell factor/lymph enhancer factor 1 (TCF/LEF1) and ultimately activate WNT target gene expression, involving cell fate determination, proliferation, survival, differentiation, and migration (12, 14, 93, 102), see Figure 1-8.

R-spondin is an enhancer of WNT signalling secreted by mesenchymal cells. It binds to leucine-rich repeat-containing G protein-coupled receptors 4, 5, and 6 (LGR4/5/6) and forms a complex with the E3 ubiquitin ligases RNF43 and ZNRF3, inhibiting a negative feedback loop in the WNT pathway via endocytosis, which results in increased WNT pathway activation for stem cell maintenance. In contrast, Dickkopf 1 (DKK1) acts as a WNT antagonist that inhibits the WNT signalling cascade (12, 93). Additionally, the non-canonical signalling pathway operates independently of ß-catenin-TCF/LEF, regulating both transcriptional and non-transcriptional responses in cells, which are associated with the Planar cell polarity (WNT-PCP) pathway and WNT-$Ca2^{+}$ signalling pathway. Therefore, WNT signalling plays an essential role in cell fate decisions, and is crucial in various

biological processes, including; embryonic development, stemness, autophagy, apoptosis, metabolism, inflammation, immunisation, and microenvironment interactions (93).

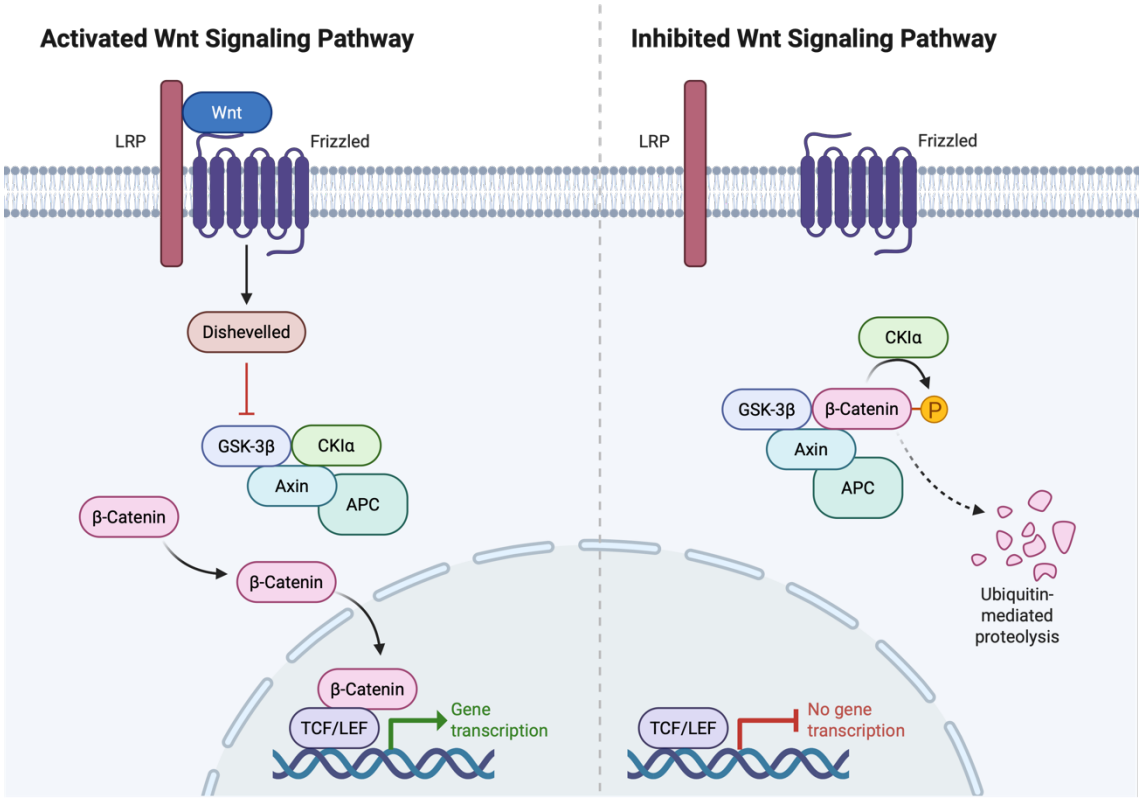


Figure 1-8 Classical Wnt signalling pathway

The diagram illustrates the activated Wnt signalling pathway, in which Wnt ligand proteins bind to Frizzled and the low-density lipoprotein receptor-related protein (LRP). This binding recruits Dishevelled (DVL) to disrupt the complex composed of GSK-3β, Axin, APC, and CKIα, as well as β-catenin. This process prevents β-catenin from being phosphorylated and degraded, enabling dephosphorylated β-catenin to translocate to the nucleus for gene regulation. This figure was created using BioRender.com.

Notch Signalling

At the bottom of intestinal crypts, physical contact between two cells facilitates Notch Signalling, one of the classical signalling mechanisms in cell fate decisions and the maintenance of the progenitor/stem cell population. A ligand-presenting cell, such as Paneth cells or deep crypt secretory cells (DSCS), expresses Notch ligands, which include Jagged-1 and 2 (JAG1-2) and Delta-like 1, 3, and 4 (DLL1, DLL3, and DLL4), to trigger Notch receptors (Notch 1–4) on a receiving receptor-producing cell, such as CBC stem cells. The heterodimerisation domain of Notch receptors on the cell membrane is cleaved in the Golgi apparatus (S1 cleavage).

Classically, the binding of extracellular domains of Notch receptors with ligands initiates endocytosis of ligands, activates the ADAM family of metalloproteases (S2 cleavage), and the γ-secretase (S3 cleavage) to cleave the transmembrane Notch receptors, leading to the release of the Notch intracellular domain (NICD). The translocation of NICD to the nucleus displaces co-repressors and forms a complex with other transcriptional regulators CBF-1/suppressor of hairless/Lag1 (CSL, also known as recombination signal binding protein-J, RBPJ), Mastermind-like protein (MAML-1), and p300/CBP, which recruits co-activators (CoA), inducing the expression of downstream target genes such as Hairy and enhancer of split gene (HES1, HES5, HES6, HES7), Hairy/Enhancer of Split related to YRPW motif (HEY1), MATH1, p21, and DELTEX to maintain stem cell identity, which regulates cell differentiation and proliferation in cooperation with Wnt signalling (12, 15, 96, 103), see Figure 1-9.

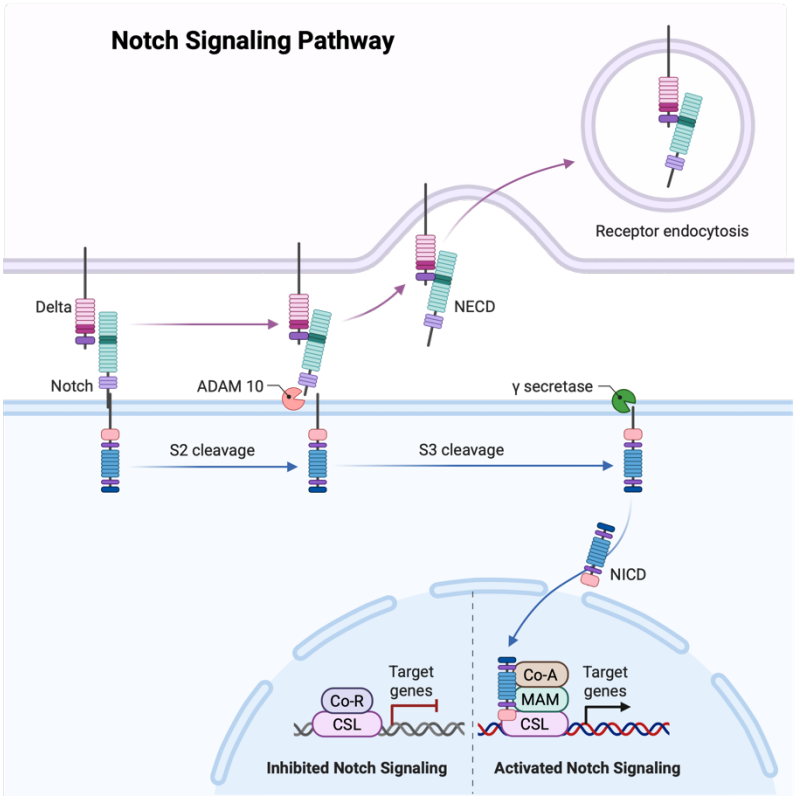


Figure 1-9 Classical Notch signalling pathway

The diagram illustrates the activated Notch signalling pathway, in which a ligand-presenting cell presents Notch ligands to activate Notch receptors on a target cell. This process involves the endocytosis of the extracellular domains of Notch receptors along with ligands, as well as the action of ADAM metalloproteases (S2 cleavage) and γ-secretase (S3 cleavage), which cleave the transmembrane Notch receptors. This cleavage releases NICD, which then translocates into the nucleus to form a complex with other transcriptional regulators. This figure was created using BioRender.com.

Epidermal Growth Factor (EGF) Signalling

EGF is a classical signalling mechanism in intestinal crypts that regulates cellular proliferation and differentiation, intracellular homeostasis, and tissue regeneration, thereby ensuring the integrity of the intestinal barrier through junctional complexes. EGF ligands are produced in intestinal cells and act in an autocrine and/or paracrine manner on neighboring target cells at the base of crypts that possess EGF receptors (EGFR). The EGF family of ligands shares similar EGF-like motifs, including EGF, TGF-α, amphiregulin (AREG), epigen (EPGN), heparin-binding EGF-like growth factor (HB-EGF), epiregulin (EREG), betacellulin (BTC), and the neuregulins (NRG1-4). EGFR members are referred to as the Erbb subclass of the Receptor Tyrosine Kinase (RTK) superfamily, which includes ErbB1 (EGFR), ErbB2 (HER2), ErbB3 (HER3), and ErbB4 (HER4). Binding of ligands to EGFR activates receptor dimerisation (homodimers or heterodimers), subsequently inducing the intrinsic kinase domain to phosphorylate specific tyrosine residues within the cytoplasmic tails of the receptors (trans-autophosphorylation) and recruiting molecules and regulatory proteins involved in intracellular signalling, including; Ras/MAPK, PI3K-AKT, and JAK-STAT. Therefore, these complex signalling cascades regulate various key cellular functions, including cell proliferation, cell survival, and stem cell maintenance (12, 14, 104-106), see Figure 1-10.

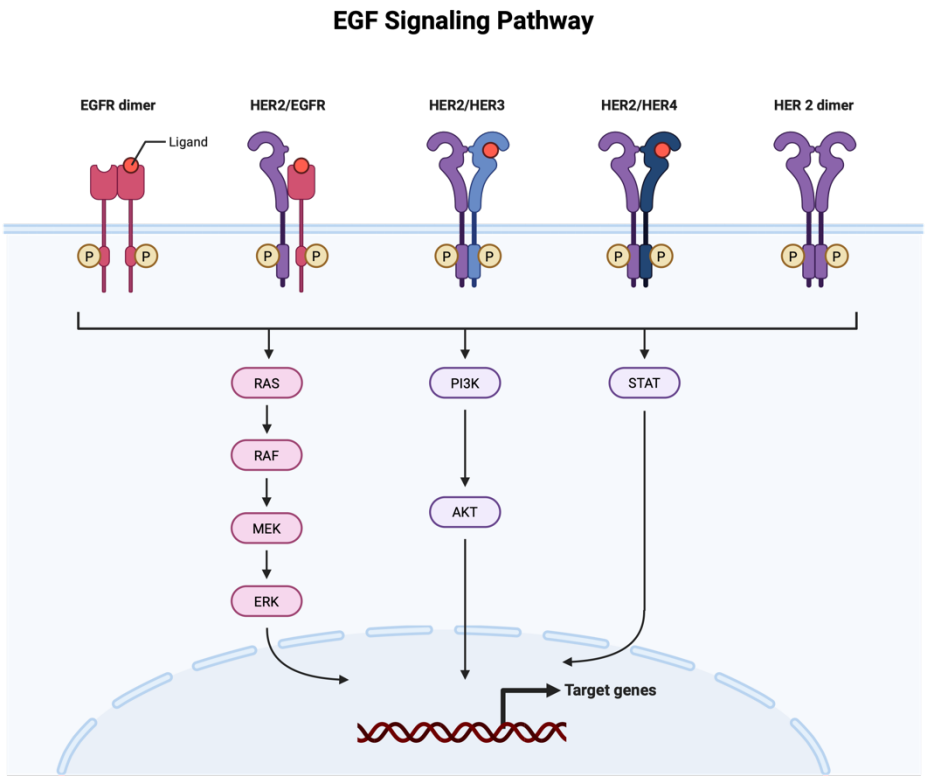


Figure 1-10 Classical EGF signalling pathway

The diagram illustrates the activation of the EGF signalling pathway, in which EGF ligands bind to EGF receptors (EGFR). Members of the EGFR family are part of the Erbb subgroup within the Receptor Tyrosine Kinase (RTK) superfamily, comprising ErbB1 (EGFR), ErbB2 (HER2), ErbB3 (HER3), and ErbB4 (HER4). When ligands bind to EGFR, they induce receptor dimerisation (either homodimers or heterodimers), leading the intrinsic kinase domain to phosphorylate specific tyrosine residues, thereby activating the Ras/MAPK, PI3K-AKT, and JAK-STAT signalling pathways. This figure was created using BioRender.com.

Bone Morphogenetic Protein (BMP) Signalling

Along the crypt-villus axis, BMP signalling is a classical pathway that increases toward the villus, contrasting with local Notch signals and Wnt and EGF gradients found at the base of the crypts. BMP signalling is part of the TGF-β superfamily. The canonical BMP signalling pathway begins with BMP ligands, such as BMP2 and BMP4, binding to BMP receptors on the cell surface and forming a heterotetrameric complex consisting of type I (BMPR1A/B) and type II (BMPR2) serine/threonine kinase receptors. Subsequently, the active type II receptor transphosphorylates the type I receptor at a glycine-serine-rich motif (GS domain) to initiate the phosphorylation of the receptor-regulated SMADS (R-SMADS, including SMAD1, SMAD5, and SMAD8; SMAD1/5/8) at a C-terminal SSXS motif prior to forming a complex with co-mediator SMAD (co-SMAD), also known as SMAD4. The complex

translocates to the nucleus and acts as a transcription factor with coactivators and corepressors to regulate gene expression, involved in cell growth, apoptosis, and differentiation, Figure 1-11. This pathway plays an essential role in embryogenesis, development, maintenance of tissue homeostasis and is important for facilitating crypt formation in the intestine. The intracellular feedback inhibition of the BMP signalling cascade is associated with inhibitory SMADS (I-SMADS, including SMAD6 and SMAD7; SMAD6/7). BMP ligands are secreted by mesenchymal cells at the villus tip, while BMP antagonists, such as Gremlins, Chordin, and Noggin, are produced by stromal cells, including myofibroblasts and smooth muscle cells, near the bottom of the crypts. Additionally, several non-canonical pathways of BMP are identified as being involved in SMAD-independent signalling, such as activating TAK-1, a serine-threonine kinase in the MAPK pathway, as well as PI3K/Akt, PKC, Rho-GTPases, and others, depending on cellular activity and extracellular environment. Furthermore, BMP signalling can also crosstalk with other pathways, such as Wnt signalling (11-15).

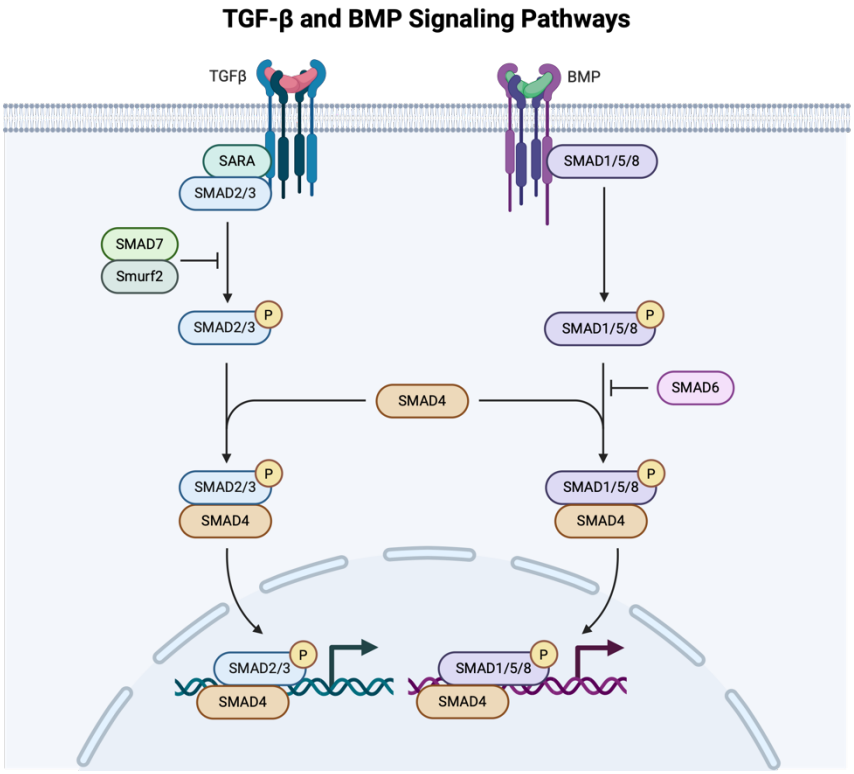


Figure 1-11 Classical BMP signalling pathway as a part of TGF-β signalling pathway

The diagram illustrates that the BMP signalling pathway is activated, with BMP ligands binding to BMP receptors to form a heterotetrameric complex consisting of type I (BMPR1A/B) and type II (BMPR2) serine/threonine kinase receptors. This triggers the phosphorylation of receptor-regulated SMADS (SMAD1, SMAD5, and SMAD8), which then combine with the co-mediator SMAD (SMAD4).

The complex translocates to the nucleus and regulates gene expression. Likewise, TGF-β ligands bind to TGF-β receptors, forming a heterotetrameric complex that leads to the phosphorylation of SMAD2 and SMAD3, which then form a complex with SMAD4 to regulate gene expression. This figure was created using BioRender.com.

Hedgehog (Hh) signalling

Intestinal epithelial cells in the crypt-villus express Hedgehog (Hh) ligands, including sonic hedgehog (Shh), Indian hedgehog (Ihh), and desert hedgehog (Dhh), which bind to the receptors patched 1 and 2 (PTCH1/2) located in mesenchymal cells (recipient cells). Upon binding, the seven-span transmembrane protein smoothened (SMO) is released, resulting in the inactivation of the two negative pathway regulators, protein kinase A (PKA) and suppressor of fused (SUFU). Subsequently, the glioblastoma family of transcription factors (GLI1/2/3) is activated, and the activated GLI2 and/or GLI3 translocate to the nucleus, driving the expression of Hedgehog target genes to regulate the development and homeostasis of the mesenchymal compartment, Figure 1-12. Due to dynamic epithelial-mesenchymal interactions, the secretion of Hedgehog ligands from epithelial cells facilitates their proliferation and regulates BMP ligand production from the surrounding stroma. Decreased Wnt secretion from epithelial and stromal cells increases Hedgehog signalling in mesenchymal cells, indirectly promoting the maintenance of intestinal stem cells (14, 15, 97).

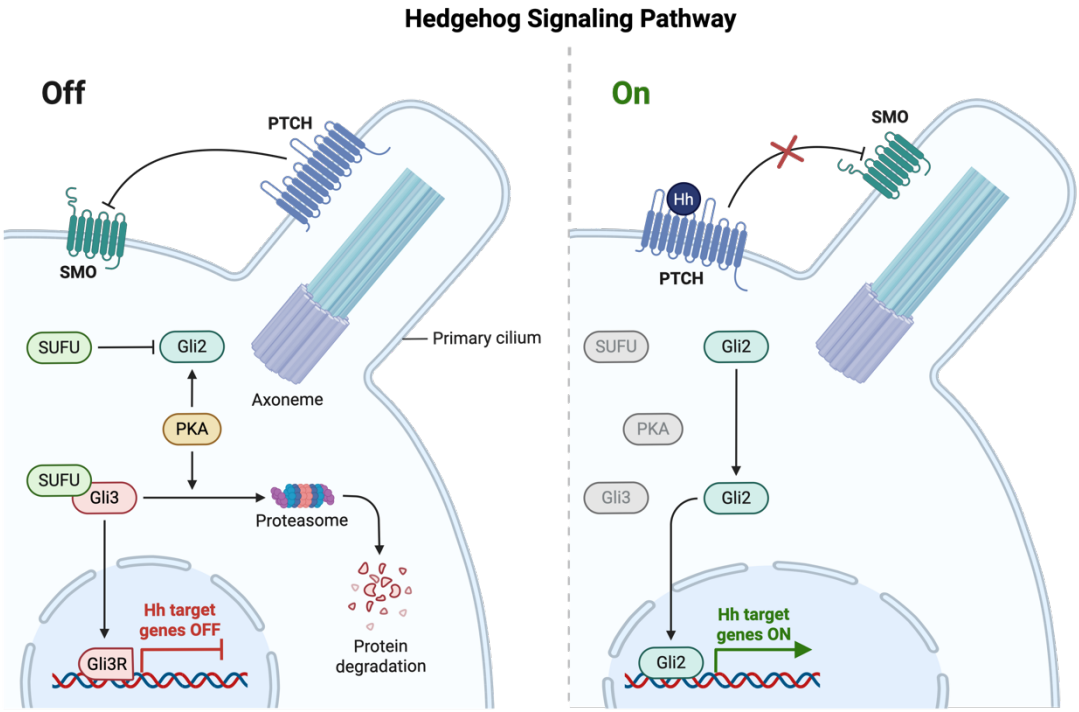


Figure 1-12 Classical Hedgehog signalling pathway

The diagram illustrates that the Hh signalling pathway is activated, with Hh ligands binding to the receptor, PTCH. SMO is then released, inactivating PKA and SUFU and preventing protein degradation. Subsequently, activated GLI2 translocates to the nucleus, driving gene expression. This figure was created using BioRender.com.

EPH/Ephrin Signalling

Along the crypt-villus axis, the Erythropoietin-producing hepatocellular (EPH) receptor tyrosine kinases (RTKs) and Ephrin signalling gradient serve as crucial moderators in controlling cell position, proliferation, differentiation, adhesion, and migration of intestinal epithelial cells. The EPH family of receptor tyrosine kinases includes EPHA (1–8, 10) and EPHB (1–4, 6) and can be activated by membrane-associated ephrin ligands, including glycosylphosphatidylinositol (GPI)-linked ephrin-A (1–5) or transmembrane ephrin-B (1–3) through cell-cell contact. However, it can also signal in a ligand-independent and context-dependent manner. Once the ligand binds to the receptor, it triggers kinase activation, the recruitment of effector molecules, and the transmission of various downstream signalling cascades, Figure 1-13. In the intestinal crypt, high expression of EPHB2 and EPHB3 is present in intestinal stem cells, which is simultaneously reduced in a gradient of EPHB2 from the transition zone, facilitating cell compartment organisation and the migration of differentiated cells toward the apical villus. While Ephrin-B1 and Ephrin-B2 are repressed in expression at the base of crypts, this gradient is regulated by Wnt signalling through TCF/β-catenin complexes (15, 107).

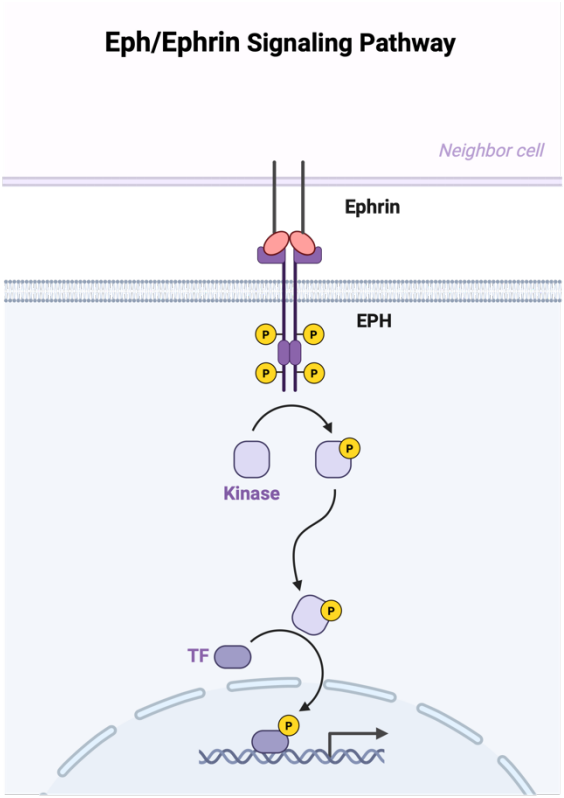


Figure 1-13 Classical EPH/Ephrin signalling pathway

The diagram illustrates that the Eph/Ephrin signalling pathway is activated, with ephrin ligands binding to the receptor, EPH. This initiates kinase activation, recruits effector molecules, and transmits signalling cascades that regulate gene expression. This figure was created using BioRender.com.

Other signalling pathways

In addition to its role in intestinal homeostasis, Hippo Signalling is an essential pathway for tissue self-renewal and regeneration. Briefly, activated Hippo signalling initiates the activation of STE20-like protein kinases 1 and 2 (MST1/2), binding with Salvador homologue 1 (SAV1) to phosphorylate large tumour suppressors 1 and 2 (LATS1/2) with NF2 interaction, as well as the phosphorylation of MOB kinase activators 1A and 1B (MOB1A/B). Activated LATS1/2 phosphorylates transcriptional coactivators Yes-associated protein (YAP) and transcriptional coactivator with PDZ-binding motif (TAZ), leading to phosphorylated YAP/TAZ binding to 14-3-3 protein, resulting in cytoplasmic retention or degradation by ß-transducin repeat-containing protein (ß-TrCP). When inactive, the Hippo signalling cascade enables unphosphorylated YAP/TAZ to translocate to the nucleus, competing with VGLL4 to bind to TEA domain family member 1 (TEAD1) transcription factors, thereby promoting the expression of target genes

involved in cell growth and survival, Figure 1-14. The interaction of Hippo signalling with classical signalling pathways in the crypts is associated with the reduction of WNT signalling through the repression of target gene expression, the induction of antagonists such as DKK1, BMP4, and WNT5α, the regulation of Notch signalling via HES1 gene expression, and the moderation of EGF signalling through the expression of EGFR ligands (108-110).

In addition EGF signalling pathway contributes to tissue homeostasis in the intestinal crypts, with important downstream intracellular signalling cascades involved in MAPKs, PI3K-AKT, and JAK-STAT signalling (104, 106). MAPK pathways can be activated by several stimuli, including growth factors, hormones, cytokines, and environmental stress, leading to diverse levels of response. The MAPK transduction cascade consists of multiple molecules, starting with the phosphorylation of MAP kinase kinase kinase (MAPKKK, such as RAF) by members of the RAS and RHO families of small GTPases. The phosphorylation loop is subsequently activated in MAP kinase kinase (MAPKK, such as MEK1/2) and MAP kinase (MAPK, such as ERK1/2) in the cytoplasm, before active MAPK is translocated into the nucleus to specific docking sites on substrates. The major MAPK pathways include the extracellular-regulated kinases 1 and 2 (ERK1/2), c-Jun NH2-terminal kinases (JNKs), stress-activated protein kinases (SAPK), p38, MAPK14 and ERK5 pathways (94, 111, 112), Figure 1-14. Additionally, MAPK cascades commonly interact with other signalling pathways, such as PI3K-AKT, TGF-ß/SMAD, and WNT/ß-catenin pathways, generating a complex signalling network that supports differential cellular responses for a broad range of cellular processes (108, 112).

Activation of the PI3K/AKT signalling pathway involves various upstream components. Different ligands trigger specific transmembrane receptors, such as the RTK family- a transmembrane protein family with intrinsic phosphotyrosine kinase activity, including EGFRs, VEGFRs, and fibroblast growth factor receptors (FGFRs), G protein-coupled receptors (GPCRs), or other types of receptors, leading to the activation of PI3Ks, plasma membrane-associated lipid kinases, which consist of three classes: Class I, heterodimers of the p110 catalytic subunit with four isoforms (p110α, p110ß, p110γ, and p110δ) and a p85 regulatory subunit, encoded by PIK3CA, PIK3CB, PIK3CG, and PIK3CD, and also divided into classes IA

and IB; Class II, monomers of three catalytic isoforms, PI3KC2α, PI3KC2β, and PI3KC2γ; and Class III, PIK3C3, also referred to as VPS34. Subsequently, active PI3Ks phosphorylate phosphatidylinositol-4,5-bisphosphate (PIP2) to generate phosphatidylinositol-3,4,5-trisphosphate (PIP3), which can be reversed by phosphatase and tensin homolog (PTEN). As a result, PIP3 recruits kinases with a pleckstrin homology domain (PH domain) to the membrane. The serine and threonine kinase AKT, also referred to as protein kinase B (PKB), is phosphorylated and activated by phosphoinositide-dependent protein kinase 1 (PDK1). Additionally, AKT can be activated through RAC (Rac-1, a small GTPase) and the mechanistic target of rapamycin (mTOR) Complex 2. Therefore, a number of downstream effectors are promoted or inhibited by AKT, such as the mTORC1 complex, Mouse double minute 2 homolog (MDM2), Glycogen Synthase Kinase-3 (GSK3), Forkhead box Os (FOXOs), and tuberous sclerosis (TSC) (95, 113-115), Figure 1-14.

Another pathway of central signalling mediators in various cellular functions, which responds to multiple cytokines and growth factors, is the Janus kinase/signal transducer and activator of transcription (JAK/STAT) signalling pathway. In the canonical pathway, the binding of the ligand to its corresponding receptor leads to receptor dimerisation, and the ligand-receptor complex induces the transphosphorylation of JAK (JAK1, JAK2, JAK3, and TYK2). Following JAK-mediated tyrosine phosphorylation of the receptors, a docking site is formed for STAT (STAT1-4, STAT5a, STAT5b, and STAT6). Subsequently, the STATs are phosphorylated; the activated STATs then dissociate from the receptor to form a dimer, translocate to the nucleus, bind to DNA, and regulate transcription, Figure 1-14. Additionally, the negative regulation of the JAK/STAT signalling pathway is related to three main groups of proteins: the protein inhibitor of activated STAT (PIAS), the suppressor of cytokine signalling (CIS/SOCS) family, and protein tyrosine phosphatases (PTPs). In the noncanonical pathway, signal transduction is more complex, as STAT can be activated by other non-receptor tyrosine kinases or directly by receptors independent of JAK, such as EGFR. Furthermore, the crosstalk of JAK/STAT signalling with other pathways, including TGF-β, MAPK, Notch, PI3K/AKT/mTOR, and NF-κB, contributes to the regulation of cellular processes and the response to inflammation (116).

Regarding cellular inflammatory reactions, immune responses, and various biological processes in the intestinal microenvironment, the nuclear factor kappa-light-chain-enhancer of activated B cells, known as NF-κB signalling, represents a crucial pathway that effectively responds to various extracellular stimuli, including TNF-α, interleukin (IL)-1ß, lipopolysaccharide (LPS), and antigens. Following the canonical pathway of signal transmission, ligands bind to their receptors, triggering the activation of multiple protein complexes. These protein kinase complexes activate the I-kappaB kinase (IKK) complex (composed of IKKα, IKKß, and IKKγ (NEMO)) through phosphorylation, ultimately leading to the subsequent phosphorylation of IκBα as part of the NF-κB complex (composed of IκBα, p50, and RELA) before its ubiquitination and proteasomal degradation. Consequently, the release of p50/RELA becomes a transcription factor that activates the transcription of target genes, Figure 1-14. In non-canonical NF-κB signalling, after the ligand binds to the receptor, it activates NF-κB-inducing kinase (NIK), which in turn activates IKKα, leading to the degradation of p100 to p52. The p52 binds to RELB and translocates to the nucleus. Furthermore, the signal transduction network of NF-κB signalling interacts with other signalling pathways, including PI3K/AKT, MAPK, JAK-STAT, TGF-ß, Wnt, Notch, and Hedgehog signalling, which are involved in cell proliferation, differentiation, survival, death, development, immunity, inflammation, and even tumourigenesis (108, 117).

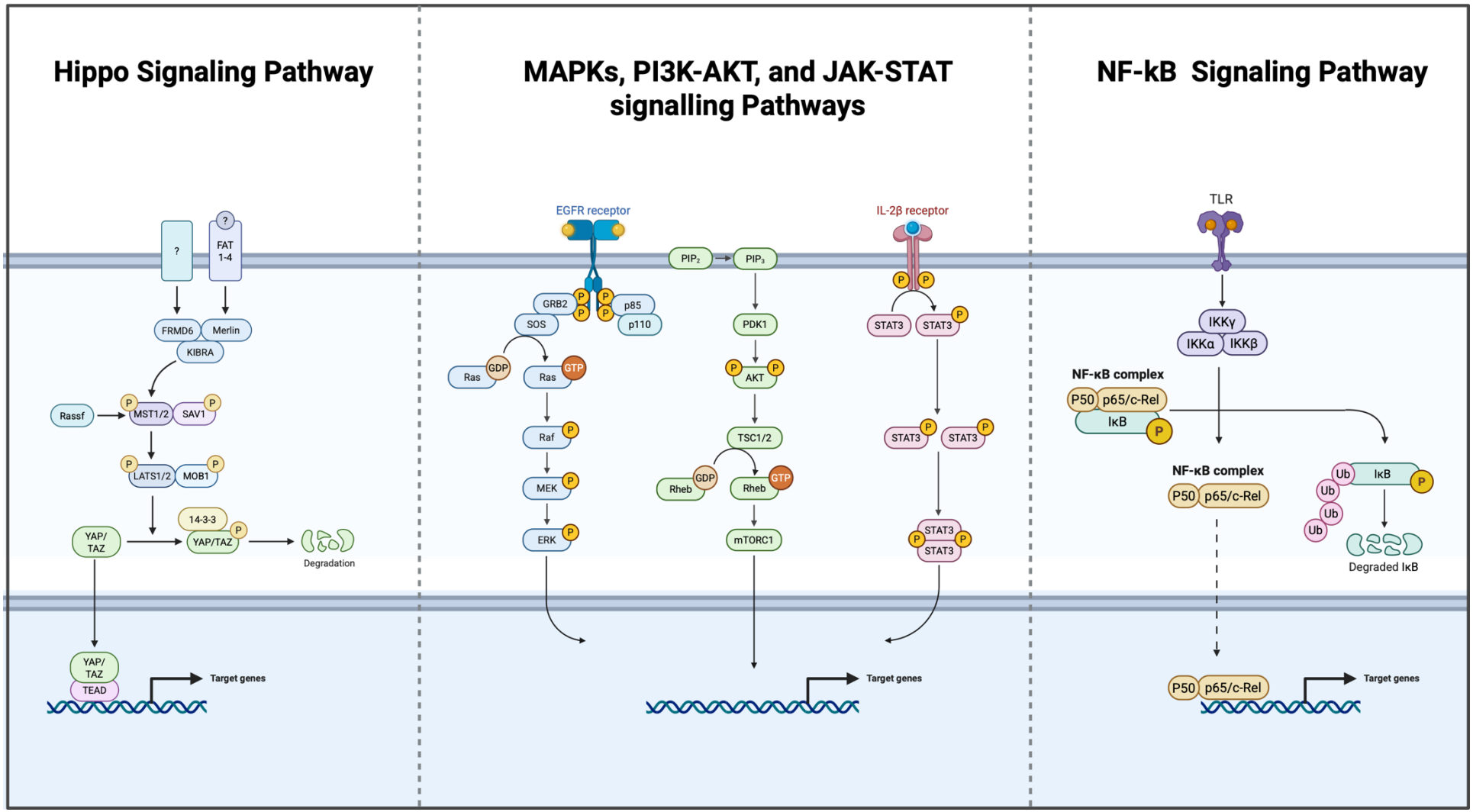

Hippo Signaling Pathway
MAPKs, PI3K-AKT, and JAK-STAT signalling Pathways
NF-kB Signaling Pathway
?
FAT 1-4
FRMD6
Merlin
KIBRA
Rassf
MST1/2
SAV1
LATS1/2
MOB1
YAP/ TAZ
14-3-3
YAP/TAZ
Degradation
YAP/ TAZ
TEAD
Target genes
EGFR receptor
GRB2
SOS
p85
p110
Ras
GDP
Ras
GTP
Raf
MEK
ERK
PIP2
PIP3
PDK1
AKT
TSC1/2
Rheb
GDP
Rheb
GTP
mTORC1
Target genes
IL-2β receptor
STAT3
STAT3
STAT3
STAT3
STAT3
STAT3
TLR
IKKγ
IKKα
IKKβ
NF-κB complex
P50
p65/c-Rel
IκB
NF-κB complex
P50
p65/c-Rel
Ub
IκB
Degraded IκB
P50
p65/c-Rel
Target genes

Figure 1-14 Other signalling pathways, including Hippo, MAPKs, PI3K-AKT, JAK-STAT, and NF-κB signalling.

The diagrams illustrate an overview of the activation of various signalling pathways that contribute to maintaining intestinal homeostasis, including the Hippo signalling pathway, the EGF signalling pathway, which mediates intracellular signalling cascades involved in MAPKs, PI3K-AKT, and JAK-STAT signalling, as well as NF-κB signalling pathways. This figure was created using BioRender.com.

1.4.2 Dysregulation of cell signalling pathways in CRC development and progression

As previously described, the intricate crosstalk among multiple signalling pathways significantly regulates cellular processes in the intestinal cryps both maintaining stem cells, their differentiation and the functionality of differentiated intestinal cells. These signals are also important for the interactions of epithelial cells with their microenvironment, including fibroblasts and immune cells, to maintain intestinal homeostasis and regeneration. Any alteration of these pathways leads to functional changes that facilitate the development of CRC, cancer progression, metastasis, and even therapeutic resistance (2, 118).
In initiating oncogenic processes, the WNT/ß-catenin signalling pathway is recognised as a key driver of CRC development (93). Mutations in *APC* are associated with a high incidence and significantly impact the alteration of the WNT/ß-catenin signalling cascade, which interacts with other signalling pathways, such as the Notch and RAS/RAF/MAPK pathways during early tumourigenesis (2, 19, 20, 118). Activation of the WNT/ß-catenin signalling pathway mediates the activation of Jagged-1, which in turn induces the Notch signalling cascade (119). Concurrent mutations of WNT/ß-catenin and mediators of the MAPK/ERK signalling cascade, including KRAS and BRAF, accelerate processes in carcinogenesis, as both promote cell cycle progression through the dysregulation of cyclin D1 and c-Myc (2, 19, 20, 120, 121).

Additionally, increased activation of EGFR leads to a significant intracellular signalling cascade involving the MAPK and PI3K/AKT pathways, which promote cancer development and progression (2, 120). Dysregulation of the RAS/RAF/MEK/ERK pathway, widely known as a crucial signal transducer controlling the downstream effects of EGFR signalling, promotes proliferation, inhibits apoptosis, degrades ECM proteins, and facilitates angiogenesis, invasion, and metastasis, including proteins such as cyclin D1, MMP1, P53, and c-Myc

(122, 123). As a cooperative network, elevated PI3K/AKT/mTOR signalling leads to aberrant cell metabolism and growth, such as enhancing cellular survival through activated NF-κB, while regulating cell cycle by inactivation of GSK3 leads to increased levels of cyclin D1 and c-Myc (2, 120).

Furthermore, increased activation of Notch signalling leads to changes in cellular processes and tumour physiology by inducing PI3K/AKT signalling activity, which is linked to PTEN inactivation in CRC. Overexpression of EGFR, c-Myc, and Krüppel-like factor 4 (KLF4) triggers tumour progression and contributes to chemoresistance, involving NF-κB activation while interrupting apoptosis through the upregulation of the apoptotic regulators BCL-2 and BCL-XL. Nevertheless, Notch signalling can suppress the activity of TGF-β, which is associated with inhibiting cell growth and acting as a tumour suppressor (96). On the other hand, the context-dependent function of the TGF-β/BMP signalling pathway switches between being a tumour suppressor in early neoplastic events and an enhancer that promotes advanced and aggressive tumour characteristics in later stages of CRC, highlighting interactions with other signalling pathways. The dysregulation of TGF-β/BMP signalling, leads to changes in crucial signalling pathways, including MAPK, PI3K/Akt, Wnt, Hedgehog, and Notch, as well as various cytokines, this complex interplay modulates the alteration of tissue homeostasis, leading to carcinogenesis and contributing to the tumour microenvironment, which involves cell proliferation and differentiation, cell survival, inflammation, angiogenesis, EMT, and eventually metastasis (2, 17, 18, 124, 125).

For instance, EMT is a cellular process that occurs during embryonic development, tissue regeneration, and cancer progression and metastasis. The mechanism involves inducing epithelial cells to change their phenotypes by losing adherent junctions, altering cell polarity, invading the basal lamina, migrating along the fibronectin matrix, reducing epithelial markers (E-cadherin, ZO-1, Claudins, Occludin, Cytokeratins, and Type IV collagen), and increasing mesenchymal markers (Vimentin, FSP-1, α-SMA, Fibronectin, N-cadherin, Type I and Type III collagen). TGF-β/BMP signalling is demonstrated as an essential activator of EMT in cooperation with RTK, WNT, Notch, Hedgehog, and PI3K/AKT, which are related to triggering transcription factors (SNAIL1/2, ZEB1/2, TWIST, LEF-1, and β-catenin) that repress the expression of epithelial markers. Thus, EMT induced by signals from tumour-associated stroma facilitates the migration of invasive cells

from the primary tumour to form a metastatic tumour at distant sites via the circulation (126, 127).

1.5 Metastasis of CRC

1.5.1 Epidemiology of Metastatic CRC

The emergence of metastasis in colorectal cancer has been observed in 20% of patients at their initial diagnosis, and 25% of patients with localised CRC will later develop metastases. After detection, patients with CRC metastasis tend to have shorter survival times: approximately 70% to 75% of patients survive for one year, 30% to 35% survive for three years, and fewer than 20% survive beyond five years from diagnosis (7). Right-sided primary tumours are associated with shorter median survival times compared to left-sided primary tumours (19 months versus 34 months) (128). The metastatic sites of CRC are commonly found in lymph nodes, the liver, the lungs, and the peritoneum, which are typically confirmed by colonoscopy and biopsy and are usually entirely removed by surgery if limited to a few metastatic foci. Additionally, bone and central nervous system metastases can be detected, although this is rare. Due to the multiple technologies available for cancer diagnosis, including contrast-enhanced computed tomography imaging (CECT) of the chest, abdomen, and pelvis, as well as magnetic resonance imaging (MRI) and positron-emission tomography (PET), these advancements may help identify tumours with metastases (7).

An extensive study found approximately 30% of fifty thousand Swedish patients had CRC metastasis, with spread to the liver at 62%, thorax at 8%, peritoneum at 4%, bone at 3%, and the nervous system at 2%. Patterns of metastasis suggest that colon cancer preferentially metastasises within the peritoneum, whereas rectal cancer tends to spread to the thorax, nervous system, and bone. CRC patients with metastasis in the nervous system and bone are associated with the worst outcomes (4 and 5.5 months, respectively) compared to peritoneum, liver and thoracic metastases (7, 9, and 14 months, respectively) (129). Similar metastatic patterns were also observed in a study using the Surveillance, Epidemiology, and End Results (SEER) Programme database (130). However, it is still limited in identifying and characterising subsets of these patients and considering more beneficial and precise treatment management (7, 131).

Regarding the molecular principles of metastasis, this could provide insight into the abnormal regulations involved in the dissemination process. Briefly, the metastatic cascade consists of cancer cells leaving the primary tumour (invasion), transitioning and invading the connective tissue before entering a blood vessel (intravasation), cancer cells circulating in the bloodstream (circulation), circulating tumour cells (CTCs) arresting at secondary sites and extravasating from blood (extravasation), and metastatic tumour cells colonising and adapting cellular mechanisms at new niches (colonisation) (132, 133). Therefore, it is crucial to have a comprehensive understanding of the mechanisms that drive the spread of colorectal metastasis to the targeted metastatic sites, including intra- and intercellular factors such as genetic abnormalities, tumour heterogeneity, alterations in biological processes, EMT, and the tissue microenvironment (7, 131, 132), Figure 1-15.

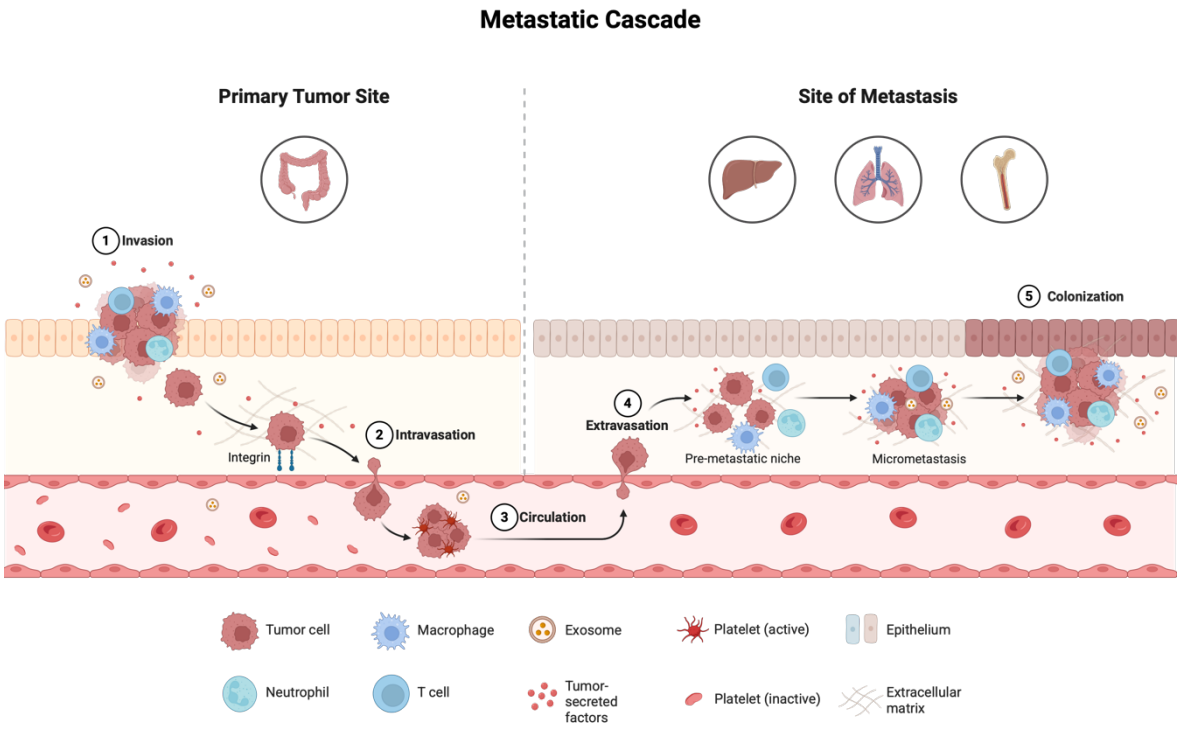


Figure 1-15 Metastatic cascade from the primary tumour to distant sites

The diagram illustrates the metastatic cascade, showing cancer cells leaving the primary tumour in the bowel (invasion), entering a blood vessel (intravasation), circulating in the bloodstream (circulation), extravasating from blood vessels (extravasation), and colonising and adapting cellular mechanisms at new sites, such as the liver, lung, or bone (colonisation). This figure was created using BioRender.com and adapted from Fares J. et al., 2020.

1.5.2 Sites of metastatic CRC

1.5.2.1 Liver

Colorectal liver metastasis (CRLM) is detected in approximately 25% to 50% of CRC patients and may present either synchronously or metachronously (134, 135). Under normal conditions, the liver is the largest organ that transforms, detoxifies, and accumulates metabolites while also contributing them to bile. Additionally, it produces and delivers various plasma proteins, including albumin, to the blood, with oxygenated blood supplied by the hepatic artery and nutrient-rich, deoxygenated blood provided by the portal vein. The liver consists of polygonal lobules containing numerous hepatocytes (parenchymal cells) and a radial arrangement of capillaries and sinusoids, each lobule separated by connective tissue. At the centre of the lobule is a central vein, while bile ducts, lymphatics, nerves, and branches of both the hepatic artery and the portal vein are located at the periphery of the lobule. Apart from hepatocytes, there are various liver sinusoidal cells, including sinusoidal endothelial cells, Kupffer cells (tissue macrophages), hepatic natural killer (NK) cells, and stellate cells, which are supported within the hepatic microenvironment (136), see Figure 1-16.

Histopathological growth patterns (HGP) of CRLM are classified into two categories: desmoplastic growth, characterised by the separation of the tumour from hepatocytes by a thick fibrous rim, and mixed/heterogeneous growth. The tumour immune microenvironment is a significant characteristic of this disease (135). Additionally, genomic abnormalities and a complex molecular mechanism- including non-coding RNAs (ncRNAs), the Notch pathway, TGF-ß signalling, and others- are emphasised in the tumour and its microenvironment in relation to immune cells, cytokines, chemokines, and exosomes, which facilitate metastasis (137). Altogether, this is likely to promise novel approaches in interventional therapy and prognosis for CRLM. Nowadays, a multidisciplinary team, comprising oncologists, surgeons, radiologists, and radiotherapists, collaborates to determine an appropriate treatment for each patient. However, some patients with heterogeneous and unresectable CRLM still require optimal systemic and locoregional chemotherapeutic, biological, and radiotherapeutic treatments to improve their life expectancy (134).

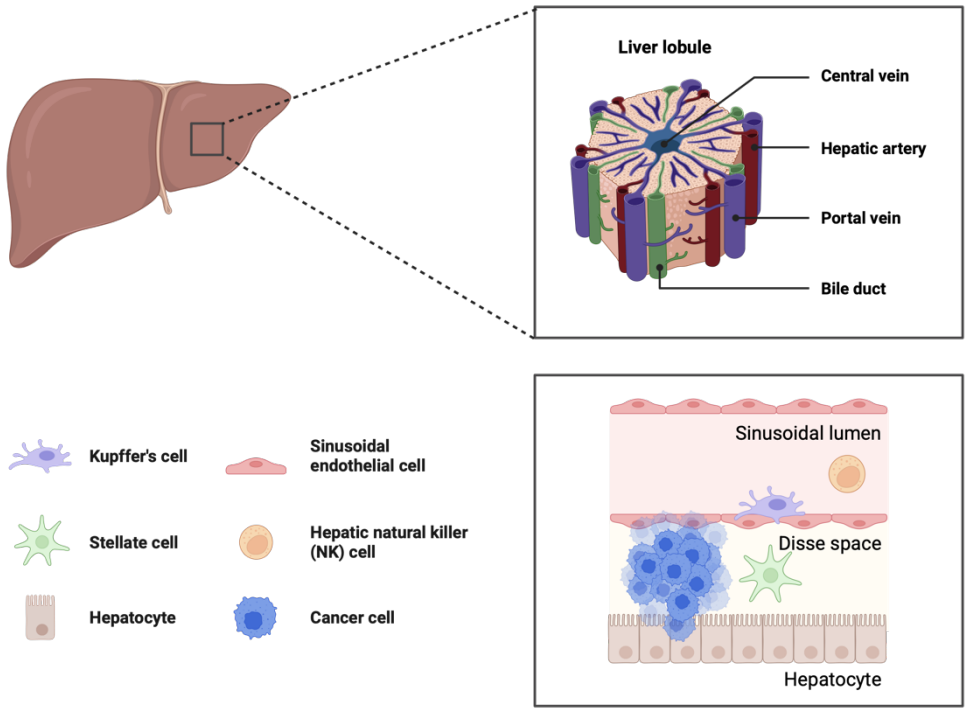


Figure 1-16 Hepatic microenvironment as a target for metastatic CRC

The diagram illustrates the liver and its microenvironment, which consists of polygonal lobules containing numerous hepatocytes, capillaries, and sinusoids. At the centre of each lobule is a central vein, while bile ducts, lymphatics, nerves, the hepatic artery, and the portal vein are located at the periphery. Additionally, various liver sinusoidal cells, including sinusoidal endothelial cells, Kupffer cells, hepatic natural killer (NK) cells, and stellate cells, are found within this hepatic niche. CRC metastasis may disrupt this homeostatic environment. This figure was created using BioRender.com and adapted from Vekemans K. and Braet F., 2005.

1.5.2.2 Peritoneum

The peritoneum, referred to as mesothelium, consists of a monolayer of mesothelial cells (flattened, squamous-like cells, approximately 25 mm in diameter) lining the peritoneal cavity. It is supported by the submesothelial stroma beneath, which contains an extracellular matrix comprised primarily of collagen fibres, fibronectin, glycosaminoglycans, and proteoglycans, along with blood and lymphatic vessels, as well as various cells such as fibroblasts and macrophages, Figure 1-17. Mesothelium protects internal organs, including the large intestine, and provides a dynamic cellular membrane that facilitates the transport and movement of fluid and particulate matter across the serosal cavities in response to inflammation and tissue repair, as well as the clearance of fibrin, and serves as a passive barrier to metastasis (138, 139).

Colorectal peritoneal carcinomatosis, or the presentation of peritoneal metastasis in colorectal cancer patients, is one of the most common occurrences in CRC metastasis, with approximately 13% of CRC patients related to poor prognosis (140). A higher risk is associated with young age, locally advanced primary tumours, lymph node involvement, positive resection margins, mucinous histology, infiltrating growth pattern, perineural invasion, or venous invasion. However, the mechanisms remain unclear. The pathophysiology of peritoneal metastasis in CRC is explained by the spontaneous shedding of cancer cells from the primary tumours or as a result of surgery. Subsequently, the detached cancer cells are transported by the peritoneal fluid in a clockwise direction, attaching to the mesothelium or submesothelial lymphatics of the peritoneum, thereby invading the submesothelium and promoting angiogenesis (141, 142). In terms of the tumour microenvironment, the loss of cell polarity in mesothelial cells, also known as the mesothelial-to-mesenchymal transition (MMT) mechanism, and the presence of CAFs are suggested to provide a tumour-promoting niche (139).

Detecting colorectal peritoneal carcinomatosis poses challenges due to the low volume of peritoneal metastases, which leads to poor sensitivity with standard imaging methods, such as CT or PET. Diffusion-weighted MRI (DW-MRI) offers enhanced sensitivity and specificity. Surgery can be curative in some cases. However, cytoreductive surgery (CRS) and hyperthermic intraperitoneal chemotherapy (HIPEC) have demonstrated promising clinical outcomes, along with palliative care for the majority of patients. Understanding the biology and novel treatment strategies for peritoneal metastasis requires further investigation (141-143).

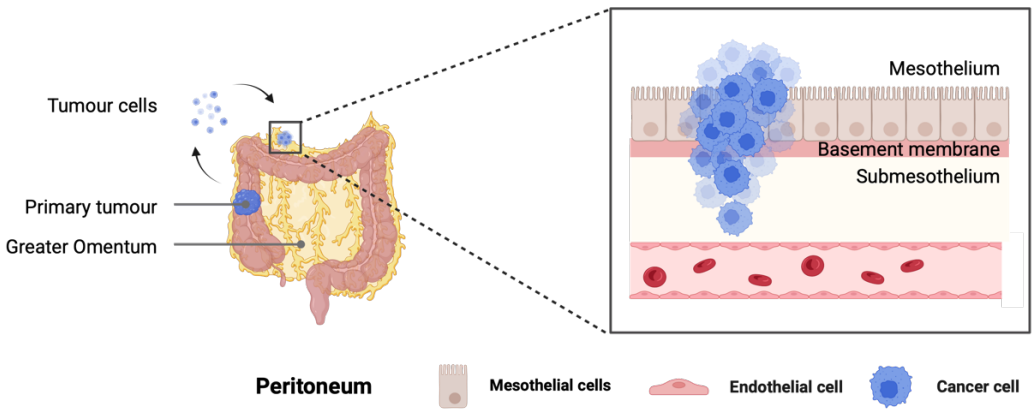


Figure 1-17 Peritoneal microenvironment as a target for metastatic CRC

The diagram illustrates the peritoneum and its microenvironment, which comprises a monolayer of mesothelial cells lining the peritoneal cavity, supported by a submesothelial stroma containing an extracellular matrix, blood and lymphatic vessels, and various cells such as fibroblasts and macrophages. CRC metastasis may disrupt this homeostatic environment. This figure was created using BioRender.com and adapted from Baaten I. et al., 2020; and Lemoine L., Sugarbaker P., and Van der Speeten K., 2016.

1.5.2.3 Lung

Regarding the essential function of the lung in gas exchange for the respiratory system, the tissue comprises the alveolar region (parenchyma), which forms clusters (acini) at the bronchioles, where it reaches the air. Additionally, a capillary network contributes to blood flow within interalveolar septa. Between the two compartments of air and blood, the interalveolar septa create a barrier consisting of an epithelium facing the alveolar lumen and an endothelium facing the capillary lumen, as well as the interstitium (connective tissue) (144). The epithelium is a mosaic of thin squamous alveolar epithelial cells (type I) for lining and cuboidal secretory cells (type II), along with various additional cells, including basal cells, club cells, ciliated cells, goblet cells, tuft cells (brush cells), pulmonary neuroendocrine cells (PNECs), and pulmonary ionocytes (145). In contrast, alveolar capillary endothelial cells form a single layer with less complex branching. Both are surrounded by interstitium, which contains abundant cells, such as myofibroblasts, and an extracellular connective tissue fibre network comprising elastic and collagen fibrils that form fibres (144), Figure 1-18.

Metastatic CRC to the lung is the second most common site, accounting for approximately 10%, after the liver (146). The incidence of synchronous lung metastases has increased over time and is more frequently associated with liver metastasis and rectal cancer (147). For patients with resectable metastatic CRC, pulmonary metastasectomy is recommended as a potential curative procedure linked to favourable survival outcomes (148, 149). The identification of genetic mutations in CRC, specifically *BRAF* and *KRAS* mutations, as part of the regulation in the MAPK/ERK and EGFR signalling pathways, is used to determine survival and disease recurrence in CRC patients after pulmonary metastasectomy, indicating poorer outcomes (149). There is limited understanding of the biological alterations underlying CRC with pulmonary metastases, unlike the more established evidence in liver metastasis. Therefore, a deeper understanding of the disease is still

necessary, along with a practical diagnostic and therapeutic approach (146, 148, 149).

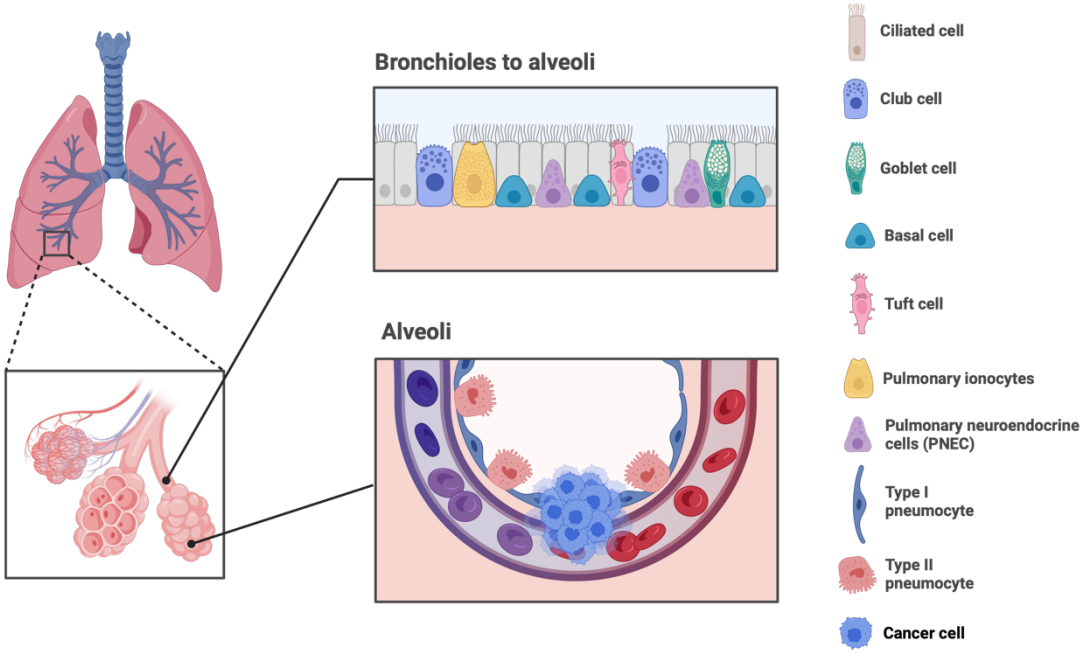


Figure 1-18 Lung microenvironment as a target for metastatic CRC

The diagram illustrates the lung and its microenvironment, comprising the alveolar region that forms clusters (acini) at the bronchioles. An epithelium lines the alveolar lumen and an endothelium lines the capillary lumen. Alveolar epithelial cells (type I), cuboidal secretory cells (type II), and other cells, such as basal cells, club cells, or ciliated cells, are found in this area. CRC metastasis may disrupt this homeostatic environment. This figure was created using BioRender.com and adapted from Meng X., Cui G., and Peng G., 2023.

1.5.2.4 Bone

There is increasing evidence of bone metastases from CRC; however, this disease is rarely detected, with an incidence rate of only 3-7%. It may occur in relation to rectal cancer, involving lymph node invasion of the primary tumour and lung metastases. CRC patients with bone metastasis have a short survival time, ranging from 5 to 21 months after diagnosis, depending on the cohort (10). In clinics, bone metastasis frequently presents in the spine (65%), pelvis or hip (34%), long bones (26%), and other bone sites (17%) (9). Skeletal-related events (SREs), which cause morbidity, are linked to bone pain, fractures, and hypercalcemia. Additionally, osteolytic, osteoblastic, or mixed metastatic bone lesions result from the dysregulation of cellular processes in bone cells, such as osteoclastogenesis and

bone resorption, which is proposed as a mechanism underlying bone metastasis (8).

Regarding the bone marrow (BM) niche, this heterogeneous area supports haematopoiesis and exhibits high signalling dynamics to maintain cell homeostasis. The distinct zones within the BM niche provide unique environments for maintaining quiescence or supporting the activation of various cells and other components in the niche. For instance, haematopoietic stem cells (HSCs) are predominant in the hypoxic environment of the endosteal zone. In contrast, highly active cells, including both haematopoietic and non-haematopoietic tissues, are found in the perivascular zone, which contains more circulating vessels that supply oxygen and nutrients to the bone cavity. Chemokines, cytokines, growth factors (GF), and oxygen levels are regulated through autocrine or paracrine signalling by different cell types in the BM niche. Sinusoidal endothelial cells and perivascular stromal cells reside in this zone, where cellular activities are more intense. Other cells present in these zones include osteo-lineage cells, sinusoidal endothelium, fibroblasts, Schwann cells, perivascular stromal cells, adipocytes, endothelial cells, mesenchymal stem cells (MSCs), and immune cells, which form a group of haematopoietic and other non-haematopoietic tissues within the BM niche. Additionally, the ECM is a crucial element that supports this heterogeneous structure (150-152), Figure 1-19.

Several signalling pathways, such as MAPK, NF-κB, and PI3K/AKT, are produced and activated between tumour cells and BM cells, which work together to create a metastasis-promoting environment (8). BMPs signalling is essential for bone differentiation and remodelling (150, 153). Alterations in BMP ligand proteins during osteoinductive processes, along with the loss of SMAD4, are implicated in metastatic CRC and malignancy in bone. However, the mechanisms linked to pathological processes are not yet fully understood. Furthermore, current Pharmacological treatments, such as bisphosphonates, denosumab, vitamin D, calcium, and dexamethasone, are used to limit the progression of metastasis and reduce skeletal-related events (SREs). It requires further investigation into suspected biological processes, potential prognostic markers, and alternative treatment approaches that would benefit not only palliative care but also curative or even preventive tools for patients (8-10).

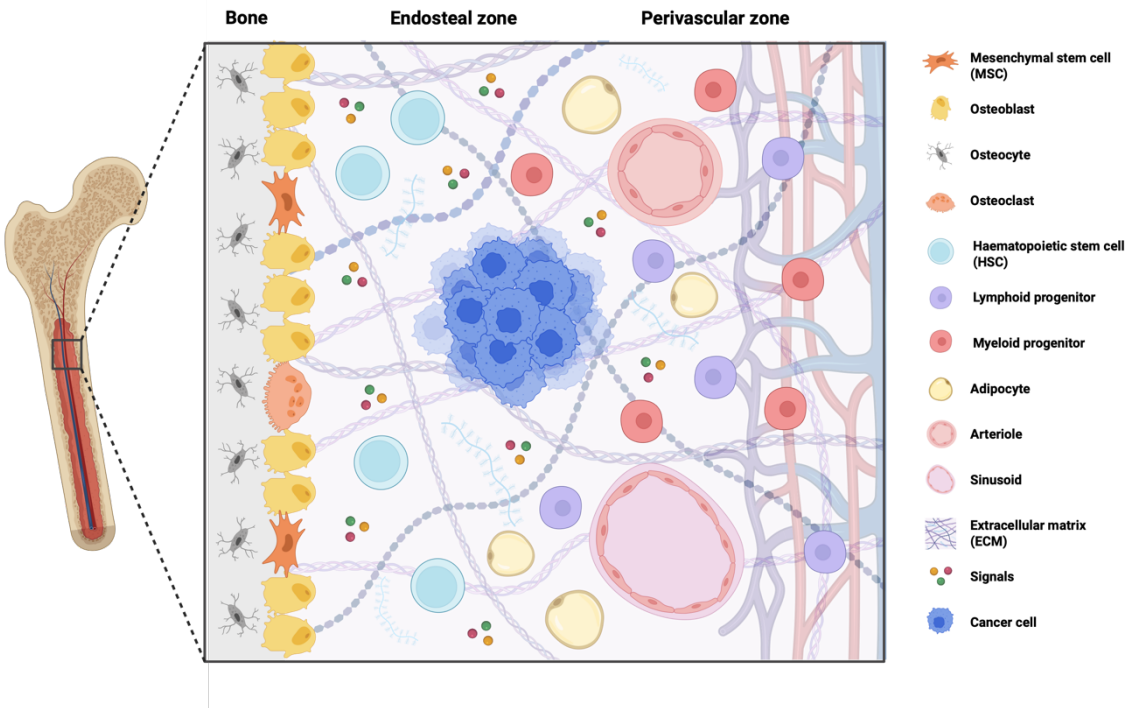


Figure 1-19 Bone microenvironment as a target for metastatic CRC

The diagram illustrates the bone and its microenvironment, comprising two distinct zones. The endosteal zone contains haematopoietic stem cells (HSCs), and the perivascular zone contains more circulating vessels and various cell types, such as osteo-lineage cells, adipocytes, endothelial cells, mesenchymal stem cells (MSCs), and immune cells. CRC metastasis may disrupt this homeostatic environment. This figure was created using BioRender.com and adapted from Busch C, and Wheadon H., 2019.

1.6 Therapeutic approaches for CRC

1.6.1 Surgery

As previously described regarding the focused screening and resection of polyps, polypectomy has been shown to provide benefits in the treatment of polyps. In cases of CRC, surgical resection remains the primary standard curative procedure. Colectomy is recommended for localised, resectable colon cancer, and corresponding lymphadenectomy is advised for the removal of regional lymph nodes. Patients with T4b tumours may consider neoadjuvant chemotherapy (42, 154-156). In rectal cancer, neoadjuvant radiation or chemoradiation is suggested for locally advanced or high-risk disease prior to surgery, whereas those with low-risk, early-stage cancer proceed directly to resection (42, 157). Currently, advanced surgical techniques, including laparoscopic surgery, robotic surgery (transanal minimally invasive surgery; TAMIS, and transanal total

mesorectal excision; TaTME), and navigation surgery (injecting fluorescence and indocyanine green (ICG) and artificial intelligence (AI)-based real-time microcirculation analysis), have been developed to enhance the efficiency of CRC treatment (42, 156). Nevertheless, surgical resection is a highly effective procedure for resectable cases of CRC; alternative standard therapies are proposed for non-resectable cases, alongside enhanced therapeutic approaches, including radiotherapy, chemotherapy, and immunotherapy, which are utilised in the management of CRC (42, 158).

1.6.2 Radiotherapy

Radiation therapy utilises high-energy X-rays or gamma rays to disrupt local tumour growth and enhance quality of life. Several advanced radiation techniques have been developed to achieve maximum precision in targeting tumour locations (42). For instance, image-guided (high-resolution imaging), intensity-modulated, stereotactic body, adaptive, and nanoparticle-mediated radiation therapy, alongside proton beam and heavy ion therapy, have been designed to address limitations and concerns regarding side effects, such as exposure to adjacent healthy tissue, local recurrence, and metastasis (42, 159). By accurately targeting tumours, neoadjuvant radiation therapy is employed to reduce the tumour burden for intermediate and advanced stage cancer, potentially in combination with chemotherapy as a total neoadjuvant therapy, thereby improving surgical outcomes (42, 158).

1.6.3 Chemotherapy

Chemotherapy and radiotherapy can be utilised both before (neoadjuvant) and after (adjuvant) surgery to achieve greater benefits from these treatments. While radiotherapy targets tumours locally, chemotherapy provides a systemic approach with cytotoxic effects, eliminating cancer cells or inhibiting cell proliferation, ultimately leading to reduced disease progression and an increased lifespan for individuals (158). Several drugs for treating CRC have been approved by the Food and Drug Administration (FDA), with Fluoropyrimidine (5-FU)-based therapies being the most common single-agent first-line treatment. Additionally, multiple-agent regimens including oxaliplatin (OX), irinotecan (IRI), and capecitabine (CAP or XELODA or XEL), along with combined therapy regimens of FOLFOX (5-FU + OX),

FOXFIRI (5-FU + IRI), XELOX or CAPOX (CAP + OX), and CAPIRI (CAP + OX) are currently employed in first-line treatment (160). Due to toxicity, the use of drug regimens is considered based on the patient's tolerability and the disease stage, such as adjuvant therapy, which is recommended for CRC patients with stage II and more advanced stages (42). Not only is systemic toxicity a concern in chemotherapy, but unsatisfactory response rates, unpredictable innate and acquired resistance, and low tumour-specific selectivity are also issues. Hence, due to the heterogeneity of CRC more targeted approaches are required with many novel drugs in development (160).

1.6.4 Targeted therapy

In relation to significant changes in the biological processes of tumour cells and their microenvironment, small molecules and monoclonal antibodies (mAbs) have been developed to target specific cancer aberrations that regulate cell proliferation, differentiation, and migration, as well as those in local blood vessels and immune cells (42, 160). The EGFR pathway is frequently cited as the target of mAbs, such as Cetuximab (a chimeric immunoglobulin G (IgG) mAb; the first mAb that the FDA approved) and Panitumumab (a humanised mAb), which inhibit the binding of EGF to EGFR, ultimately leading to the inhibition of proliferation and the induction of apoptosis, along with the inhibition of angiogenesis. Similarly, Bevacizumab (a humanised mAb), Aflibercept (a fusion recombinant protein), Ramucirumab (a fully human mAb), and Regorafenib (a small molecule multi-targeted tyrosine kinase inhibitor) disrupt the binding of ligands to their receptors in the VEGF/VEGFR signalling axis, resulting in the inhibition of angiogenesis, a vital component of tumour growth and metastasis. Furthermore, targeting BRAF and Human EGFR-2 (HER2) also inhibits the phosphorylation cascade involved in regulating cell proliferation and tumour development. Encorafenib and Binimetinib are small-molecule inhibitors of BRAF and MEK, while Trastuzumab (a recombinant humanised mAb) and tucatinib (an oral small molecule) target HER2. Although they demonstrate promising specific targeting inhibitory effects, some common side effects associated with these targeted therapies have been reported, such as hypertension, electrolyte imbalance, nausea, diarrhoea, and asthenia (42).

1.6.5 Immunotherapy

Other advanced treatment have been developed to prevent immunosuppression, which arises from crosstalk between immune and non-immune cells in the tumour microenvironment that confers tumour cells with immune evasion (161). To boost humoral and cellular immunity against cancer cells, current approaches have focused on immune checkpoint (IC) inhibition, adoptive cell transfer (ACT), cytokine therapy, and dendritic cell vaccines. Immune checkpoint molecules are designed to disrupt the ligand-receptor complexes involved in dysregulated signal transduction during the immune response. As is commonly known, the programmed death 1 (PD-1) and programmed death ligand 1 (PD-L1) complex, which is present on various immune cells, including activated T cells, NK cells, and dendritic cells, serves to suppress the immune response. Pembrolizumab, Nivolumab, and Dostarlimab are three drugs that have been used as inhibitors of this pathway. Additionally, the inhibition of cytotoxic T-lymphocyte-associated antigen-4 (CTLA-4), an immune checkpoint molecule expressed on the surface of activated T cells, is achieved with Ipilimumab, which upregulates effector T cell activity and suppresses regulatory T cells, thereby inhibiting tumour progression (42, 158).

Adaptive cell transfer (ACT) is an approach that activates an intrinsic molecular pathway *ex vivo* by stimulating autologous lymphocytes to promote tumour eradication. This immunotherapy-based procedure involves tumour-infiltrating lymphocytes or T-cell receptor gene therapy, commonly known as chimeric antigen receptor (CAR) modified T cells. Recently, new forms of immunotherapy have emerged, including cytokine-induced killer (CIK) cells combined with dendritic cells and cancer vaccines targeting tumour-associated antigens such as CEA, mucin 1, Her2, which may offer improved survival benefits. However, its research applications remain limited (42).

1.7 TGF-β superfamily signalling in normal and cancer cells

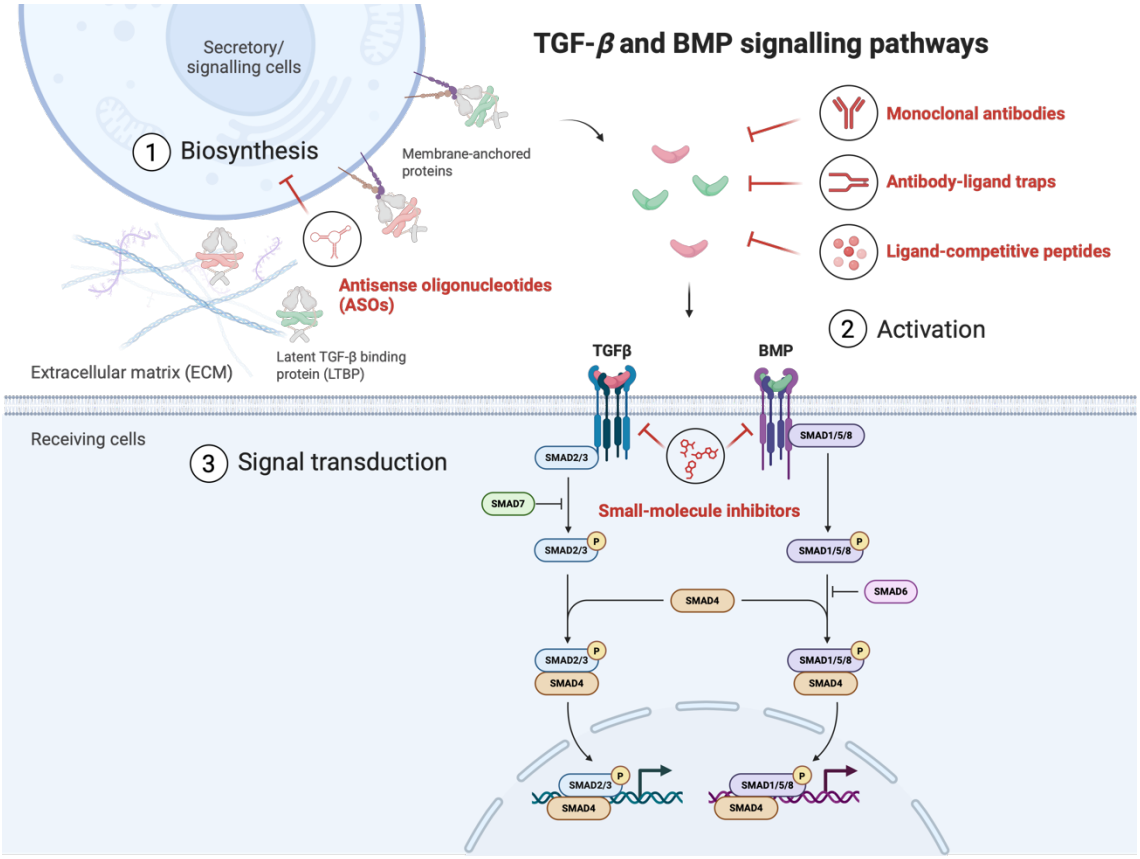


Figure 1-20 TGF-β and BMP signalling pathways and targeted inhibition

The diagram illustrates the canonical TGF-β and BMP signalling pathways, where TGF-β and BMP precursors are synthesised and released into the ECM before ligands bind to receptors, activating the signalling cascades. Phosphorylation occurs on R-SMADs, including SMAD2 and SMAD3 for TGF-β signalling, as well as SMAD1, SMAD5, and SMAD8 for BMP signalling, followed by complex formation with SMAD4. This complex then translocates to the nucleus to control the transcription of target genes. To inhibit these signalling pathways, various therapeutic agents have been developed to target different aspects, including biosynthesis, activation, and signal transduction. These agents include antisense oligonucleotides (ASOs), monoclonal antibodies, antibody-ligand traps, ligand-competitive peptides, and small-molecule inhibitors. This figure was created using BioRender.com and adapted from Massague J. and Sheppard D., 2023; Deng Z. et al., 2024; and Liu S., Ren J., and Ten Dijke P., 2021.

1.7.1 TGF-β signalling members and their function

One of the critical signalling pathways involved in embryonic development, tissue homeostasis, and regeneration is the transforming growth factor ß (TGF-ß) family. The TGF-ß family comprises two subfamilies of cytokines: the TGF-ß/Nodal subfamily (including TGF-ß 1-3; Nodal; Activins A, B, C, and E; growth and differentiation factors (GDF) 1, 3, 8, 9, and 11; Inhibin; as well as Lefty-1 and 2) and the bone morphogenetic protein (BMP) subfamily (which includes BMP 2, 3, 4,

5, 6, 7, 8, 8B, 9, 10, and 15; GDF 5, 6, 7, and 10; and anti-Müllerian hormone, AMH). For signal activation, after precursor complexes of dimeric TGF-β with latency-associated peptide (LAP) are produced and secreted from secretory cells into the ECM, they become a latent form by binding to latent TGF-β binding protein (LTBP) in the ECM or membrane-anchored proteins on the cell surface. Subsequently, LAP binding to integrins on adjacent cells, which leads to the release of TGF-β. TGF-β ligands bind to pairs of transmembrane serine/threonine protein kinase subunits, including type I receptors (TGFBR1 (also known as activin receptor-like kinase 5; ALK5), ACVR1A (ALK2), ACVR1B (ALK4), ACVR1C (ALK7), ACVRL1 (ALK1), BMPR1A (ALK3), and BMPR1B (ALK6)) and type II receptors (TGFBR2, ACVR2A, ACVR2B, BMPR2, and AMHR2), with the assistance of the co-receptor (membrane proteoglycan betaglycan) presenting the signals, which leads to the assembly of the receptor complex that activates signal transduction in the receiving cells. Once ligand-receptor binding occurs, the type II receptor phosphorylates and activates the type I receptor kinase, subsequently recruiting and phosphorylating the SMAD transcription factors at the docking site of the type I receptor (162).

Following the step of receptor-mediated phosphorylation, receptor-regulated SMADs (R-SMADs), specifically SMAD2 and SMAD3 from the TGF-β subfamily, along with SMADs 1, 5, and 8 from the BMP subfamily, form a complex with the common-partner SMAD (co-SMAD, also known as SMAD4). This complex translocates into the nucleus as part of a canonical signalling cascade, functioning as a transcription factor (TF) that binds to specific DNA sequences (SMAD-binding elements, SBEs) and collaborates with transcriptional cofactors such as histone acetyltransferases (HATs) and histone deacetylases (HDACs) to activate or repress the downstream transcriptional activity of target genes, Figure 1-20. In addition to transcriptional cofactors that regulate SMAD signalling, inhibitory SMADs (I-SMADs) comprise SMAD6 and SMAD7, which act as antagonists of SMAD4 and the type I receptors, respectively. Negative feedback through SMAD7 is induced by TGF-β, BMP, interferon-γ (IFN-γ), and other signals. Furthermore, post-translational modifications can also influence the functions of SMADs. R-SMADs and co-SMADs can be phosphorylated by other protein kinases at many different sites to enhance or attenuate SMAD activity. For instance, the additional phosphorylation of active SMAD complexes, triggered by glycogen synthase kinase 3β (GSK3β), leads to the

degradation of SMADs through binding to the ubiquitin ligases, such as SMAD ubiquitination regulatory factors (SMURF1/2) and neural precursor cell expressed, developmentally downregulated 4-like (NEDD4L). The dephosphorylation and dissociation of SMADs are mediated by small C-terminal domain phosphatases (SCP1 and SCP2), as well as poly(ADP-ribose) polymerase-1 (PARP-1), allowing SMADs to engage in subsequent signal transduction (162, 163).

The TGF-ß/BMP signalling pathways are also activated through SMAD-independent mechanisms, known as non-canonical pathways. TGF-ß can stimulate various intracellular signal transduction pathways, including ERK signalling, Rho guanosine triphosphatase (GTPase) signalling, p38 MAPK signalling, JNK signalling, NF-κB signalling, PI3K/AKT signalling, and JAK/STAT signalling. Both canonical and non-canonical TGF-ß/BMP signalling pathways are recognised as mediating a broad range of cellular events, including cell survival, growth, proliferation, differentiation, adhesion, morphogenesis, migration, metabolism, immunity, and death. Furthermore, these signalling pathways can regulate multiple cells within the tissue environment, depending on the context-specific functional response. The regulation of cell differentiation, epithelial/endothelial-mesenchymal transition (EMT/EndMT), and apoptosis is emphasised during embryonic development. Inflammation, re-epithelialisation, angiogenesis, and fibroblast activation are facilitated in wound healing. Various mechanisms involved in suppressing cell proliferation and inducing cell apoptosis are regulated by the TGF-ß/BMP signalling pathway, contributing to tissue homeostasis. These signalling pathways regulate multiple immunocompetent cells, maintaining the balance of their activities in immunoreaction and immunosuppression for immune homeostasis (162, 163).

1.7.2 Dysregulation of TGF-β/BMP signalling in cancer

Several pathological processes, including developmental defects, aberrant healing, fibrotic diseases, inflammatory diseases, infectious diseases, and cancer, are associated with the dysregulation of TGF-ß/BMP signalling (162, 163). The significance of TGF-ß signalling pathways is particularly highlighted in various cancers, such as CRC, as they can either suppress or promote carcinoma progression depending on the stage of the disease. In the pre-malignant stage, these signalling pathways contribute to tumour suppression by inhibiting

tumourigenic inflammation and eliminating mutant cells through apoptosis. However, mutant cells may evade tumour suppression by acquiring mutations or alterations that inactivate the TGF-β pathway, as well as by decoupling apoptosis, leading to tumour promotion and progression. Following their role in tumour promotion, TGF-β signalling can induce an immunosuppressive TME by disrupting the functions of immune cells, thereby limiting the effectiveness of immunotherapy. Furthermore, CAFs can also be activated by TGF-β, resulting in extensive ECM remodelling and paracrine signalling within the TME, creating intra-tumoural fibrosis and ECM stiffness, which promote cancer cell growth, invasion, immune evasion, and angiogenesis. Similarly, EMT is induced by TGF-β to facilitate tumour dissemination through enhanced tumour invasion and metastasis via circulation (162).

In the metastatic dormancy stage, disseminated cancer cells may endure immune attacks, physical barriers, and metabolic stresses. To withstand these considerable stresses, TGF-β may protect metastatic progenitors by regulating the balance between quiescent and proliferative states, facilitating immune evasion and preserving these cells for long-term relapse. When disseminated cancer cells emerge at metastatic sites, TGF-β in the TME may promote organ-specific interactions between cancer cells and stroma, which are necessary for the colonisation traits of metastatic tumours, by regulating immune suppression, EMT, invasion, and further dissemination (162).

In relation to CRC, the dysregulation of TGF-β signalling pathways is emphasised during CRC progression. The *TGFBR2* mutation is observed in approximately 30% of CRC, with a higher prevalence of 80% occurring in frameshift mutations within MSI-H and a greater incidence in the proximal colon. The inactivation of TGFBR2 in MSI-H CRCs is associated with a dMMR and demonstrates a robust immune response, thereby promoting inflammation within the TME. This mutation is significant in the later stages of MSI adenomas transitioning to cancer, alongside other pathways such as Wnt-β-catenin, Hippo, and MAPK, which stimulate proliferation. Furthermore, improved survival rates are linked to the *TGFBR2* mutation. As part of TGF-β signal transduction, a *TGFBR1* mutation has been noted to be associated with an increased risk of CRC (17, 124, 164).

At the same stage in the transition from adenoma to carcinoma, *SMAD4* mutations have been observed in 10% of CRC cases. The loss of heterozygosity (LOH) on chromosome 18 may lead to the inactivation of SMAD4, as well as other genes in this region, such as SMAD2 and DCC. LOH not only affects SMAD4 inactivation, but various mechanisms in post-transcriptional and post-translational regulation, including ubiquitylation, sumoylation, and microRNA interference, may also contribute to the SMAD4 defect. Interestingly, later stages of malignancy and metastasis are recognised to result in the loss of SMAD4 function, while early stages are associated with SMAD2. Regarding the function of SMAD4 as a tumour suppressor, inactivation of SMAD4 leads to the disruption of canonical TGF-ß signalling. Furthermore, the loss of SMAD4 facilitates the activation of other pathways that promote tumour progression, and the dysregulation of target genes promotes aggressive phenotypes, such as angiogenesis, the acquisition of stem cell properties, and chemoresistance. SMAD4 deficiency tends to have a poor prognosis, with a worse response to fluorouracil-based therapies (17, 124, 164-166).

Although TGF-ß signalling has been extensively shown to be dysregulated in the development of CRC, mutations and alterations in the BMP signalling pathway, which also play a bifunctional role in CRC, are increasingly reported. Mutations in *BMPR1A* can be detected in both premalignant and malignant stages (22, 23). Additionally, the expression of BMP9 and BMP4 increased during carcinogenesis and cancer, respectively (167, 168). Poor prognosis and survival outcomes are linked to the loss of SMAD4 in tumours, alongside typical BMP receptor expression in these tumours and elevated BMP2 expression in the stroma (169-171). At this time, the dysregulation of these signalling pathways still requires further investigation to gain a better understanding of these comprehensive biological processes, thereby clarifying their role in carcinogenesis, promoting CRC metastasis and facilitating potential therapeutic strategies.

1.7.3 Inhibition of TGF-β/BMP signalling

To target TGF-ß/BMP signalling, several therapeutic agents have been developed to regulate various aspects of the signalling activity, including biosynthesis, activation, and the signalling axis. For instance, antisense oligonucleotides (ASOs), antisense RNA (such as small interfering RNAs (siRNAs) and short-hairpin RNA

(shRNA)), ligand-competitive peptides, and monoclonal antibodies are designed to modify processes during TGF-β/BMP biosynthesis and activation by targeting TGF-β mRNAs, the latent TGF-β complex, or other components involved in signalling activation. Additionally, the signalling cascade may be disrupted by directly targeting each component in the signal transduction pathway. Antibody-ligand traps can be employed to target TGF-β ligands, while the kinase activity of TGF-β receptors can be inhibited by various small-molecule inhibitors. Targeting SMADs and their interacting factors can be selectively accomplished using antisense oligonucleotides, small-molecule inhibitors, or SMAD-binding peptide aptamers, Figure 1-20. Nevertheless, it still has limited therapeutic applications in clinical practice and requires further development, as many approaches in clinical trials have failed to demonstrate clinical benefit in terms of efficacy and safety (163, 172).

Considering the properties of small-molecule inhibitors, various compounds have been designed with molecular structures that specifically target the type I receptor of TGF-β/BMP, thereby inhibiting the phosphorylation of SMADs and suppressing canonical or non-canonical pathway signal transduction. Furthermore, these drug compounds provide advantages in terms of ease of administration, stability, cost-effective production, and the potential for rapid withdrawal in the event of adverse effects, owing to their short half-lives and high likelihood of success in preclinical studies (173, 174). To enhance an advanced therapeutic approach for anti-CRC by targeting TGF-β/BMP signalling pathways, specific targeted therapeutic candidates from other diseases and malignancies may be evaluated in CRC models. For instance, fibrodysplasia ossificans progressiva (FOP) and paediatric intrinsic diffuse glioma (PIDG) are caused by activating mutations of ALK2. Thus, inhibitors of the type I BMP receptor (ALK1/2/3/6) are being developed and tested for these diseases. Several compounds have been developed and investigated in preclinical studies, such as Dorsomorphin (Dor, the first kinase inhibitor targeting Pan-type 1 BMP receptor), as well as LDN-212854 and K02288 (both have enhanced selectivity for ALK2) but also target ALK1 and ALK6 whilst showing weaker inhibition of other ALKs (175-178).

1.8 Modelling in cancer research

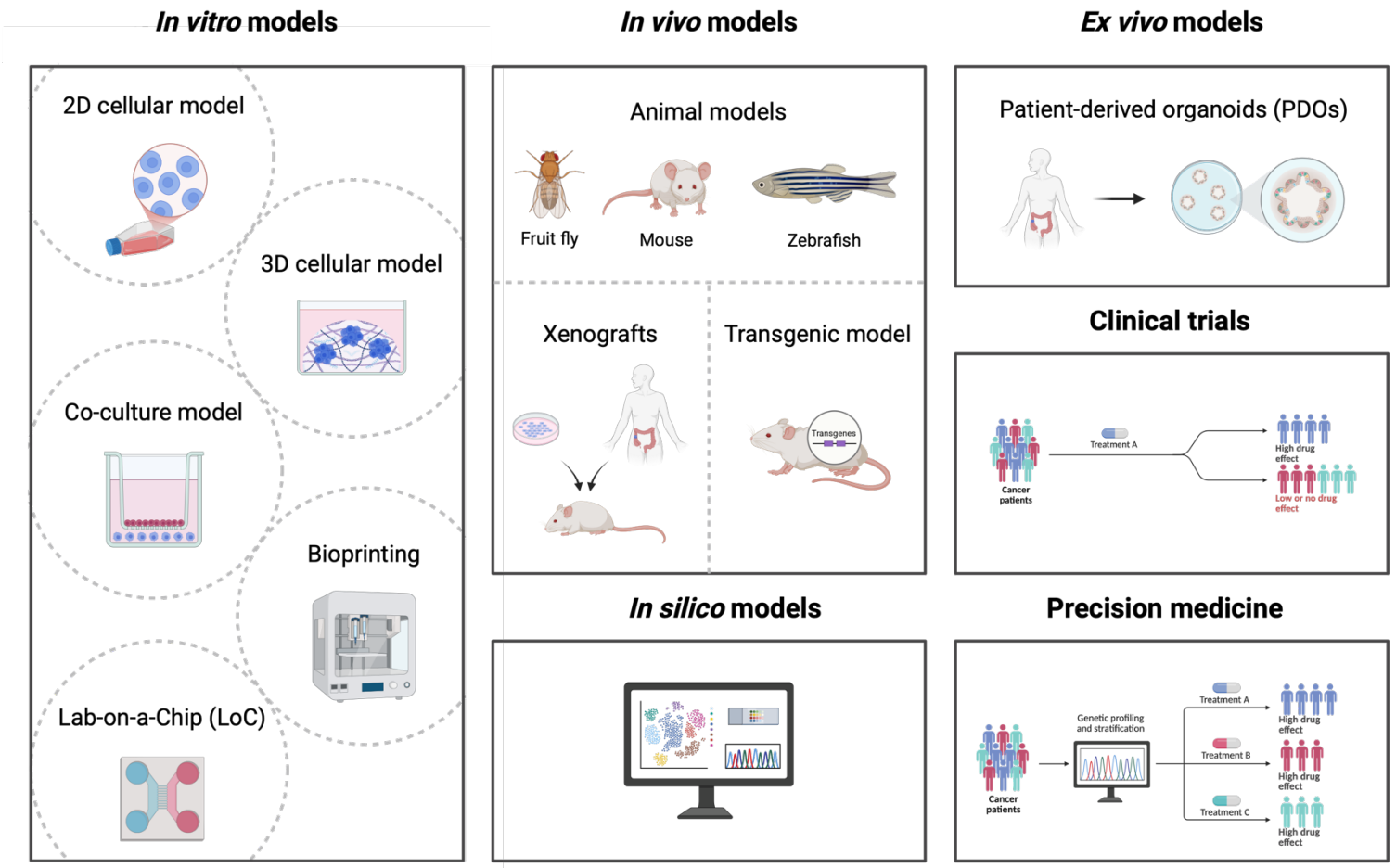


Figure 1-21 Overview of modelling in cancer research

The diagrams illustrate modelling in cancer research, which includes *in vitro* models (e.g., 2D cellular model, 3D cellular model, co-culture model, bioprinting, and lab-on-a-chip), *in vivo* models (e.g., animal models, xenografts, transgenic models), *in silico* models (computational or programming modelling), *ex vivo* models (e.g., patient-derived organoids), clinical trials, and precision medicine. This figure was created using BioRender.com.

1.8.1 *In vitro* models

In vitro models have been widely utilised as experimental tools in preclinical studies, enabling the investigation of cancer characteristics and mechanisms at the cellular and molecular levels. Their applications range from cancer biomarkers to drug screening and the development of novel therapeutic approaches, see Figure 1-21. The first fundamental tool involves two-dimensional (2D) cellular models, which are developed by isolating various origins and types of cells from patients or animals, such as intestinal cells and CRC cells. The established cell lines are immortal, rapidly proliferating, and highly consistent; however, most cells are cultured in monoculture and maintained in artificial media on culture plates, lacking the matrix components found in the TME. As a result, the continuous proliferation of these cell lines may lead to a loss of the characteristics of the original tumour (179).

Three-dimensional (3D) cellular models have been developed to replicate the physiologically heterogeneous nature and complex interactions among various cells, thereby addressing the limitations of 2D cellular models by growing them as spheroids and organoids. Spheroids consist of a mixture of single-cell or multicellular systems that can incorporate stem cells, cancer cells, immune cells, and stromal cells from suspended cell lines or isolated tumour tissues from patients into a 3D structure under aggregated and compacted conditions (stiffness) in a hydrogel, as well as in non-to-low-adherent environments (179).

Organoids represent a relatively complex 3D culture compared to spheroids, involving self-organising organotypic cultures derived from various stem cells, including tissue-specific adult stem cells (ASCs), embryonic stem cells (ESCs), induced pluripotent stem cells (iPSCs), and minced human CRC tumours for patient-derived organoids (PDOs). The organoids require a specific medium containing stem cell niche restorative factors and tissue-specific growth factors,

along with culturing in 3D Matrigel conditions to provide stable, physiologically epithelial structures that resemble the *in vivo* growth microenvironment. There is potential to develop a co-culture model of organoids with multiple cell types, such as immune cells and mesenchymal cells. However, the model remains limited in its ability to maintain the blood vessels, immune system, and peripheral nervous system of tumour cells (179).

Due to advanced technology, various techniques have been developed to facilitate *in vitro* studies of CRC and TME interactions through co-culture modelling. For instance, a transwell provides physical separation via a membrane, enabling soluble factors to permeate and communicate between different cell types within the same environment. In contrast, tumouroids consist of integrated 3D cultures with multiple components of the tumour niche, resembling cancer cell features such as cell-cell interaction and paracrine signalling in a 3D structure observed in tumours. Additionally, employing natural, synthetic, and hybrid hydrogel substrates to construct scaffolds or matrix droplets in culture, along with 3D bioprinting, can also facilitate the microarchitecture of complex co-cultures (180).

Advances in biomedical engineering technology has enhanced applications in biomedical research. Microfluidic devices represent an innovative technique that offers advantages over conventional experiments. Microfluidic fabrications are designed to incorporate microchannels that facilitate liquid flow at dimensions ranging from 1 μm to 1 mm, allowing for the manipulation of fluid behaviours and the customisation of materials through various fabrication methods, such as moulding and 3D printing. Consequently, this platform supports the advancement of tissue engineering (microfibre) or 3D cell culture with dynamic and physical properties that mimic essential cellular microenvironments, also referred to as Lab-on-a-Chip (LoC). Nevertheless, challenges persist in future research trends (180-182).

1.8.2 *In vivo* models

In vivo models, often referred to as animal models, can replicate disease occurrences while considering genetic diversity and immune function, thereby enhancing the investigation of the pathophysiological mechanisms underlying the

development of CRC. Mouse models are the most frequently employed in biomedical research. Other animals, such as the fruit fly, zebrafish, and pigs, have also been utilised as preclinical models of CRC, Figure 1-21. Mouse models, which offer various advantages such as low cost, manageability, a short gestation period, anatomical similarity to humans, and ease of genetic manipulation, have been used to elucidate disease pathogenesis, diagnosis, prognosis, and the identification of predictive markers, as well as to validate new therapies in preclinical studies (179, 183).

Among numerous mouse models, various techniques have been established, including carcinogen-induced models (CIMs), genetically engineered mouse models (GEMMs), cell line-derived xenografts (CDX), patient-derived xenografts (PDX), patient-derived organoid xenografts (PDOX), and orthotopic injection-induced metastasis models, all providing significant benefits for specific research purposes. Following the development of transgenic models, practical tools are now available to create knock-out and knock-in animal models, thereby facilitating the study of gain-of-function or loss-of-function mutations in oncogenes and tumour suppressor genes. Although these models have served as relevant tools in preclinical studies, limitations have been reported, including animal-to-animal variation (179, 183, 184).

1.8.3 Other models

In silico models, also referred to as computational modelling or programming modelling, are constructed using domain-specific languages, innovations in software engineering, and data from biological and chemical reactions, along with mathematical representations, to predict mechanistic insights into cellular processes, Figure 1-21. Consequently, novel programming-based approaches are being stimulated, creating a dynamic framework for specific biological systems that offers greater benefits in cancer pathogenesis investigation, the development of biomarkers for early detection, diagnosis, drug discovery, and the enhancement of therapeutic strategies. Although this technology is straightforward in CRC model development, complexity remains a challenge concerning experimental designs and the calibration, analysis, and interpretation of data (185-190).

Ex vivo models in CRC involve obtaining tissues from patients through biopsy and maintaining them for further investigation. Patient-derived organoids (PDOs), as mentioned previously, are sourced from stem cells and develop self-organising structures using specific media conditions, (Figure 1-21). Cancer-on-a-chip (CoC) is another innovative platform that can be used to culture patient-derived cells. Additionally, decellularised scaffolds (dECMs) are created using specific detergents and enzymatic treatments to remove cells, preserve tissue architecture, and retain ECM components from the original tissues, enabling recellularisation with CRC cells and facilitating specific mutations and modulation with cytokines. Patient-derived explants (PDEs) are described as *ex vivo* systems that provide a complete original tumoural composition. Fresh tumour tissues are fragmented or sliced into various thicknesses, ranging from 150 µm to several centimetres, depending on the experimental setting. Consequently, different culture protocols are employed to maintain the explants, such as 3D microfluidic organ culture. Systems are under development to overcome the short-term lifespan, lack of vascularisation, and limitations in oxygen and nutrient supply, as well as waste removal in order to develop robust *ex vivo* culture systems for longer-term studies. These methods facilitate advancements in translational research, drug development, and therapeutic selection in personalised medicine for cancer therapy (180, 191).

1.8.4 Clinical trials and Precision medicine

Clinical trials, also known as Randomised Controlled Trials (RCTs), are prospective studies involving volunteers that aim to evaluate new interventions or treatments. As a relevant tool for examining the relationship between an intervention and an outcome, minimising bias is achieved through randomisation systems that assign participants to study groups with balanced characteristics. Although this research is time-consuming and costly, and some participants may be unavailable during follow-up, the significant advantage is that it demonstrates the effectiveness of the intervention prior to its application in the clinic. Clinical trials can be conducted for cancer screening and treatment management, see Figure 1-21. For instance, numerous clinical trials have been designed to study immunotherapy and targeted therapies in CRC. However, the limitation of treatment efficacy in patients persists and requires further novel therapeutic strategies (192-195).

Recently, the emerging concept of tailored treatment and prevention in oncology, known as Personalised Medicine (PM), also referred to as precision medicine is being implemented. Regarding each patient, the inter- and intra-tumour variability in genes, tumour microenvironment, lifestyle, and comorbidities associated with cancer varies from person to person. With a better fundamental understanding of molecular oncology and innovative technologies, relevant signatures of genetic and epigenetic events from a multi-omics approach contribute to the identification of tumour-specific biomarkers and the development of treatment strategies, such as targeted and immune therapies. Therefore, optimal medical management can be tailored to each specific patient, incorporating effective treatment options and supportive care to improve quality of life (196, 197), Figure 1-21.

1.9 The research aims and hypotheses

1.9.1 Research aims

The aims of this thesis include:

1) To explore the intrinsic expression patterns of the TGF-β and BMP signalling pathways in CRC and metastatic sites, such as the bone marrow niche, and to determine which components are associated with dysregulation in CRC prognosis.

2) To investigate the association between clinical and pathological data and to explore which histological subtype is related to the dysregulation of the TGF-β and BMP signalling pathways in premalignant lesions of CRC.

3) To investigate the association with clinical and pathological data and explore what sites of tumour development are associated with dysregulation of the TGF-β and BMP signalling pathways in CRC

4) To evaluate novel inhibitors of the TGF-β and BMP signalling pathways, both alone and in combination with chemotherapeutic agents in CRC, to investigate changes in the signalling mechanism, and to develop a model of metastasis.

1.9.2 Research hypotheses

This thesis hypothesises that deregulation of components in the TGF-β and BMP signalling pathways is associated with a poor prognosis, and consequent to CRC carcinogenesis and metastasis. Blocking the TGF-β and BMP signalling pathways with novel inhibitors, either alone or in combination with chemotherapeutic agents, could prevent and/or treat mCRC.

Chapter 2 Materials and Methods

2.1 Materials

2.1.1 Open access gene expression databases

2.1.1.1 Cancer Dependency Map (DepMap) portal

Access link: https://depmap.org/portal/

Table 2-1 Gene members in transforming growth factor-beta (TGF-β) superfamily

Genes	Ensembl	Pathways	Aliases
TGFB1	ENSG00000105329	TGF-β signalling	
TGFB2	ENSG00000092969	TGF-β signalling	
TGFB3	ENSG00000119699	TGF-β signalling	
TGFBR1	ENSG00000106799	TGF-β signalling	ALK5
TGFBR2	ENSG00000163513	TGF-β signalling	
TGFBR3	ENSG00000069702	TGF-β signalling	
BMP1	ENSG00000168487	BMP signalling	
BMP2	ENSG00000125845	BMP signalling	BMP2A
BMP3	ENSG00000152785	BMP signalling	
BMP4	ENSG00000125378	BMP signalling	BMP2B
BMP5	ENSG00000112175	BMP signalling	
BMP6	ENSG00000153162	BMP signalling	
BMP7	ENSG00000101144	BMP signalling	
BMP8A	ENSG00000183682	BMP signalling	
BMP8B	ENSG00000116985	BMP signalling	
BMP10	ENSG00000163217	BMP signalling	
BMP15	ENSG00000130385	BMP signalling	GDF9B
GDF1	ENSG00000130283	BMP signalling	
GDF2	ENSG00000263761	BMP signalling	BMP9
GDF3	ENSG00000184344	BMP signalling	BMP10
GDF5	ENSG00000125965	BMP signalling	BMP14
GDF6	ENSG00000156466	BMP signalling	BMP13
GDF7	ENSG00000143869	BMP signalling	BMP12
GDF9	ENSG00000164404	BMP signalling	
GDF10	ENSG00000266524	BMP signalling	BMP3B
GDF11	ENSG00000135414	BMP signalling	BMP11
GDF15	ENSG00000130513	BMP signalling	Paralog BMP6
BMPR1A	ENSG00000107779	BMP signalling	BMPR1A/ALK3
BMPR1B	ENSG00000138696	BMP signalling	BMPR1B/ALK6
BMPR2	ENSG00000204217	BMP signalling	

Table 2-1 Gene members in transforming growth factor-beta (TGF-β) superfamily (continue)

Genes	Ensembl	Pathways	Aliases
INHA	ENSG00000123999	Activin signalling	Activin-α /INHA
INHBA	ENSG00000122641	Activin signalling	Activin-βA /INHBA
INHBB	ENSG00000163083	Activin signalling	Activin-βB /INHBB
INHBC	ENSG00000175189	Activin signalling	Activin-βC /INHBC
INHBE	ENSG00000139269	Activin signalling	Activin-βE /INHBE
NODAL	ENSG00000156574	Activin signalling	
ALK	ENSG00000171094	Activin signalling	
ACVR1	ENSG00000115170	Activin signalling	ACVR1A, ALK2
ACVR1B	ENSG00000135503	Activin signalling	ALK4
ACVR1C	ENSG00000123612	Activin signalling	ALK7
ACVR2A	ENSG00000121989	Activin signalling	
ACVR2B	ENSG00000114739	Activin signalling	
SMAD1	ENSG00000170365	SMAD signalling	
SMAD2	ENSG00000175387	SMAD signalling	
SMAD3	ENSG00000166949	SMAD signalling	
SMAD4	ENSG00000141646	SMAD signalling	
SMAD5	ENSG00000113658	SMAD signalling	
SMAD6	ENSG00000137834	SMAD signalling	
SMAD7	ENSG00000101665	SMAD signalling	
SMAD9	ENSG00000120693	SMAD signalling	SMAD8

2.1.1.2 Stemformatics

Access link: https://www.stemformatics.org

Table 2-2 Eligible microarray datasets consisting of bone marrow cells obtained from Stemformatics

No.	Study	Dataset ID	Platform	Cell types	No. of samples	Sample ID	Pubmed ID
1	Affer et al., 2011	6324	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	HPC	8	6324_GSM609[354-357, 366-369]	21436996
2	Aghajanova et al., 2010	6284	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	3	6284_GSM518[039, 043-44]	20147733
3	Alves et al., 2012	6183	Affymetrix Human Genome U133A 2.0 Array [HG-U133A_2]	MSC	61	6183_GSM971[351-411]	22927939
4	Andersen et al., 2015	6965	Illumina HumanRef-8 v3.0 Expression BeadChip	SC	16	6965_GSM1698[519-534]	26450135
5	André et al., 2013	6395	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	3	6395_GSM894[405-407]	23555770
6	Benisch et al., 2012	6216	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	9	6216_GSM878[095-103]	23028809

Table 2-2 Eligible microarray datasets consisting of bone marrow cells obtained from Stemformatics (continue)

No.	Study	Dataset ID	Platform	Cell types	No. of samples	Sample ID	Pubmed ID
7	Dani et al., 2013	6206	Illumina HumanRef-8 v3.0 Expression BeadChip	MSC	4	6206_GSM698[427, 433, 437, 439]	22367914
8	Danieli et al., 2015	6624	Illumina HumanHT-12 v3.0 Expression BeadChip	MSC	10	6624_GSM1498[671-680]	25824141
9	Delorme et al., 2010	6037	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	3	6037_GSM250[019-021]	19905894
				EPC	3	6037_GSM250[022-024]	
				GC	3	6037_GSM250[025-027]	
				LC	3	6037_GSM250[028-030]	
10	Desterke et al., 2015	6629	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	MSC	6	6629_GSM1084[994-999]	25840601
11	Dickinson (2017)	6625	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	8	6625_GSM1322[103-110]	28833807
12	Espagnolle N, Guilloton F, 2013	6203	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	4	6203_GSM1180[553, 556, 559, 562]	24188055
13	Gerber et al., 2013	6322	Affymetrix Human Exon 1.0 ST Array [probe set (exon) version] [HuEx-1_0-st]	HPC	10	6322_GSM1070[369-370, 374-376, 379-380,384-386]	23651669
14	Granchi et al., 2010	6518	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	12	6518_GSM308[050-051, 055-056, 060-061, 064-065, 067, 071, 075, 079]	19686055
				MSCDI	7	6518_GSM308[049, 053-054, 058-059, 062-063]	
				OB	4	6518_GSM308[030, 034, 037, 041]	
				ODI	23	6518_GSM308[031-033, 035-036, 038-040, 042-044, 068-070, 072-074, 076-078, 080-082]	
15	Hamidouche et al., 2009	6480	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	3	6480_GSM250[019-021]	19843692
16	Hu et al., 2014	6385	Agilent-028679 SurePrint G3 Human Exon 4x180K Microarray	MSC	6	6385_GSM1347[854-859]	24758227
17	James et al., 2015	6778	Agilent SurePrint G3 Human GE 8x60K Microarray G4851A	MSC	3	6778_FH181-[1-3]	26070611
18	Jansen et al., 2010	6101	Affymetrix Human Exon 1.0 ST Array [transcript (gene) version] [HuEx-1_0-st]	MSC	3	6101_GSM464[392, 394, 395]	19788395
19	Koch et al., 2013	6199	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	6	6199_GSM909[604-609]	23080539
20	Larson BL, Ylöstalo J, and Prockop DJ, 2008	6332	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	6	6332_GSM241[198-203]	17916801
21	Lim et al., 2012	6450	Agilent-028004 SurePrint G3 Human GE 8x60K Microarray	MSC	3	6450_GSM952[617-619]	22665977
22	Lojewski et al., 2015	6865	Affymetrix Human Genome U133A Array [HG-U133A]	PC	3	6865_GSM1527[968-970]	26304036
23	Markov et al., 2007	6370	Affymetrix Human Genome U133A Array [HG-U133A]; Affymetrix Human Genome U133B Array [HG-U133B]	MSC	3	6370_GSM139[891_907-893_909]	17348805
24	Martins et al., 2014	6393	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	2	6393_GSM1289[026-027]	24438697

Table 2-2 Eligible microarray datasets consisting of bone marrow cells obtained from Stemformatics (continue)

No.	Study	Dataset ID	Platform	Cell types	No. of samples	Sample ID	Pubmed ID
25	Matigian et al., 2014	6064	Illumina HumanRef-8 v2.0 Expression BeadChip	MSC	4	6064_4294077[037_A, 037_D, 037_G, 038_B]	26484151
26	Menge et al., 2012	6123	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	4	6123_GSM997[928-931]	23175708
27	Mrugala et al., 2009	6119	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	6	6119_GSM260[657-662]	19196040
28	Pandya et al., 2017	7277	Affymetrix Human Gene 2.0 ST Array [HuGene-2_0-st]	MO	3	7277_GSM2067[560-562]	28253233
29	Pang et al., 2011	6661	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	HPC	27	6661_GSM81[2988-3014]	22123971
30	Papadimitropoulos et al., 2014	6416	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	10	6416_GSM1277[638-647]	25020062
31	Paul et al., 2013	6266	Illumina HumanHT-12 v3.0 Expression BeadChip	MSC	4	6266_GSM1174[437-440]	24090675
32	Rapin et al., 2014	6358	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	HPC	4	6358_GSM1041[380-383]	24363398
				MPP	2	6358_GSM1041[384-385]	
				CMP	3	6358_GSM1041[386-388]	
				GMP	5	6358_GSM1041[389-393]	
				MEP	2	6358_GSM1041[394-395]	
				PM	6	6358_GSM1041[396-401]	
				MY	2	6358_GSM1041[402-403]	
				MM	3	6358_GSM1041[404-406]	
				BC	4	6358_GSM1041[407-410]	
				PMN	3	6358_GSM1041[411-413]	
33	Reinisch et al., 2014	6343	Affymetrix Human Gene 2.0 ST Array [HuGene-2_0-st]	MSC	3	6343_GSM1376[241-243]	25406351
34	Roobrouck et al., 2011	6786	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	MSC	6	6786_GSM725[428, 431, 435, 438, 440, 458]	21433224
				MAPC	10	6786_GSM725[429-430, 433, 436-437, 439, 442, 447-448, 457]	
35	Royer-Pokora et al., 2010	6470	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	MSC	5	6470_GSM451[463-467]	20106868
36	Ryota et al., 2012	6208	Illumina HumanWG-6 v3.0 Expression BeadChip	MSC	3	6208_GSM928[749-751]	23142027
37	Tan et al., 2015	7120	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	6	7120_GSM1328[082-087]	26528540
38	Tanabe et al., 2008	6254	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	30	6254_GSM184[636-665]	18550633
39	Ullaha M, Sittingera M, and Ringe J, 2014	6334	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	3	6334_GSM906[367-369]	24269783
40	von-der-Heide et al., 2017	7173	Affymetrix Human Genome U133 Plus 2.0 Array [HG-U133_Plus_2]	MSC	4	7173_GSM2253[075-078]	27833093
41	Walenda et al., 2013	6345	Affymetrix Human Gene 1.0 ST Array [HuGene-1_0-st]	MSC	6	6345_GSM909[604-609]	24147049

2.1.2 Patient cohorts

2.1.2.1 Integrated Technologies for Improved Polyp Surveillance (INCISE) cohort

The INCISE project retrospectively identified patients who had participated in the Scottish Bowel Screening Programme in NHS Greater Glasgow & Clyde (GG&C) (49, 198). Patients were screened for a biannual guaiac faecal occult blood test (gFOBT), and a polypectomy at a screening colonoscopy after a subsequent positive stool test was known. The patients who have confirmed a histological premalignant polyp with adenoma or serrated polyp at their index screening colonoscopy between May 2009 and December 2016 were included in the INCISE cohort. If patients had been diagnosed with CRC, inflammatory bowel disease, a known inherited polyposis or CRC syndrome, they were excluded from the project. After having index endoscopy, patients were followed with further colonoscopy from 6 months to 6 years to identify metachronous polyps or CRC development. A total of 2,642 eligible precancerous colonic polyps were included in this cohort, which consisted of 1,256 Tissue MicroArray (TMA) cores. Ethics approval was obtained for the INCISE project (GSH/20/CO/002)

2.1.2.2 Early colorectal cancer stages cohort

This cohort retrospectively comprised of 159 cases of colorectal cancers from patients that had undergone surgical resection to remove their cancers which were detected via the cancerous polyps of Scottish Bowel Screening Programme in NHS GG&C between April 2009 and March 2011. Patients received a confirmed pathological diagnosis of CRC following colonoscopy screening. Tumours were staged using the 5th edition of TNM staging. Clinical data, including sex, age, TNM staging, tumour site, and histological differentiation grade, were deposited in the database.

2.1.2.3 Glasgow Royal Infirmary (GRI) cohort

The GRI cohort was recruited from retrospective cases of 787 patients diagnosed with stage I-III colorectal cancer between 1997 and 2012 at Glasgow Royal Infirmary, the Greater Glasgow and Clyde Health Board, under Safe Haven number GSH/21/ON/009. Tumours were staged using the 5th edition of TNM staging. Clinical data, including age, tumour site, TNM staging, histological differentiation

grade, and Mismatch repair (MMR) status, as well as pathological parameters such as Petersen Index (PI), Tumour budding (TB), Disease-free survival (DFS) status, Glasgow microenvironment score (GMS), Modified Glasgow Prognostic Score (mGPS), Klintrup-Mäkinen grade (KMG), and Tumour stroma percentage (TSP), were deposited in the database.

2.1.3 Cell lineage and tissue culture

2.1.3.1 Cell lineage

Table 2-3 CRC cell line information

Cell lines	DLD1	HCT116	SW620	SW837	COLO205	SW480
Lineage Subtype	Colorectal Adenocarcinoma	Colorectal Adenocarcinoma	Colorectal Adenocarcinoma	Colorectal Adenocarcinoma	Colorectal Adenocarcinoma	Colorectal Adenocarcinoma
Cancer Type Details	Colon Adenocarcinoma	Colon Adenocarcinoma	Colon Adenocarcinoma	Rectal Adenocarcinoma	Colon Adenocarcinoma	Colon Adenocarcinoma
Primary/Metastasis	Primary	Primary	Metastasis to lymph node	Primary	Metastasis to ascites	Primary
Media	RPMI, 10% FBS, and 1% L-Glutamine	RPMI, 10% FBS, and 1% L-Glutamine	RPMI, 10% FBS, and 1% L-Glutamine	DMEM, 10% FBS, and 1% L-Glutamine	DMEM, 10% FBS, and 1% L-Glutamine	RPMI, 10% FBS, and 1% L-Glutamine
Molecular characteristics						
MSI	MSI	MSI	MSS	MSS	MSS	MSS
BRAF(V600)	WT	WT	WT	WT	p.V600E	WT
KRAS	p.G13D	p.G13D	p.G12V	p.G12C	WT	p.G12V
TP53	p.S241F	WT	p.R273H & p.P309S	p.R248W	p.Y107Ter	p.R273H & p.P309S
CMS group	CMS1	CMS4	CMS4	CMS4	N/A	CMS4
CIMP	CIMP+	CIMP+	CIMP-	CIMP+	CIMP+	CIMP-
Clinical Information						
Gender	Male	Male	Male	Male	Male	Male
Ethnicity	N/A	N/A	White	White	White	White
Age at sampling	67	48	51	53	70	50

Information from
1) https://www.atcc.org
2) https://depmap.org/portal/
3) Berg et al. Molecular Cancer (2017) 16:116

Table 2-4 Other cell line information

Cell lines	HeLa	HepG2	CAL-72	MSC
Lineage Subtype	Cervical Adenocarcinoma	Hepatocellular carcinoma	Osteosarcoma	Mesenchymal stem cell
Cancer Type Details	Cervical Adenocarcinoma	Hepatocellular carcinoma	Osteosarcoma	Healthy
Primary/Metastasis	Primary	Primary	Primary	Primary
Media	DMEM, 10% FBS, and 1%L-Glutamine	DMEM, 10% FBS, and 1%L-Glutamine	DMEM, 10% FBS, 1 mM L-Glutamine, 1% Pen/Strep, 1% Insulin-Transferrin-Selenium	MSC Growth Medium 2, Supplement Mix, and 1% Pen/Strep
Clinical Information				
Gender	Female	Male	Male	N/A
Ethnicity	Black	White	N/A	N/A
Age at sampling	31	15	10	N/A

Information from 1) https://www.atcc.org
2) https://depmap.org/portal/
3) https://www.dsmz.de

2.1.3.2 Tissue culture plastics

Table 2-5 Tissue culture plastics

Product	Manufacturer	Catalogue number
Bijou bottles (7 mL, 20 mL)	Greiner Bio-One, UK	189176, 201172
Cell scrapers	Greiner Bio-One, UK	541080
Conical Centrifuge Tubes (15 mL, 50 mL)	Fisher Scientific, UK	14-959-49B, 14-432-22
Corning 96 well plate TC-treated flat bottom	Scientific laboratory supplies, UK	3596
Cryovials (2 mL)	Greiner Bio-One, UK	122263
Flat-bottomed multi-well culture plates (6, 12, 24, 96 wells)	Greiner Bio-One, UK	M8562, M9187, M8812, M8687
Nunc™ Cell-Culture Treated Multidishes, 6-well plate	Thermo Fisher, UK	140685
Nunc™ EasYFlask™ Cell Culture Flasks with Filter (25, 75, and 175 cm^2)	Thermo Fisher, UK	156367, 156499, 159910
Pasteur plastic pipettes (3 mL)	Greiner Bio-One, UK	612398
Pipette tips (p10, p200, p1000)	Greiner Bio-One, UK	765288, 775350,777350
Reaction Tubes (0.5 mL, 1.5 mL, 2 mL)	Greiner Bio-One, UK	667201, 616201, 623201
Syringes (5 mL, 10 mL, 20 mL, 60 mL)	Fisher Scientific, UK	14955452, 14955453, 14955454, 14955455
Syringe filters (0.2 µm, 0.45 µm)	Fisher Scientific, UK	15206869, 15216869
Tissue culture pipettes, sterile (5 mL, 10 mL, 25 mL)	Fisher Scientific, UK	10127400, 10677341, 10606151

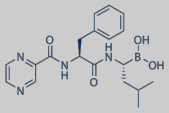
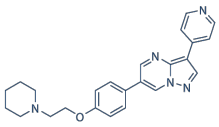
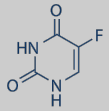
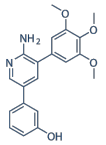
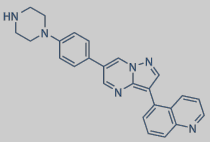
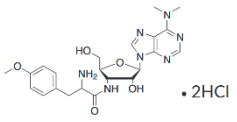
2.1.3.3 Tissue culture reagents

Table 2-6 Tissue culture reagents

Product	Manufacturer	Catalogue number
2-propanol	Honeywell, Germany	24137-251
Bovine Albumin Fraction V (7.5% solution)	Gibco, UK	15260037
Cell Proliferation Reagent WST-1 solution	Roche, UK	5015944001
Crystal Violet solution	Sigma-Aldrich, US	HT90132-1L
Cultrex Reduced Growth Factor Basement Membrane Extract, Type R1	Bio-Techne Ltd (R&D systems), UK	3433-005-R1
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich, UK	D8418, D2438
Distilled Water	Thermo Fisher, UK	15230204
DMEM	Gibco, UK	11965092
DPBS, powder, calcium, magnesium	Gibco, UK	21300058
Ethanol absolute	Sigma-Aldrich, UK	459844
Fetal Bovine Serum (FBS)	Gibco, UK	A5256701
Hoechst 33342 solution (20 mM)	Thermo Fisher, UK	62249
Insulin-Transferrin-Selenium (ITS -G) (100X)	Gibco, UK	41400045
L-Glutamine (200 mM)	Gibco, UK	A2916801
Mesenchymal Stem Cell Growth Medium 2	PromoCell, Germany	C-28009
Methanol	Sigma-Aldrich, France	32213-2.5L-M
Penicillin-Streptomycin (10,000 U/mL)	Gibco, UK	15140122
Polybrene Infection / Transfection Reagent	Sigma-Aldrich, UK	TR-1003
RPMI 1640 Medium	Gibco, UK	21870076
Trypan blue solution	Sigma-Aldrich, UK	T8154
Trypsin-EDTA (0.25%)	Gibco, UK	25200072

2.1.3.4 Inhibitors, antibiotics and compounds

Table 2-7 Inhibitors, antibiotics and compounds

Product	Chemical structure	Manufacturer	Catalogue number
Bortezomib		Stratech, UK	S1013
Dorsomorphin (DOR)		Abcam, UK	ab120843
5-Fluorouracil (5-FU)		Stratech, UK	S1209
K02288		Stratech, UK	S7359
LDN-212854		Stratech, UK	S7147-SEL
Puromycin (10 mg/mL solution)		InvivoGen, UK	P9620
Human bone morphogenetic protein 4 (hBMP4), Animal-Free Recombinant Protein	N/A	Peprotech, UK	AF-120-05ET
Human epidermal growth factor (EGF), Animal-Free Recombinant Protein	N/A	Peprotech, UK	AF-100-15-100UG

2.1.4 Immunohistochemistry (IHC)

2.1.4.1 Tissue Micro Arrays (TMAs)

Colorectal tissues in this study were collected under research projects of the University of Glasgow. Glasgow Tissue Research Facility (GTRF) constructed TMAs from Formalin-Fixed Paraffin-Embedded (FFPE) tissue blocks. Four 0.6 mm cores were removed from each specimen to construct the precancerous polyp TMA, and three 0.6 mm cores were taken from each patient for the cancer arrays to account for tissue heterogeneity. TMAs were constructed using 3DHISTECH TMA Grandmaster. All projects have ethics approval from the West of Scotland Research Ethics Service (WoSRES) with REC reference 22/WS/0020.

2.1.4.2 IHC consumables

Table 2-8 IHC consumables

Product	Manufacturer	Catalogue number
Acid Alcohol 0.5%	Leica Biosystems, US	3803650E
Casein Solution 10X	Vector Laboratories, US	NC9965082
Citric acid	Sigma Aldrich, UK	27109
Dako Antibody diluent	Agilent Technologies, US	S0809
Dako pen	Agilent Technologies, UK	S2002
Ethanol absolute	VWR Chemicals, France	20821.330
Harris haematoxylin	Leica Biosystems, US	3801560E
Histoclear	National Diagnostics, US	HS-202
Hydrogen peroxide 30%	VWR Chemicals, France	23619.297
ImmPACT® DAB Substrate Kit, Peroxidase (HRP)	Vector Laboratories, US	SK-4105
ImmPRESS® HRP Universal Antibody (Horse Anti-Mouse/Rabbit IgG) Polymer Detection Kit, Peroxidase	Vector Laboratories, US	MP-7500-50
Pertex mounting medium	CellPath, UK	SEA-0100-00A
Scott's Tap Water Substitute	Leica Biosystems, US	3802901E
Sodium chloride	VWR Chemicals, Belgium	27810.295
Tri-sodium citrate dihydrate	Fisher Scientifics, UK	S/3320/53
Tris Base	Fisher bioreagents, UK	BP152-1

2.1.5 Flow cytometry (FACS)

2.1.5.1 FACS consumables

Table 2-9 FACS consumables

Product	Manufacturer	Catalogue number
7-AAD	BD Biosciences, UK	559925
BD Cytofix/Cytoperm™ Fixation and Permeabilization Solution	BD Biosciences, UK	554722
Bovine Serum Albumin (BSA), Fraction V	Roche, Germany	10735086001
CellTrace™ Violet Cell Proliferation Kit, for flow cytometry	Invitrogen, UK	C34557
Colcemid (10 µg/mL)	Roche, UK	10295892001
DAPI solution	BD Biosciences, UK	564907
Falcon™ Round-Bottom Polypropylene Test Tubes With Cap	Fisher Scientific, UK	352063
Falcon™ Round-Bottom Polystyrene Test Tubes	Fisher Scientific, UK	352008
FITC Annexin V	BD Biosciences, UK	556419
HBSS (10X), calcium, magnesium, no phenol red	Gibco, UK	14065056

Table 2-9 FACS consumables (continue)

Product	Manufacturer	Catalogue number
HBSS (10X), no calcium, no magnesium, no phenol red	Gibco, UK	14185052
Ki-67 Monoclonal Antibody (SolA15), PE-Cyanine7, eBioscience™, 100 µg	Invitrogen, UK	25-5698-82
Sodium azide	Sigma-Aldrich, UK	S8032
Sterile Cell Strainer 70 µm Nylon Mesh	Fisher Scientific, UK	22363548
Triton™ X-100 solution	Sigma-Aldrich, UK	93443

Table 2-10 1X wash/perm solution

Reagents	Final concentration	1X
BSA	1%	1 g
Sodium azide	0.10%	100 mg
Triton-X	0.50%	0.5 mL
PBS or dH2O	-	100 mL

2.1.6 Molecular biology

2.1.6.1 RNA extraction, cDNA synthesis, and qPCR consumables

Table 2-11 Molecular biology materials and reagents

Product	Manufacturer	Catalogue number
RNeasy Plus Micro Kit	Qiagen, US	74034
Ethanol absolute	Sigma-Aldrich, UK	459844
High-Capacity cDNA Reverse Transcription Kit	Applied Biosystems, UK	4368814
Water, nuclease-free	Thermo Fisher, UK	R0581
PowerTrack™ SYBR Green Master Mix for qPCR	Applied Biosystems, UK	A46110
MicroAmp™ Optical 8-Tube Strip with Attached Optical Caps, 0.2 mL	Applied Biosystems, UK	A30588
MicroAmp™ Optical Adhesive Film	Applied Biosystems, UK	4311971
MicroAmp™ Optical 384-Well Reaction Plate with Barcode	Applied Biosystems, UK	4309849
RNaseZap™ RNase Decontamination Solution	Invitrogen, UK	AM9780

2.1.6.2 Primers

Table 2-12 Primers

Targeted genes	Primer sequence	Product size (bp)	Reference transcripts
BMP2 F	AGACCTGTATCGCAGGCACT	138	NM_001200.2
BMP2 R	CCACTCGTTTCTGGTAGTTCTTCC		
BMP5 F	GGACTCCTCCAGAATGTCCAG	123	NM_021073.4
BMP5 R	GGTGCTATAATCCAGTCCTGCC		
BMP7 F	CCAACGTGGCAGAGAACAGC	111	NM_001719.3
BMP7 R	CTTCAGGCGCGATGATCCAG		
BMPR1A F	ACGCCGGACAATAGAATGTTGTC	118	NM_004329.2
BMPR1A R	GAGCAAAACCAGCCATCGAATG		
BMPR1B F	GAAGATCAATTGAATGCTGCACAGAA	112	NM_001203.2
BMPR1B R	GCCCTGTGGTGTATAGGTCCA		
BMPR2 F	TTGCCCTCCTGATTCTTGGC	121	NM_001204.7
BMPR2 R	CTTGATTCTGCGAAGCAGCC		
SMAD1 F	GCTGCTCTCCAATGTTAACCG	113	NM_001003688.1
SMAD1 R	CACTAAGGCATTCGGCATACAC		
SMAD2 F	CCACGGTAGAAATGACAAGAAGG	123	NM_001003652.3
SMAD2 R	GATTACAATTGGGGCTCTGCAC		
SMAD3 F	GTCTGCGTGAATCCCTACCAC	125	NM_001145102.1
SMAD3 R	GGGATGGAATGGCTGTAGTCG		

Table 2-12 Primers (continue)

Targeted genes	Primer sequence	Product size (bp)	Reference transcripts
SMAD4 F	GGGTCAACTCTCCAATGTCCAC	121	NM_005359.5
SMAD4 R	GTCACTAAGGCACCTGACCC		
SMAD5 F	TGGGTCAAGATAATTCCCAGCCT	106	NM_001001419.2
SMAD5 R	GGCTCTTCATAGGCAACAGGC		
SMAD6 F	CTCCCTACTCTCGGCTGTCT	120	NM_005585.5
SMAD6 R	AGAATTCACCCGGAGCAGTG		
SMAD7 F	CCATCACCTTAGCCGACTCT	147	NM_005904.4
SMAD7 R	CCAGGGGCCAGATAATTCGT		
SMAD8/9 F	CTTATCATGCCACAGAAGCCTCT	120	NM_001127217.2
SMAD8/9 R	GCTCCTCGTAACAAACTGGTCG		
TGFB1 F	CAGCAGGGATAACACACTGC	109	NM_000660.7
TGFB1 R	CATGAGAAGCAGGAAAGGCC		
TGFB2 F	GCGACGAAGAGTACTACGCC	78	NM_003238.6
TGFB2 R	GCGGGATGGCATTTTCGGAG		
TGFBR1 F	CGTTCGTGGTTCCGTGAGG	128	NM_001130916.2
TGFBR1 R	TAATCTGACACCAACCAGAGCTG		
TGFBR2 F	CTAACAGTGGGCAGGTGGG	130	NM_003242.6
TGFBR2 R	ATTTCCCAGAGCACCAGAGC		
ACTB F	CGCCGCCAGCTCACC	120	NM_001101.5
ACTB R	CACGATGGAGGGGAAGACG		
B2M F	TTGTCTTTCAGCAAGGACTGG	172	NM_004048.2
B2M R	ATGCGGCATCTTCAAACCTCC		

Primers were manufactured by Integrated DNA Technologies

2.1.7 Western blotting

2.1.7.1 Western blotting consumables

Table 2-13 Western blotting consumables

Product	Manufacturer	Catalogue number
2-Mercaptoethanol	Sigma-Aldrich, Germany	M6250-250ML
Acrylamide/Bis-acrylamide solution 30%	Sigma-Aldrich, US	A3699-100ML
Albumin from chicken egg white	Sigma-Aldrich, UK	A5378-10G
Ammonium Persulfate, APS (228.20 g/mol)	Sigma-Aldrich, UK	A3678-25G
Bovine Serum Albumin Fraction V	Roche, Germany	10735086001
Bromophenol blue sodium salt (691.94 g/mol)	Sigma-Aldrich, US	B8026-5G
Combs, 15 wells, 0.75 mm	Bio-Rad, US	1653355
cOmplete™ ULTRA Tablets	Roche, UK	5892791001
Flat bottom 96-well microplate	Greiner, Germany	655101
Glycerol	Sigma-Aldrich, US	G5516-100ML
Glycine (75.07 g/mol)	Sigma-Aldrich, US	G8898-1KG
Immobilon Forte Western HRP substrate, 100 mL	Millipore, UK	WBLUF0100
Methanol	Sigma-Aldrich, France	32213-25L
N,N,N′,N′-Tetramethylethylenediamine, TEMED	Sigma-Aldrich, China	T9281-25ML
Nitrocellulose membrane, 0.45 µm	Thermo Fisher, Germany	88018
Nonidet™ P 40 Substitute (1.56 M, 1.06 g/mL)	Fluka, Switzerland	74385
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa	Thermo Fisher, UK	26619
PhosSTOP™	Roche, UK	4906837001
Polyclonal Goat Anti-Mouse Immunoglobulins/HRP	Agilent Technologies, UK	P0447
Polyclonal Goat Anti-Rabbit Immunoglobulins/HRP	Agilent Technologies, UK	P0448
Ponceau S solution	Sigma-Aldrich, US	P7170-1L
Quick Start Bovine Serum Albumin Standard (2 mg/mL)	Bio-Rad, US	5000206
Quick Start Bradford 1X Dye reagent	Bio-Rad, US	5000205
Restore™ Western Blot Stripping Buffer	Thermo Fisher, UK	21059
Short and spacer plates with 0.75 mm	Bio-Rad, US	1653308, 1653310
Sigma 7-9®, Trizma base (121.14 g/mol)	Sigma-Aldrich, US	T1378-5kg
Skim Milk Powder	Merck, UK	70166-500G
Sodium azide (65.01 g/mol)	Sigma, Switzerland	S8032-100G
Sodium Chloride, NaCl (58.44 g/mol)	Sigma-Aldrich, UK	S9888-1KG
Sodium deoxycholate (414.55 g/mol)	Sigma-Aldrich, UK	D6750-25G
Sodium dodecyl sulfate, SDS (288.38 g/mol)	Sigma-Aldrich, UK	L4509-500g

Table 2-13 Western blotting consumables (continue)

Product	Manufacturer	Catalogue number
Sodium Fluoride, NaF (41.99 g/mol)	Sigma-Aldrich, UK	215309
SuperSignal™ West Femto Maximum Sensitivity Substrate	Thermo Fisher, UK	34094
Tween20	Sigma-Aldrich, France	P1379-1L
Whatman® gel blotting paper	Whatman, US	WHA10427826

Table 2-14 Preparation of western blotting stock buffer solutions

Buffers	Preparation
1M Tris-HCL pH 6.8 (121.14 g/mol)	60.6 g Trisma base in 500 mL dH_2O (Adjusted pH)
1M Tris-HCl pH 7.4 (121.14 g/mol)	60.6 g Trisma base in 500 mL dH_2O (Adjusted pH)
1M Tris-HCl pH 7.5 (121.14 g/mol)	60.6 g Trisma base in 500 mL dH_2O (Adjusted pH)
1M Tris-HCl pH 8.8 (121.14 g/mol)	60.6 g Trisma base in 500 mL dH_2O (Adjusted pH)
10% (v/v) NP40	25 mL NP40 in 250 mL dH_2O
10% (w/v) Sodium deoxycholate	1 g Sodium deoxycholate in 10 mL dH_2O
10% (w/v) Sodium dodecyl sulfate (SDS)	10 g SDS in 100 mL dH_2O
5 M Sodium Chloride (NaCl) (58.44 g/mol)	2.922 g in 10 mL dH_2O
1 M Sodium Fluoride (NaF) (41.99 g/mol)	420 mg in 10 mL dH_2O
10% (w/v) APS	0.1 g APS in 1000 µL dH_2O
10X SDS-Page Running Buffer (0.25 M Tris, 1.92 M Glycine and 1% (w/v) SDS)	30.3 g Trizma base with 144.4 g Glycine and 10 g SDS in 1000 mL dH_2O
10X Semi-Dry Transfer Buffer (0.48 M Tris, 0.39 M Glycine and 0.375% (w/v) SDS)	58.1 g Trizma base with 29.3 g Glycine and 3.75 g SDS in 1000 mL dH_2O
10X Wet-Transfer Buffer (0.25 M Tris and 1.92 M Glycine)	30 g Trizma Base and 144 g Glycine in 1000 mL dH_2O
10X Tris-Buffered Saline (TBS) pH 7.5 (200 mM Tris and 1.5 M NaCl)	24.2g Trisma base and 87.6 g NaCl in 1000 mL dH_2O (Adjusted pH)

Table 2-15 Preparation of protein lysis buffer (RIPA buffer)

Reagents	Stock concentration	Final concentration	1X
dH_2O	-	-	525 µL
NP40	10% (v/v)	1% (v/v)	100 µL
Sodium deoxycholate	10% (w/v)	1% (w/v)	100 µL
Sodium dodecyl sulfate (SDS)	10% (w/v)	0.1% (w/v)	10 µL
Tris-HCl pH 7.4	1 M	25 mM	25 µL
Sodium Chloride (NaCl)	5 M	150 mM	30 µL
Sodium Fluoride (NaF)	1 M	10 mM	10 µL
Add inhibitors below before use			
Protease inhibitor cocktail	10X	1X	100 µL
Phosphatase inhibitor	10X	1X	100 µL
Total volume	-	-	1000 µL

Table 2-16 Preparation of 5X sample buffer

Reagents	Stock concentration	Final concentration	1X
Sodium Dodecyl Sulphate (SDS)	288.38 g/mol	10% (w/v)	100 mg
dH_2O	-	-	240 µL
Tris-HCL pH 6.8	1 M	200 mM	200 µL
Glycerol	-	50% (v/v)	500 µL
Bromophenol blue	-	-	10 µL
Add reducing reagent below before use			
2-mercaptoethanol	-	5% (v/v)	50 µL
Total volume	-	-	1000 µL

Table 2-17 Preparation of a 0.75 mm acrylamide gel for hand casting

Gel	Running gel			Stacking gel
Protein size	> 50kDa	35-50 kDa	< 40 kDa	
% Gel	7.50%	10%	12%	5%
ddH_2O	5.6 mL	4.35 mL	3.35 mL	6 mL
1M Tris-HCl pH 8.8	5.6 mL	5.6 mL	5.6 mL	-
1M Tris-HCl pH 6.8	-	-	-	1.25 mL
10% (w/v) SDS	0.25 mL	0.25 mL	0.25 mL	0.15 mL
30% Acrylamide	3.75 mL	5.0 mL	6.0 mL	1.67 mL
Add reagents below before use				
10% (w/v) APS	100 µL	100 µL	100 µL	50 µL
TEMED	20 µL	20 µL	20 µL	20 µL
Total volume	~ 15 mL	~ 15 mL	~ 15 mL	~ 10 mL

Table 2-18 Preparation of Western blot working buffer solutions

Buffers	Final concentration	Preparation
1X SDS-Page Running Buffer	25 mM Trizma base with 192 mM Glycine and 0.1% (w/v) SDS	100 mL of 10X SDS-Page Running Buffer and 900 mL of dH2O
1X Semi-Dry Transfer Buffer	39 mM glycine with 48 mM Trizma base, 0.0375% SDS , and 20% (v/v) Methanol	40 mL of Methanol, 20 mL of 10X Semi-Dry Transfer Buffer and 140 mL of dH2O
1X Wet-Transfer Buffer	25 mM Tris, 192 mM glycine, and 20% (v/v) Methanol	100 mL of 10X Wet-Transfer Buffer, 200 mL of Methanol and 700 mL of dH2O
1X Tris-Buffered Saline, TBS	20 mM Tris-HCL pH 7.5 and 150 mM NaCl	100 mL of 10X TBS in 900 mL dH2O
1X Tris-Buffered Saline NP40, TBSN	20 mM Tris-HCL pH 7.5 and 150 mM NaCl and 0.05% (v/v) NP40	5 mL of 10% NP40 in 1000 mL of 1X TBS
1X Tris-Buffered Saline Tween, TBST	20 mM Tris-HCL pH 7.5 and 150 mM NaCl and 0.1% (v/v) Tween	1 mL of Tween in 1000 mL of 1X TBS
5% BSA blocking solution	5% (w/v) BSA, 1% (w/v) Ovalbumin and 0.02% (w/v) Sodium azide	2.5 g BSA, 500 mg Ovalbumin and 10 mg Sodium azide in 50 mL of 1X TBSN/TBST
1% BSA blocking solution	1% (w/v) BSA, 0.2% (w/v) Ovalbumin and 0.004% (w/v) Sodium azide	2 mL of 5X BSA blocking solution and 8 mL of 1X TBSN/TBST
5% milk blocking solution	5% (w/v) milk	2.5 g milk powder in 50 mL of 1X TBST
0.5% milk blocking solution	0.5% (w/v) milk	1 mL of 5X milk blocking solution and 9 mL of 1X TBST

2.1.7.2 Antibodies

Table 2-19 Antibodies

Antibodies	Protein size (kDa)	% Gel	Blocking	Primary Antibody	Secondary Antibody	Manufacturer	Catalogue number
p-MEK 1/2 (S217/221)	45	10%	5X BSA in TBSN	1:1000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	9121S
MEK 1/2 (L38C12)	45	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	Cell Signaling Technology, US	4694S
Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E)	44-42	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	4370S
p44/42 MAPK (Erk1/2)	42-44	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	9102
Phospho-GSK-3β (Ser9)	46	10%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Cell Signaling Technology, US	9336S
Anti-GSK-3β (PY216)	46	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	BD Transduction Laboratories, UK	612313
GSK-3β (E-11)	47	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	Santa Cruz Biotechnology, US	SC-377213
Anti-Active-β-Catenin (Anti ABC), clone 8E7	92	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	Millipore, UK	05-665
Anti-β-Catenin	92	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	BD Transduction Laboratories, UK	610154
β-TrCP (D13F10)	62	7.5%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Cell Signaling Technology, US	4394S
phospho-smad2(ser465-467)/smad3(ser423-425)	52-60	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	9510S
Anti-SMAD2-clone EP567Y	58	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Millipore, UK	04-1029
Anti-phospho-Smad3 (Ser423/425)	50	7.5%	5% Milk in TBST	1:2000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Millipore, UK	07-1389
Anti-SMAD3, clone EP568Y	58	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Millipore, UK	04-1035
Phospho-SMAD1 (Ser463/465)/ SMAD5 (Ser463/465)/ SMAD9 (Ser465/467) (D5B10)	60	7.5%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Cell Signaling Technology, US	13820S
SMAD1/5/8 (N-18)-R	52-56	7.5%	5% Milk in TBST	1:2000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Santa Cruz Biotechnology, US	SC-6031-R
SMAD4 (D3R4N) XP®	70	7.5%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Cell Signaling Technology, US	46535S
Phospho-Akt Substrate (RXXS*/T*) (110B7E)	25-180	7.5%	5X BSA in TBSN	1:1000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	9614S
Cleaved Caspase-3 (Asp175) (5A1E)	17, 19	12.0%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Rabbit HRP in 0.5% Milk	Cell Signaling Technology, US	9664S
Caspase-3 (3G2)	17, 19, 35	12.0%	5% Milk in TBST	1:1000 in 0.5% Milk	1:10000 Mouse HRP in 0.5% Milk	Cell Signaling Technology, US	9668S

Table 2-19 Antibodies (continue)

Antibodies	Protein size (kDa)	% Gel	Blocking	Primary Antibody	Secondary Antibody	Manufacturer	Catalogue number
Phospho-Histone H2A.X (Ser139) (20E3)	15	12.0%	5X BSA in TBST	1:1000 in 1X BSA	1:10000 Rabbit HRP in TBST	Cell Signaling Technology, US	9718S
Histone H2A.X	15	12.0%	5X BSA in TBST	1:1000 in 1X BSA	1:10000 Rabbit HRP in TBST	Cell Signaling Technology, US	2595S
β-Tubulin	55	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Rabbit HRP in TBSN	Cell Signaling Technology, US	2146
SH-PTP2 (B-1)	70	7.5%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	Santa Cruz Biotechnology, US	SC-7384
β-Actin (C4)	43	10%	5X BSA in TBSN	1:2000 in 1X BSA	1:10000 Mouse HRP in TBSN	Santa Cruz Biotechnology, US	SC-47778

2.1.8 Spheroids culture

2.1.8.1 Spheroids culture plastics

Table 2-20 Spheroids culture plastics

Product	Manufacturer	Catalogue number
AggreWell™ 400 24-well plates	StemCell Technologies, Canada	#34415
General Purpose Dispense Tips (21G needle)	Nordson Corporation, US	7005005
Microscope Slides	Fisher Scientific, UK	11572203
Optimum Syringe Barrel Pistons	Nordson Corporation, US	7366134
Optimum Syringe Barrels (Cartridges)	Nordson Corporation, US	7366092
Petri dish	Greiner Bio-One, UK	632181
Reversible strainer 37 μm	Stem Cell Technologies, Canada	27250
Syringes (1 mL)	BD Biosciences, UK	303172

2.1.8.2 Spheroid culture and viability reagents

Table 2-21 Spheroids culture reagents

Product	Manufacturer	Catalogue number
Anti-Adherence Rinsing solution	StemCell Technologies, Canada	#07010
Calcium chloride hexahydrate	Sigma-Aldrich, Croatia	442909
LIVE/DEAD™ Viability/Cytotoxicity Assay Kit	Invitrogen, UK	L3224
Porcine Gelatin, type A	Sigma-Aldrich, UK	G1890
Sodium alginate	Sigma-Aldrich, China	W201502

2.2 Methods

2.2.1 *in silico* analysis - Accessing online gene expression databases and Visualisation

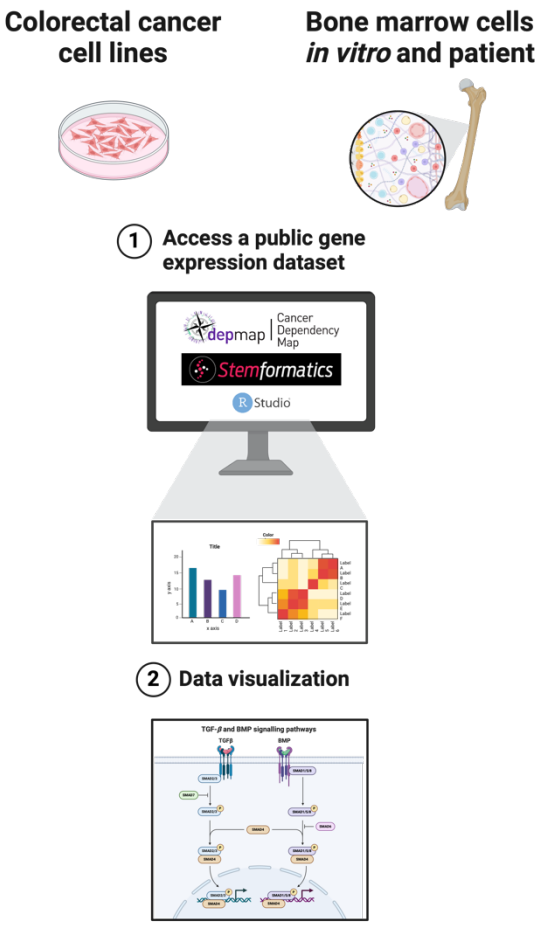


Figure 2-1 Workflow for bioinformatic analysis of TGF-β and BMP signalling pathways in colorectal cancer (CRC) and bone marrow (BM) niche

The DepMap and Stemformatics databases were assessed for expression data of TGF-β and BMP gene members in colorectal cancer cell lines as well as in healthy bone marrow cells.

2.2.1.1 Cancer Dependency Map (DepMap) portal for colorectal cancer

The Cancer Dependency Map (DepMap) integrates large-scale data from various research platforms on cancer cell lines (199). It aims to provide comprehensive data exploration tools in multi-omics and compound sensitivity profiles to reveal cancer vulnerabilities, landscape, and precision oncology insights. To identify which components of the TGF-β and BMP signalling pathways are dysregulated across CRC cells, we accessed this database portal to explore the gene expression

profile of CRC cell lineage. The public expression dataset from DepMap, transcription per million (TPM), was analysed through RStudio version 2024.12.0 (posit, US) using Bioconductor's ExperimentHub function by the query for EH6120 (200). Boxplots were created from the mRNA expression levels of interesting genes (Table 2-1) across cell lineages in the DepMap dataset. A bar chart was plotted to show gene expression levels comparing each CRC cell line. CSV files were retrieved and selected, and the data was collected from only the cell lines available in our laboratory for further interpretation. A heatmap was generated to visualise the expression levels across the different CRC cell lines. The R packages of ggplot2 and complexHeatmap package were utilised in data visualisation.

2.2.1.2 Stemformatics for healthy bone marrow

The Stemformatics web database was explored for projects containing gene expression datasets from bone marrow samples under normal conditions, both *in vitro* and *in vivo*. Various bone marrow (BM) cell types were selected, including Band cell (BC), Common myeloid progenitor (CMP), Erythropoietic cell (EPC), Granulocyte (GC), Granulocyte-macrophage progenitor cell (GMP), Haematopoietic precursor cell (HPC), Leukocyte (LC), Multipotent adult progenitor cell (MAPC), Megakaryocyte-erythroid progenitor cell (MEP), Metamyelocyte (MM), Macrophage (MO), Haematopoietic multipotent progenitor cell (MPP), Mesenchymal stromal cell (MSC), MSC differentiation intermediate (MSCDI), Myelocyte (MY), Osteoblast (OB), Osteogenic differentiation intermediate (ODI), Pericyte cell (PC), Promyelocyte (PM), Polymorphonuclear cell (PMN), and Stromal cell (SC). Log-normalised expression values for gene ID and sample tables from each project were downloaded as TSV files. Forty-one datasets were included up to September 29, 2022, and were categorised by BM cell types (Table 2-2). RStudio was utilised to average the gene expression levels for each BM cell type, which were visualised in a heatmap using the ComplexHeatmap package.

2.2.2 Patient tissue studies - Establishing prognostic significance

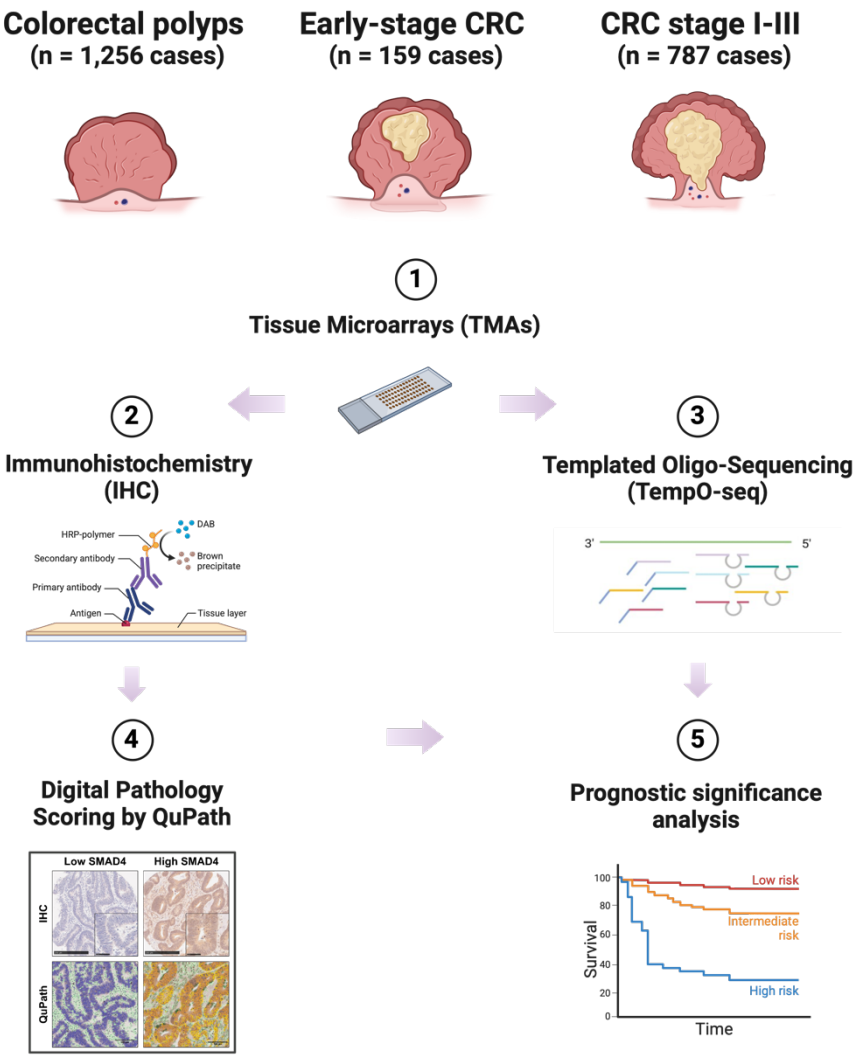


Figure 2-2 Workflow for determining the prognostic significance of SMAD4 expression in pre-malignant and CRC patient tissues

The TMAs of patient tissues, including pre-malignant and CRC samples, were examined for protein expression using immunohistochemistry. Additionally, an assessment of patient tissues for gene expression was conducted using TempO-seq. Protein expression levels were evaluated with QuPath for digital pathology scoring, and prognostic significance was determined.

2.2.2.1 Antibody specificity

To determine antibody specificity before using antibody staining on patient cohort tissues (201), we performed a western blot experiment in triplicate to confirm the specific binding protein target using Smad4 overexpression lysate as a positive control and CRC cell line lysates. Subsequently, FFPE CRC cell line pellets (provided by Dr Kathryn Pennel) and patient tissue TMA optimisation slides were stained with an IHC assay to assess staining patterns. This assay could detect staining on cellular cytoplasm and/or nuclei and was used to optimise the primary antibody concentration.

2.2.2.2 Immunohistochemistry (IHC)

Pre-existing TMAs sections were heated at 65°C for 30 mins to anneal tissue to the glass slides. All tissue slides were dewaxed in Histoclear (5 mins x 2 times), rehydrated through a decreasing ethanol gradient (Absolute ethanol 2 mins x 2 times, 90% ethanol 2 mins, 70% ethanol 2 mins), and slides rinsed in running water for 10 mins. Antigen retrieval was induced using Citrate buffer pH 6.0 (2 mM Tri-sodium citrate dihydrate (2.41 g) and 8 mM Citric acid (0.34 g) in 1 L distilled water) under pressure heating with a microwave, following the standard antigen retrieval method. The process involved pre-heating the buffer in a pressure cooker without a lid in a microwave for 14 mins, then placing slides into the pressure cooker, sealing the lid, and heating using the microwave for 7 mins. Notably, antigen retrieval for the cell pellet is shorter than for archival FFPE tissues—5 mins after adding slides to the pressure cooker. Afterwards, the slides were cooled to room temperature for 30 mins and rinsed in running water. Endogenous peroxidase activity was blocked in 3% H_2O_2 (10X dilution of 30% H_2O_2 in 400 mL distilled water) for 15 mins and rinsing in running water. Each slide was circled surrounding the tissue section using a Dako pen, and nonspecific antibody binding was blocked in 200 µL of 10% casein (10X dilution of casein solution in 1 mL Dako antibody diluent) for 30 mins at room temperature in dark. The tissues were stained with SMAD4 antibody at 1:100 in Dako antibody diluent overnight at 4°C and washed in 1X TBS for 5 mins twice. After that, the section slides were incubated in ImmPRESS® HRP Universal Antibody (Horse Anti-Rabbit IgG) Polymer Reagent (3-4 drops) for 30 mins at room temperature in dark and washed in 1X TBS for 5 mins twice. To visualise the staining, the sections were incubated

with ImmPACT® DAB Substrate, Peroxidase (HRP) (1 dye drop in 1 mL diluent) for 5 mins until brown colour developed and then rinsed in running water for at least 10 mins. Next, the slides were counterstained with Harris Haematoxylin for 1 min, excess staining was eliminated with Acid Alcohol for 5 secs, followed by Scott's Tap Water Substitute for 2 mins, and the slides were rinsed in running water for 2 mins during these steps. The staining tissues were dehydrated in an increasing gradient of ethanol (70% ethanol 1 min, 90% ethanol 1 min, and absolute ethanol 1 min x 2 times) and Histoclear for 1 min x 2 times and then were mounted with Pertex mounting medium before being scanned using Hamamatsu NanoZoomer Digital Slide Scanner and visualised in NDP.view2 Image viewing software (U12388-01, Hamamatsu, UK). Positive and negative control (no antibody) slides were done in the full section of colorectal cancer FFPE slides to confirm the staining and rule out nonspecific staining.

2.2.2.3 Weighted Histoscore (WHS) method

The Weighted Histoscore (WHS) method is a traditional procedure for evaluating IHC staining intensity (202). Stained tissue sections were assessed for levels of staining intensity on cellular cytoplasm, nuclei, and membranes depending on antibody binding target protein patterns. For the SMAD4 expression, we evaluated the stained intensity from cytoplasmic and nuclear epithelial cells at 20X magnification. Ten percent of TMA cores were manually scored by a trained observer blinded to clinical data. The score of each core was calculated from the equation below. The WHS is between 0 and 300.

$$\begin{aligned}\text{WHS} = {} & (\%\ \text{of unstained cells} \times 0) + (\%\ \text{of weakly stained cells} \times 1) \\ & + (\%\ \text{of moderately stained cells} \times 2) + (\%\ \text{of strongly stained cells} \times 3)\end{aligned}$$

2.2.2.4 Digital path and analysis

Quantitative Pathology & Bioimage Analysis (QuPath) has been developed to offer a comprehensive, practical, and reliable tool for the image analysis of whole tissue slides with functional features such as automated TMA dearraying, cell detection, and trainable cell classification, which supports large-scale image data in digital pathological assessment (203). All TMAs were evaluated for the IHC stained intensity of each TMA core using QuPath version 0.3.2. Briefly, a TMA was dearrayed, relying on the cohort TMA construction template. Cell detection,

annotation and classification were processed. For example, tumour cells (or epithelial cells), stroma, and/or necrosis were identified (Figure 2-3). After that, the intensity of staining was defined from the DAB channel and classified using accurate thresholds, being negative, weak, moderate, and strong staining for cytoplasm, nucleus, or total cellular (Table 2-22). Digital weighted histoscoring or QuPath score was confirmed as a reliability score with 10% WHS scores. The correlation coefficient and interclass correlation coefficient (ICC) were calculated and visualised in a Scatter plot and a Bland Altman plot using IBM SPSS statistics version 28.0 (IBM, US). An ICC of 0.7 is the minimum acceptable correlation between scores (202)

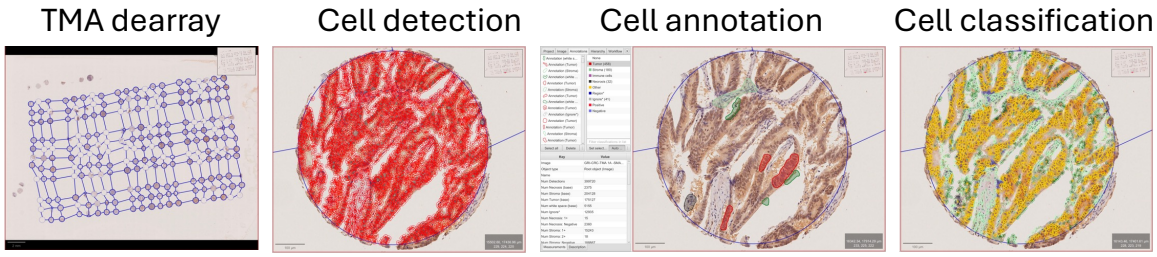


Figure 2-3 Image processing for digital pathology scoring of IHC staining on tissue microarrays (TMAs) using QuPath

An image of the IHC staining TMA was dearrayed, detecting all cells, annotating cell types, and classifying based on staining intensity.

Table 2-22 QuPath thresholds for classifying epithelial SMAD4 staining

Thresholds	1	2	3
Cytoplasm	0.18	0.37	0.58
Nuclei	0.24	0.7	0.92
Cell	0.21	0.48	0.83

2.2.2.5 Prognostic significance analysis

The training group of the INCISE cohort accounts for 867 cases that were excluded due to missing tissue cores, poor quality of tissue cores such as folding and large debris covering the tissue, and fewer than 10% of epithelial cells in the tissue cores. At least ten-millimetre polyps were included in 788 of the 822 cases with valid SMAD4 staining cores. We focused on analysing adenomatous polyps. Only 777 cases of adenomatous polyps were selected, while serrated polyps were excluded. The QuPath scores were obtained from 2 cores of epithelial cells from

basal and luminal polyps for each patient and then averaged to acquire a SMAD4 score. Given the strong correlation between cytoplasmic, nuclear, and cellular scores of SMAD4 expression in this cohort, we decided to use the cellular SMAD4 scores in the subsequent analysis. Cut points were determined using the survminer package in RStudio, which stratified patients into two groups: low and high SMAD4 expression. Kaplan–Meier (KM) survival curves and Cox regression were employed to examine the relationship between SMAD4 expression and metachronous polyp development in the total cohort, as well as in histological subgroups: tubular, tubulovillous, and villous polyps. Chi-squared tables were used to test the correlation between SMAD4 expression and clinicopathological characteristics using IBM SPSS statistics. Data visualisation was performed using RStudio packages, including survival, gplots, corrplot, and ggplot2.

In the cohort of early-stage colorectal cancer, which consisted of a smaller cancer group, only 129 cases of colorectal cancer stage T1/T2 with valid SMAD4 staining tissue cores were examined. The cytoplasmic and nuclear SMAD4 tumour scores were averaged from three tissue cores for each patient. Cut points for the cytoplasmic and nuclear scores were established, and both scores were combined, categorising them into three groups: low SMAD4 in both the cytoplasm and nuclei, high SMAD4 in either the cytoplasm or nuclei, and high SMAD4 in both the cytoplasm and nuclei. Due to the limitations of a short follow-up period for survival time, only the Chi-squared table was analysed.

The GRI cohort compiled data on CRC stages I-III from 787 cases, while the remaining 727 cases with valid SMAD4-staining tissue cores were excluded due to mortality within 30 days of surgery or receipt of neoadjuvant therapy. A total of 673 carcinoma cases were analysed. For each patient, three tissue cores were evaluated to calculate an average SMAD4 score. Following a similar procedure to the INCISE cohort, we utilised the cellular SMAD4 scores to identify the cut-off point. The relationship between SMAD4 expression, cancer-specific survival (CSS) time, and clinicopathological characteristics was examined using Kaplan–Meier (KM) survival curves for the total cohort, as well as in combination with SMAD4 expression and tumour stromal percentage (TSP) scores, Cox regression, and Chi-squared tables, respectively.

2.2.2.6 Transcriptomic data analysis

The full sections of FFPE tissues, with 4 μm in thickness, were cut by GTRF. The epithelium of adenomatous colorectal polyps from the INCISE cohort was annotated by GTRF. Colorectal tumour epithelium from the GRI cohort was annotated by Dr Kathryn Pennel and Dr Phimmada Hatthakarnkul. Subsequently, RNA extraction of the area of interest was performed by BioClavis (UK). The samples were analysed using high-throughput gene expression through Templated Oligo-Sequencing (TempO-Seq™, BioClavis, UK). This assay has been developed to provide whole transcriptomic profiling through a novel ligation-based targeted technology that does not require RNA purification, cDNA synthesis, or targeted RNA capture processes (204). This approach is beneficial for identifying gene expression signatures, such as those found in patient FFPE or TMAs (205).

The raw counts of INCISE cohort data were examined for data quality control by bioinformatician Dr Natalie Fisher. Only 1,765 cases from the training group were eligible. To obtain more precise biological information regarding dysregulation in colorectal polyps and based on prognostic factors, we selected patients with tubulovillous polyps from the first quartile for the low SMAD4 group (n = 93) and the fourth quartile for the high SMAD4 group (n = 94). The R package DESeq2 was used to compare differential gene (DE) expression between the low and high SMAD4 groups, and a Volcano plot and boxplot were created using ggplot2. Gene Set Enrichment Analysis (GSEA) was conducted using GSEA 4.3.3 (Broad Institute, Inc., US).

Another bioinformatician, Leonor Patricia Schubert Santana, evaluated the quality of the raw count data from the GRI cohort and completed the pre-processing. Only 318 cases remained for subsequent analysis. CRC patients with high stroma were selected from the first quartile for the low SMAD4 group (n = 37) and from the fourth quartile for the high SMAD4 group (n = 30). The R package limma was used for differential expression analysis. The Volcano plot and boxplot were created to compare differential gene expression levels between the low and high SMAD4 groups using the EnhancedVolcano package and ggplot2. GSEABase was utilised in the GSEA analysis with R.

2.2.3 *In vitro* 2D cell line studies

2.2.3.1 Adherent cell culture

Colorectal cell lines were provided by Edwards Group, Wolfson Wohl Cancer Research Centre. HepG2 was provided by the Bird Group, Cancer Research UK Scotland Institute. CAL72 and MSC were grown in the Wheadon Group, Paul O'Gorman Leukaemia Research Centre. All cells were stored in liquid nitrogen for long-term preservation.

To grow cell lines in tissue culture, the frozen cells in cryovials were thawed in a water bath for a few mins. The cell suspension was then transferred to a 15 mL Falcon tube containing 10 mL of medium. The cells were centrifuged at 300 x g for 5 mins and subsequently resuspended in 5 mL of fresh medium for a T25 flask or 10 mL for a T75 flask. DLD1, HCT116, and SW620 were cultured in RPMI with 10% FBS and 1% L-glutamine, while SW837, COLO205, and HepG2 were cultured in DMEM with 10% FBS and 1% L-glutamine. CAL-72 was cultured in DMEM with 10% FBS, 1% L-glutamine, 1% Pen/Strep, and 1% Insulin-Transferrin-Selenium. MSC was cultured in Mesenchymal Stem Cell Growth Medium 2 with Supplement Mix and 1% Pen/Strep. All cell lines were maintained in a humidified incubator at 37°C with 5% CO_2, and the medium was changed every 2-3 days by replacing it with 5-6 mL of fresh medium for a T25 flask, 10-12 mL for a T75 flask, and 25-30 mL for a T175 flask. Once the cells reached 70-80% confluence, the cells were split by following these steps: aspirated off the medium, rinsed each flask with at least 5 mL of sterile PBS three times, added trypsin (2 mL for a T25 flask, 3 mL for a T75 flask, and 5 mL for a T175 flask), incubated at 37°C for 3-5 mins until the cells detached, were neutralised using the trypsin with 1-3 mL of FBS, we then added twice the volume of medium to the trypsin volume, transferred the cell suspension to a 15 mL or 50 mL tube, centrifuged at 300 x g for 5 mins, and then resuspended in fresh medium. We determined cell numbers by counting them with a trypan blue assay or by seeding the cells into a new flask using a 1/10 to 3/10 ratio of cell suspension and topping up with fresh medium.

COLO205 is a semi-adherent cell line. We used a cell scraper to remove any adherent cells from the flask surface. The cell suspension was then transferred to

a 15 mL tube and centrifuged at the same speed as outlined above to harvest the cell pellet, which was then resuspended in the fresh medium.

The remaining cell suspension could be used in further experiments or cryopreserved. The 2X freezing media consisted of culture medium solution with 20% FBS and 20% DMSO. Cells were resuspended at 2-4 million cells per mL in complete media and an equal volume of 2X freezing media added dropwise. 1 mL of cells were transferred to a cryovial tube and placed in a Freezing container in a -80°C freezer for rate controlled freezing overnight before transferring to long term storage in a liquid nitrogen tank.

2.2.3.2 Trypan blue exclusion assay for cell viability

To assess cell viability, the cell suspension was diluted 1:1 with trypan blue. 10 µL of the cell dilution were added to a haemocytometer. Cells with clear cytoplasm, which exclude dye infiltration due to an intact membrane, were counted as viable cells, whereas nonviable cells presented blue cytoplasm under a light microscope at 10X magnification (206), Figure 2-4. We calculated the number of cells using the equations below.

$$\text{Number of cells } \left(\frac{\text{cells}}{\text{mL}}\right) = \frac{\text{No. of counted viable or nonviable cells x Dilution factor x } 10^4 \frac{\text{cells}}{\text{mL}}}{\text{No. of counting square}}$$

$$\text{Total Number of cells (cells)} = \text{No. of cells } \left(\frac{\text{cells}}{\text{mL}}\right) \text{x Volume of cell suspension (mL)}$$

$$\text{Percentage of viable cells (\%)} = \frac{\text{No. of viable cells } \left(\frac{\text{cells}}{\text{mL}}\right) \text{ x 100}}{\text{No. of viable and nonviable cells } \left(\frac{\text{cells}}{\text{mL}}\right)}$$

$$\text{Total volume of cell suspension needed (mL)} = \frac{\text{No. of cells needed (cells)}}{\text{No. of having cells } \left(\frac{\text{cells}}{\text{mL}}\right)} + \text{Adding dilution volume(mL)}$$

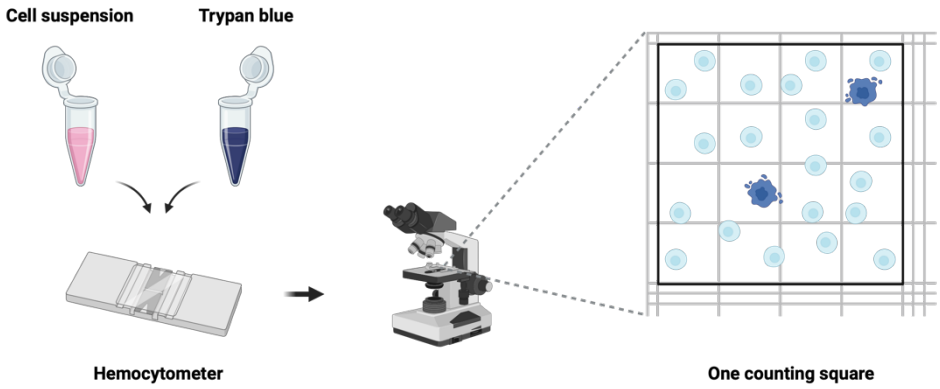


Figure 2-4 Assessment of cell viability using the Trypan blue exclusion assay

For example, we harvested 10 mL of cell suspension, then aliquoted 10 µL of this cell suspension and added it to 10 µL of trypan blue (2X dilution factor). We counted the number of cells from four counting squares. Viable cells and nonviable cells were counted as 200 and 20 cells, respectively. Therefore, the number of viable cells was equal to 100 x 10^4 cells/mL, while nonviable cells were 10 x 10^4 cells/mL. The total number of viable cells in this cell suspension was 10 x 10^6 cells. The percentage of viable cells was 90.91%. If we aim to seed the cells at 2 x 10^6 cells in a T75 flask (10 mL of cell suspension needed), 2 mL of the cell suspension were aliquoted, and 8 mL of medium was added.

2.2.3.3 Drug and compound preparation

In this study, we used the following drugs; Dorsomorphin (DOR), K02288, and LDN-212854, kinase inhibitors targeting the type I receptors of the TGF-ß and BMP signalling pathways (BMPR). In addition to these individual treatments, we evaluated their efficacy in combination with Fluorouracil (5-fluorouracil, 5-FU), a standard chemotherapeutic agent for CRC. Initially, the BMPR inhibitors were reconstituted using dimethyl sulfoxide (DMSO) in a 10 mM stock solution, while 5-FU was prepared in a 200 mM stock solution.

Bortezomib was introduced to examine the effects of 20S proteasome inhibition on protein levels related to the activities of the NF-κB and ERK signalling pathways. This inhibitor was prepared using DMSO in a 10 mM stock solution. The growth factors BMP4 and EGF were dissolved in 7.5% Bovine Albumin Fraction V at concentrations of 10 and 100 µg/mL. The stocks were aliquoted and stored at temperatures between -20°C and -80°C.

The drug treatment was freshly prepared on the days of the experiment. Each drug concentration was diluted in a cell culture medium according to the equations below, and a serial dilution of the drug was also prepared at a 1:1 volume, as shown in Figure 2-5.

$$\text{Volume of drug stock } (\mu L) = \frac{\text{Final concentration } (\mu M) \text{ x Final volume } (mL)}{\text{Concentration of drug stock } (mM)}$$

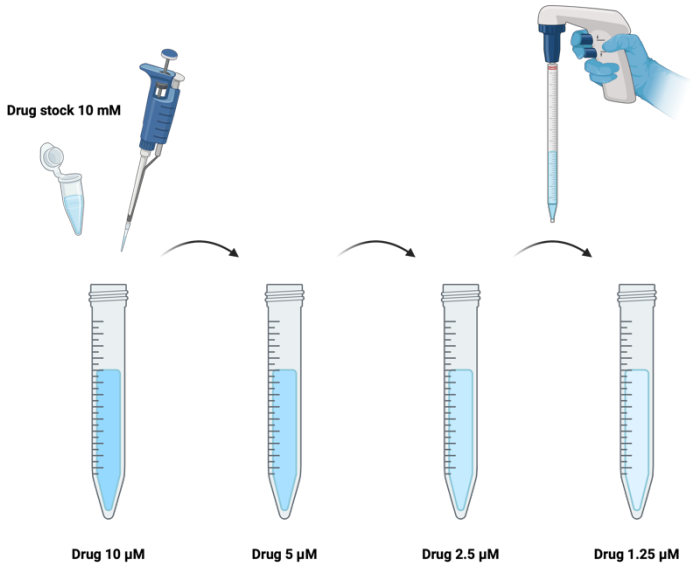


Figure 2-5 Preparation of drug serial dilutions

For example, we aim to prepare 10 µM of LDN-212854 for a final solution of 10 mL, so we added 10 µL of the drug stock to the medium. If we aim to prepare a serial dilution from 10 µM down to 1.25 µM, five mL of the 10 µM drug were added to 5 mL of medium to reach a concentration of 5 µM, and we repeated the same step with the same aliquot volumes until we reached the last concentration.

2.2.3.4 WST-1 Proliferation assay

To investigate the inhibitory concentration of 50% (IC50) of small molecule inhibitors and chemotherapy agents on CRC cell lines, DLD1 and SW620 cells were plated at 5,000 cells in 200 µL of medium per well on a 96-well plate for 24 hours before treatment. A 0.1% DMSO solution was used as vehicle control. Serial dilutions were prepared from drug stocks of BMPR inhibitors: DOR, K02288, and LDN-212854, as well as 5-FU, ranging from 1 to 50 µM and 1 to 200 µM, respectively. In triplicate samples, the medium was replaced with each inhibitor concentration and cultured in a humidified incubator for 72 hours. Representative images were captured using the ZEISS Axiocam 705 mono microscope and ZEN 3.5

blue edition (ZEISS, Germany). Afterwards, 10 µL of Water-Soluble Tetrazolium (WST) -1 reagent was added to each well, and the plate was incubated covered with foil at 37°C for at least 1 hour (ranging from 1 to 3 hours). The plate was examined for metabolic cell proliferation through optical density (OD) measurement at 450 nm using the SpectraMax microplate reader (Molecular Devices, US), Figure 2-6. NAD+ is derived from mitochondrial NADH in living cells, producing dark red formazan crystals from tetrazolium salt, which reflects the number of viable cells (207). The percentage of cell viability under each condition was compared to the vehicle. The half-maximal inhibitory concentration (IC50) was calculated by using GraphPad Prism 10 (GraphPad Software, US).

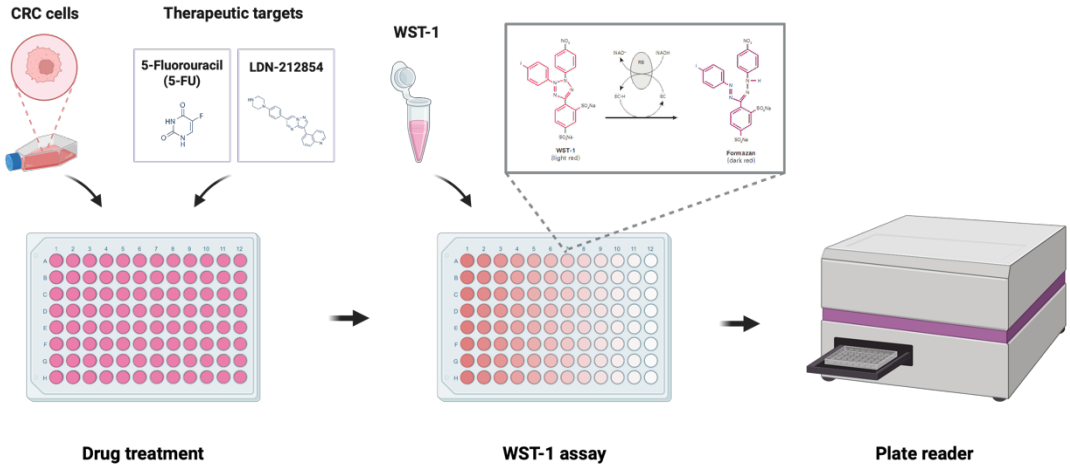


Figure 2-6 Assessment of cell proliferation using the WST-1 assay

CRC cell lines were seeded at the optimal density in 96-well plates and treated with therapeutic agents. Following the specified incubation periods, WST-1 was added, and the optical density (OD) was measured using a plate reader.

2.2.3.5 Chou-Talalay method using CompuSyn for drug combination

Among BMPR inhibitors, we focused on LDN-212854 due to its most potent inhibitory effect on cell proliferation, as demonstrated by the WST-1 assay. To study the drug in combination with 5-FU, DLD1 and SW620 cells were seeded at 5,000 cells per well on a 96-well plate. The cells were treated with 0.5X, 1X, and 2X of the IC50 of the first drug alone and 1X of the IC50 of the first drug combined with 0.5X, 1X, and 2X of the IC50 of the second drug (Figure 2-7). The plates were incubated for 48 and 72 hours, and then cell viability was examined using the

WST-1 assay. The combination effects were investigated using the Chou-Talalay method based on the Median-Effect Equation (208, 209), as shown below, utilising CompuSyn version 1.0 (ComboSyn Incorporated, US).

$$\frac{fa}{fu} = \left(\frac{D}{Dm}\right)^{m}$$

*f*a is the fraction affected or inhibition fraction in the range of 0 to 1. It is calculated from % inhibition/100

$$\%\ \text{inhibition} = \left[1 - \left(\frac{\text{OD treatment} - \text{OD media blank}}{\text{OD vehicle} - \text{OD media blank}}\right)\right] \text{x } 100$$

*f*u is fraction unaffected (*f*u = 1- *f*a)

D is the dose of the drug

Dm is the dose that is required to produce a median effect e.g. IC50

M is the coefficient signifying the shape of the dose-effect curve: flat sigmoidal (m < 1), hyperbolic (m = 1), and sigmoidal (m > 1)

The combination Index (CI) was implemented from the programme, which indicated the drug combination results: synergism (CI < 1), additivity (CI = 1), and antagonism (CI > 1)

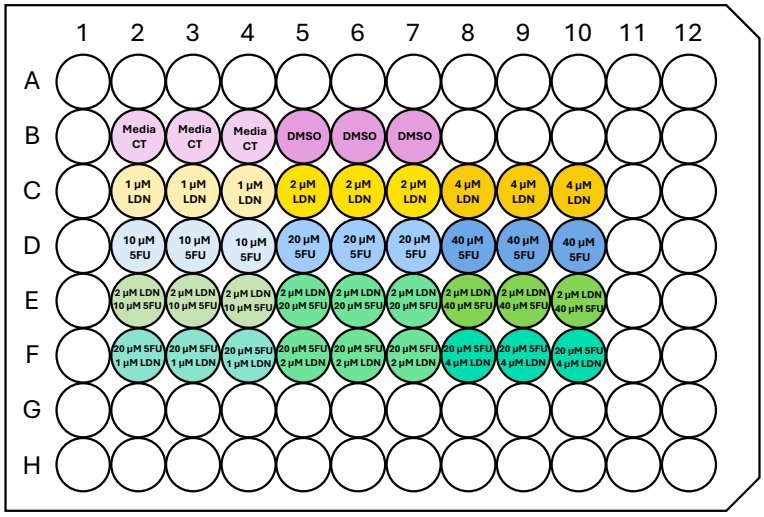


Figure 2-7 Conditions for the drug combination study applying the Chou-Talalay method

For example, CRC cell lines were seeded at the optimal density in 96-well plates and treated with 1 to 4 μM of LDN-212854 and 10 to 40 μM of 5-FU for single-drug treatment. Additionally, the combinations of 2 μM of LDN-212854 with 10 to 40 μM of 5-FU and 20 μM of 5-FU with 1 to 4 μM of LDN-212854 were performed. Following the WST-1 analysis, the combined effects were evaluated using the Chou-Talalay method.

2.2.3.6 Clonogenic assay

The clonogenic assay, often known as the colony formation assay, studies a single cell's ability to form a colony *in vitro*. This survival assay is commonly used to evaluate cell reproductive death and the efficacy of treatments following radiation and chemotherapy. The optimal dilution of a cell fraction, before or after treatment, is seeded to produce colonies (containing at least 50 cells per colony) within 1-3 weeks. Low cell densities at seeding may lead to inadequate colony formation. Thus, alternative methods, such as the addition of a conditioned medium (the medium collected from a growing culture that contains growth factors produced by dividing cells) and the growth of the cells on soft agar with or without a fibroblast feeder layer, could facilitate colony formation (210).

To begin, seeding density and conditions were optimised for the cell properties of SW620, a metastatic cell with a slightly smaller size and moderate adherence to the culture plate compared to DLD1. Before seeding SW620, 6-well plates were coated with 1 mL of 1% Basement Membrane Extract (BME) Type R1 in cold serum-free medium per well and incubated at 37°C overnight. Subsequently, cell suspensions were prepared with 500, 750, and 1,000 cells per well. Prior to seeding any excess BME coating solution was removed from the plates. The culture plates, both with and without BME coating, were then seeded with 4 mL of the cell suspension. After 24 hours of incubation, the cells were treated for the next 96 hours with 0.1% DMSO diluted in normal medium or 10% conditioned medium (collected from SW620 normal culture for 2-3 days) as vehicle controls and compared to controls without DMSO (normal medium culture and 10% conditioned medium). Following 96 hours treatment with DMSO the plates were then replaced with normal medium or 10% conditioned medium and refreshed every 2-3 days until reaching 2 weeks in culture. The colonies were fixed with 2 mL of 50% methanol in PBS and 100% methanol at room temperature for 15 mins each, stained with 0.25% v/v crystal violet solution in PBS for 45 mins, and the plates were turned upside down to dry overnight. The plates were scanned using the GelCount™ Cell Colony Counter (Oxford Optronix, UK), Figure 2-8.

After we achieved an optimal seeding condition, the cells were seeded and treated with 0.25X, 0.5X, and 1X of the drug's IC50, both alone and in combination, for 96 hours in triplicate. The number of colonies was counted manually, and the percentage of colony formation under each condition was then compared to the vehicle.

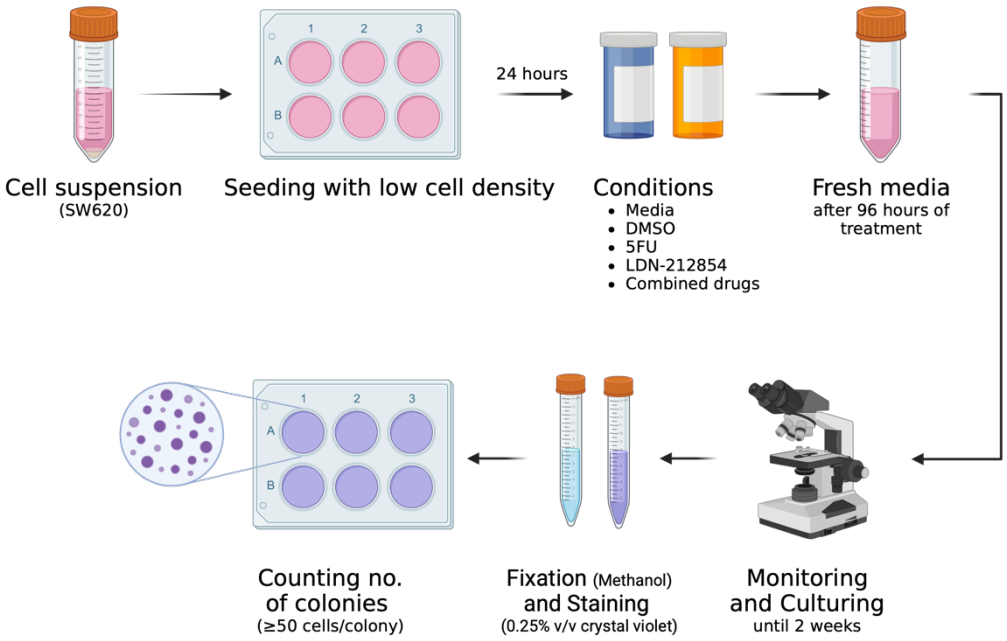


Figure 2-8 Assessment of cell proliferation and recovery through the clonogenic assay

CRC cell lines were seeded at low density in 6-well plates and incubated with drug treatment. After 96 hours, fresh medium was added, and the culture was maintained for two weeks. Finally, the colonies were fixed and stained with a 0.25% v/v crystal violet solution to visualise and count them.

2.2.3.7 Growth factor stimulation *in vitro*

To investigate whether the crosstalk pathways of TGF-β and BMP signalling, Wnt-β-catenin signalling, and MAPK/ERK signalling regulate the loss of SMAD4 function in CRC, DLD1 and SW620 were seeded at 2 to 3 x 10^6 cells in T75 flasks and incubated at 37°C for 24 hours or until reaching 80% confluence. Next, the cells were rinsed three times and treated with 1X HBSS, devoid of calcium and magnesium (diluted 10X HBSS in dH_2O, pH 7.2) for one hour to reduce signals from the normal medium. The growth factors, BMP4 and EGF, were diluted to 20 ng/mL in serum-free medium, and 10 mL of this solution was added to the cells and incubated for four hours. The cells were harvested by trypsinisation and using

1X HBSS buffer instead of 1X PBS. Subsequently, protein lysis and western blotting were performed for further investigation.

2.2.3.8 Establishing fluorescent cell lines

Fluorescent cells using labelled fluorescent proteins have been widely produced to assist in cell tracking. One method is lentiviral transduction, which provides efficient transgene delivery to mammalian cells (211). In this study, a stock of viral suspension containing EGFP (19319, addgene, US) and BFP (163169, addgene, US) was produced by Dr Caroline Busch and Xenia Bubnova. DLD1 and SW620 were seeded at 5 x 10^5 cells per well into 6-well plates 24 hours prior to lentiviral transduction. The medium was replaced with RPMI containing 20% FBS, 1% L-glutamine, and 16 µg/mL Polybrene (PB) and combined with the virus suspension containing EGFP or BFP at a 1:1 concentration. The plates were centrifuged at 900 x g for 30 mins at room temperature with slow deceleration and then incubated at 37°C overnight. Next, the cells were recovered with fresh normal medium for 24 hours and selected for transfected cells using 2 µg/mL Puromycin. Subsequently, the confluent transfected cells were sorted for high fluorescent expressing cells using a BD FACSARIA cell sorter (BD Biosciences, UK), and their fluorescence was detected after expansion by BD FACS Canto II (BD Biosciences, UK), utilising channels at 488-530/30 for EGFP and 405-450/50 for BFP. Representative images were captured using the ZEISS Axiocam 705 mono microscope and ZEN 3.5 blue edition, utilising the bright field, EGFP channel, and Syt40 or DAPI channels for BFP, Figure 2-9.

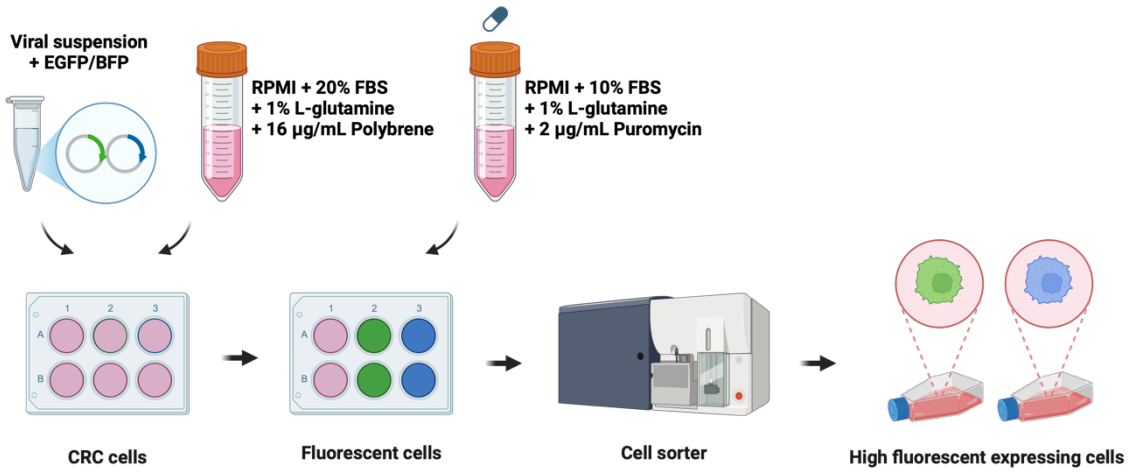


Figure 2-9 Establish fluorescent CRC cell lines expressing EGFP and BFP

CRC cell lines were seeded at the optimal density in 6-well plates and incubated with a 1:1 ratio of viral supernatant containing EGFP or BFP mixed with transfection medium. Subsequently, the fluorescent cells were selected by culturing them in a medium containing puromycin, and the cells were sorted for high fluorescent expression.

2.2.3.9 Conditioned medium from different cell types

The conditioned medium consists of secreted proteins produced by dividing cells. Enzymes, growth factors, cytokines, hormones, and other soluble mediators are released to regulate interactions between cells and the extracellular matrix, facilitating cell growth, differentiation, invasion, and angiogenesis (212). To investigate the effects of conditioned media produced by different cell types on CRC proliferation, the conditioned media were collected from the growing cultures of bone marrow cell lines (CAL72 and MSC), a hepatic cell line (HepG2), with CRC cell lines (DLD1 and SW620) conditioned media used as controls after incubation for 2-3 days. The cultured media were filtered using a 0.45 µm filter, aliquoted and then stored at -20°C. Subsequently, DLD1 and SW620 were seeded at 5,000 cells in 200 µL of normal medium per well on a 96-well plate for 24 hours. The conditioned media were thawed at 37°C and prepared in a serial dilution ranging from 1.56% to 100% of conditioned media. Complete normal medium or 0% of conditioned media served as a positive control. Cell-free wells filled with the individual conditioned media were designated as negative controls for each concentration. The normal medium was replaced with conditioned media and incubated for 72 hours before assessing CRC cell proliferation using the WST-1 assay. Following this protocol, we prepared conditioned media at concentrations of 10%, 20%, and 30% in normal medium to demonstrate a clearer growth trend, Figure 2-10. Furthermore, we examined the inhibitory effects of single drug treatments at 0.5X, 1X, and 2X of the IC50, diluted in 20% conditioned medium for 72 hours. The treatments prepared with normal medium served as a reference, The percentage of cell proliferation under each conditioned medium was compared to the normal medium, as calculated from the equation below.

$$\%\ \text{cell proliferation} = \frac{\text{OD conditioned medium} - \text{OD cell} - \text{free conditioned medium}}{\text{OD normal medium} - \text{OD cell} - \text{free normal medium}} \text{ x } 100$$

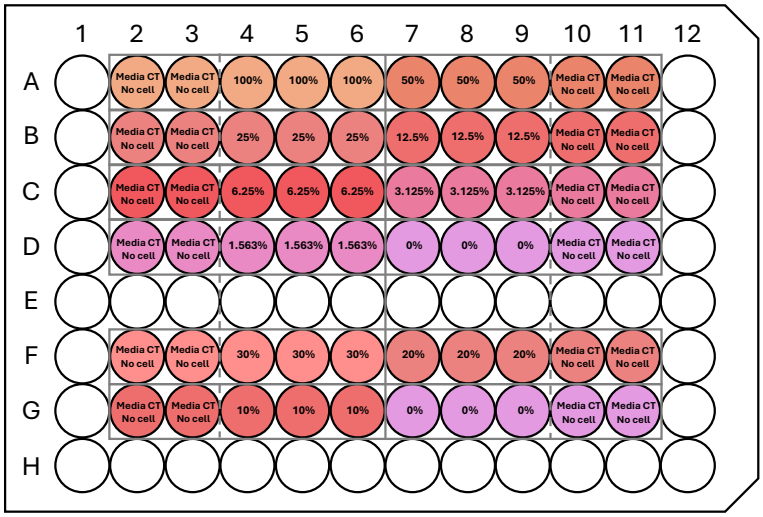


Figure 2-10 Conditions for studying conditioned media from different cell types

CRC cell lines were seeded at the optimal density in 96-well plates and treated with various conditioned media from other cell types, ranging from 0% to 100% or 0% to 30%.

2.2.4 Flow cytometry (FACS)

2.2.4.1 Annexin V apoptosis analysis

Cell death results in changes to the plasma membrane, including the translocation of phosphatidylserine (PS), a negatively charged phospholipid, from the inner side to the outer cell surface. In the early stages of apoptosis, the cell membrane remains intact; however, during necrosis, the membrane loses its integrity and starts to leak. Annexin V has a high affinity for binding to PS in a Ca^{2+}-dependent manner. Propidium iodide (PI), 4′,6-diamidino-2-phenylindole (DAPI) and 7-Aminoactinomycin D (7-AAD) are non-vital exclusion dyes that bind strongly to A-T-rich regions in DNA (PI and DAPI) or double stranded DNA by intercalating between base pairs in G-C rich regions (7-AAD). Therefore, using fluorescence-activated cell sorting (FACS) via flow cytometry to detect fluorescein isothiocyanate (FITC)-labelled Annexin V (green fluorescence) combined with either PI, 7-AAD (red fluorescence) or DAPI (blue fluorescence) staining allows for the discrimination of cell populations into intact live cells ($FITC^-$and PI^-/ $DAPI^-$/ $7\text{-}AAD^-$), apoptotic cells ($FITC^+$ and PI^- /$DAPI^-$/$7\text{-}AAD^-$), and necrotic cells ($FITC^-$ and PI^+/$DAPI^+$/$7\text{-}AAD^+$) (213, 214).

To investigate the inhibitory effect of drug treatments, inducing apoptosis in metastatic cells, SW620 cells were seeded at a density of 3 x 10^6 cells in T75 flasks

for one day before being treated with 20 µM of 5-FU, 2 µM of LDN-212854, and a combination of the two drugs (20 µM of 5-FU and 2 µM of LDN-212854) for 48 and 72 hours. A 0.1% DMSO solution was utilised as the vehicle control. Floating and adherent cells were collected using trypsinisation and centrifuged at 300 x g for 5 mins. A total of 5 x 10^5 cells were aliquoted into FACS tubes, washed with 1X HBSS containing calcium and magnesium (10X HBSS diluted in dH_2O), and stained with 1X FITC Annexin V and 1X 7-AAD (2.5 µL of each fluorochrome in a total of 100 µL of 1X HBSS) for 20 mins at room temperature in the dark. The cell mixture was also aliquoted for unstained and staining with a single fluorochrome of 1X FITC Annexin V and 1X 7-AAD, as referenced in the FACS control setting. To terminate the reaction, 100-200 µL of 1X HBSS was added to the FACS tubes. Flow cytometry model BD FACSCanto™ II was used to detect cell populations using channels at 488-538/30 for FITC Annexin V and 488-670LP for 7-AAD, as shown in Table 2-23. The results were analysed by FlowJo™ v10 Software (BD Biosciences, UK), as illustrated in Figure 2-11.

Table 2-23 FACS parameters for Annexin V apoptosis analysis

Parameter	Voltage	Log	Area (A)	Height (H)	Width (W)
Forward scatter (FSC)	240	-	✓	✓	✓
Side scatter (SSC)	370	-	✓	✓	✓
488-530/30 (FITC Annexin V)	320	✓	-	-	-
488-670LP (7-AAD)	430	✓	-	-	-

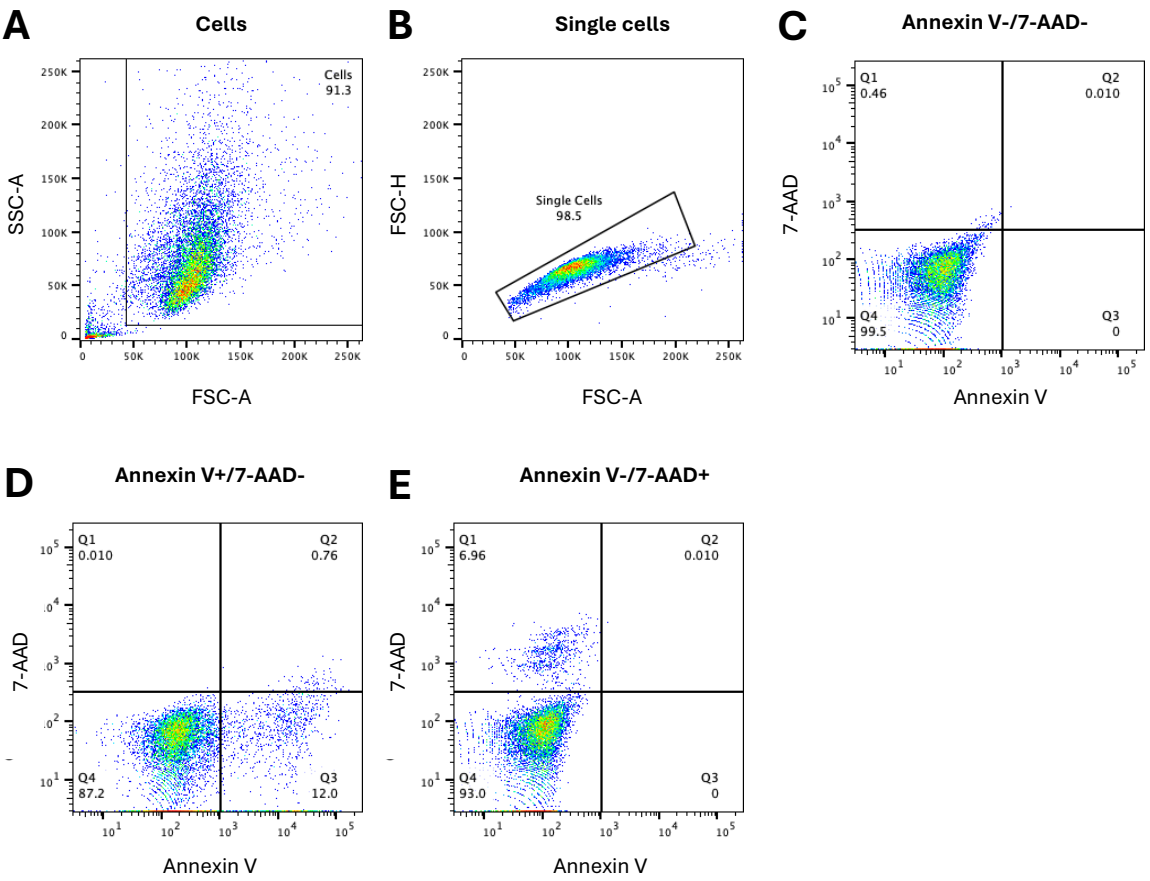


Figure 2-11 Gating strategy for Annexin V apoptosis analysis

The FACS results were defined for the cell population area, excluding debris (A). Single cells were selected (B). The unstained sample was gated as AnnexinV negative and 7-AAD negative (C). The sample stained only with AnnexinV was identified as AnnexinV positive and 7-AAD negative (D). The sample stained only with 7-AAD was identified as 7-AAD positive and AnnexinV negative (E).

Additionally, we set up a parallel experiment to examine the amount of cell death using the Trypan blue exclusion assay. SW620 cells were seeded at a density of 2 x 10^5 cells in 2 mL medium per well on 6-well plates and then treated with the same drug concentrations for 48, 72, and 96 hours.

2.2.4.2 Ki-67 cell cycle analysis

Among cell populations, various cell cycle statuses exist, including proliferating cells that progress through the DNA synthesis phases of G1, S, G2, and M, as well as resting or quiescent cells (G0). Distinct cellular DNA content is observed in the G0/G1 (2n), S (2n to 4n), and G2/M (4n) phases, which are evaluated using fluorescent DNA dyes, such as PI and DAPI, via flow cytometry. To distinguish G0 cells from G1 cells, a proliferation-specific marker, Ki-67, is used to identify

nuclear proliferation antigens. Higher Ki-67 expression is found during cell proliferation, reaching a maximum level in the G2 and early M phases. Extremely low or undetectable levels are presented in the G0 phase and during the anaphase and telophase of the mitotic processes (215).

In a tissue culture method similar to apoptosis analysis, SW620 were seeded at 3 x 10^6 cells in T75 flasks and then treated with the drug conditions: 20 µM of 5-FU, 2 µM of LDN-212854, and the drug combination for 48 hours to observe the inhibitory effects on the cell cycle, as the doubling time of SW620 is approximately 30 hours. The cells were harvested by trypsinisation and centrifugation at 300 x g for 5 mins. Next, 2 x 10^5 cells were fixed with 250 µl of BD Cytofix/Cytoperm™ Fixation and Permeabilization Solution at room temperature for 20 mins in dark. After centrifugation, the fixed cells were washed twice with 500 µL of 1X wash/perm solution. The fixed cells in 1X wash/perm solution were stored at 4°C overnight. Subsequently, 50 µL of Ki-67 PECy7 at a concentration of 1:100 in 1X wash/perm solution was added to the cells and incubated at 4°C for 30 mins before staining with 100 µL of DAPI at a 1:1000 concentration in 1X wash/perm solution at room temperature for 10 mins in the dark. Unstained and staining with a single fluorochrome were prepared to establish FACS reference controls. The cells were detected by flow cytometry (BD FACSCanto™ II) using channels at 488-780/60 for Ki-67 PECy7 and 405-450/50 for DAPI, as shown in Table 2-24. The results were analysed using FlowJo™ v10 Software, as illustrated in Figure 2-12.

Table 2-24 FACS parameters for Ki-67 cell cycle analysis

Parameter	Voltage	Log	Area (A)	Height (H)	Width (W)
Forward scatter (FSC)	200	-	✓	✓	✓
Side scatter (SSC)	400	-	✓	✓	✓
488-780/60 (Ki-67 PECy7)	620	✓	✓	-	-
405-450/50 (DAPI)	315	-	✓	-	-

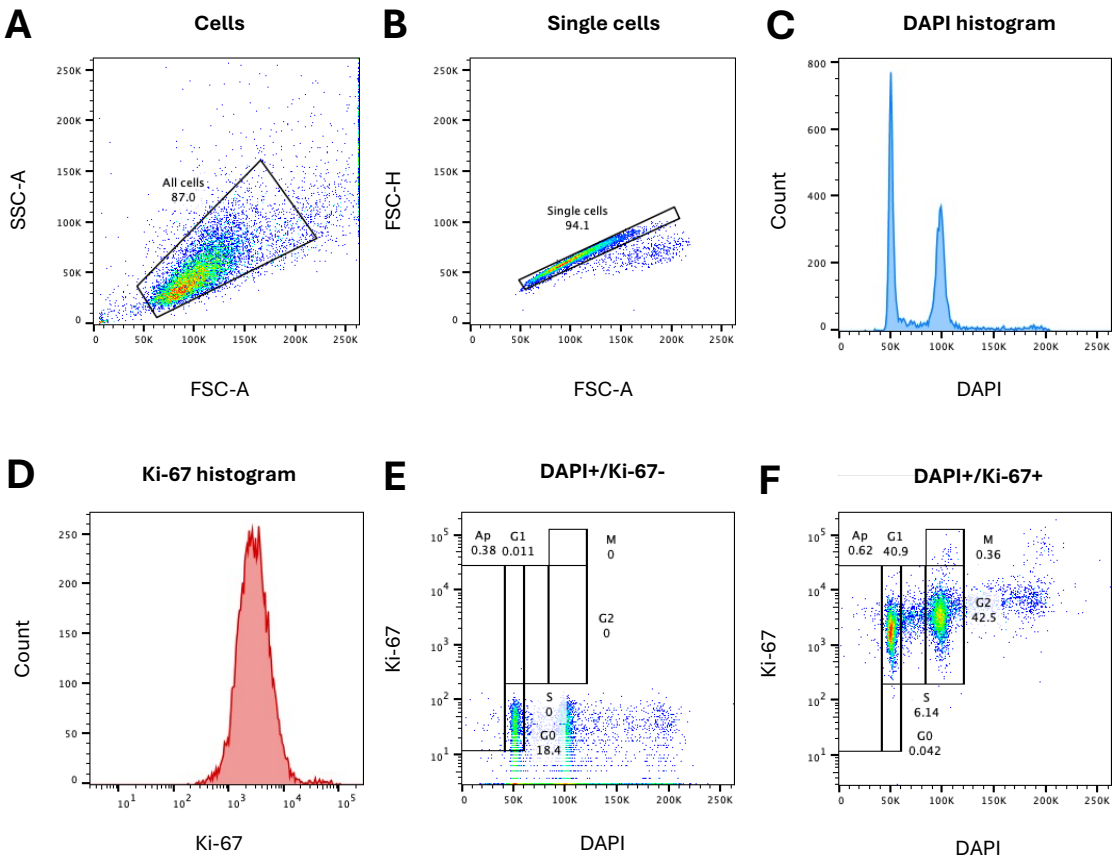


Figure 2-12 Gating strategy for Ki-67 cell cycle analysis

The FACS results were defined for the cell population area, excluding debris (A). Single cells were selected (B). The DAPI histogram displayed two peaks of cells, identified as 2n and 4n, at 50K and 100K (C). The Ki-67 histogram revealed a peak of cells expressing Ki-67, detected between 10^3 and 10^4 (D). The sample stained exclusively with DAPI was identified as DAPI positive and Ki-67 negative (E). The G2 phase was gated as double the size of the G1 phase, with both phases being Ki-67 positive (F).

2.2.4.3 CellTrace™ Violet (CTV) analysis

CellTrace™ Violet (CTV) is a succinimidyl ester that covalently binds to cellular proteins through interaction with all free intracellular amine groups. This dye is widely used to assess cell proliferation by fluorescently labelling cellular proteins. Once cell division occurs, intracellular esterases cleave the CTV, distributing the fluorescence intensity (MFI) equally from parent cells to daughter cells. Thus, the number of cell divisions can be quantified (216). Colcemid is employed to terminate cell division by preventing microtubule polymerisation, leading to cell-cycle arrest (217).

To evaluate the cell divisions resulting from drug treatments on SW620 cells, the cells were seeded in T25 flasks, at 1 x 10^6 cells overnight. The adherent cells were rinsed once with 1X PBS before being treated with 5 µM of CTV in 1X PBS with 2% FBS (5 mL of the CTV solution per flask) and incubated at 37°C for 20 mins. Next, the CTV-stained cells were washed twice and incubated with fresh medium for 10 mins to terminate the reaction. Subsequently, the medium was replaced with 100 ng/mL of Colcemid in culture medium as a CTV max control in one flask or with the following drug treatments: 20 µM of 5-FU, 2 µM of LDN-212854, and the drug combination. The cells were harvested by trypsinisation after incubating for 96 hours, which encompasses three to four cell divisions. Afterwards, 2 x 10^5 cells were stained with 100 µL of 1X 7-AAD at room temperature for 10 mins in the dark to exclude non-viable cells. Cell aliquots from the unstained control (non-stained CTV and 7-AAD), stained 7-AAD on non-viable cells (incubated unstained cells in 70% ethanol on ice for 30 mins), stained CTV max control, and stained CTV max and 7-AAD were prepared to establish FACS reference controls. The cells were detected using BD FACSCanto™ II with channels set at 405-450/50 for CTV and 488-670LP for 7-AAD, as shown in Table 2-25. The results were analysed using proliferation modelling on FlowJo™ v10 Software, as illustrated in Figure 2-13.

Table 2-25 FACS parameters for CTV analysis

Parameter	Voltage	Log	Area (A)	Height (H)	Width (W)
Forward scatter (FSC)	240	-	✓	✓	✓
Side scatter (SSC)	370	-	✓	✓	✓
405-450/50 (CTV)	300	✓	✓	✓	✓
488-670LP (7-AAD)	500	✓	✓	✓	✓

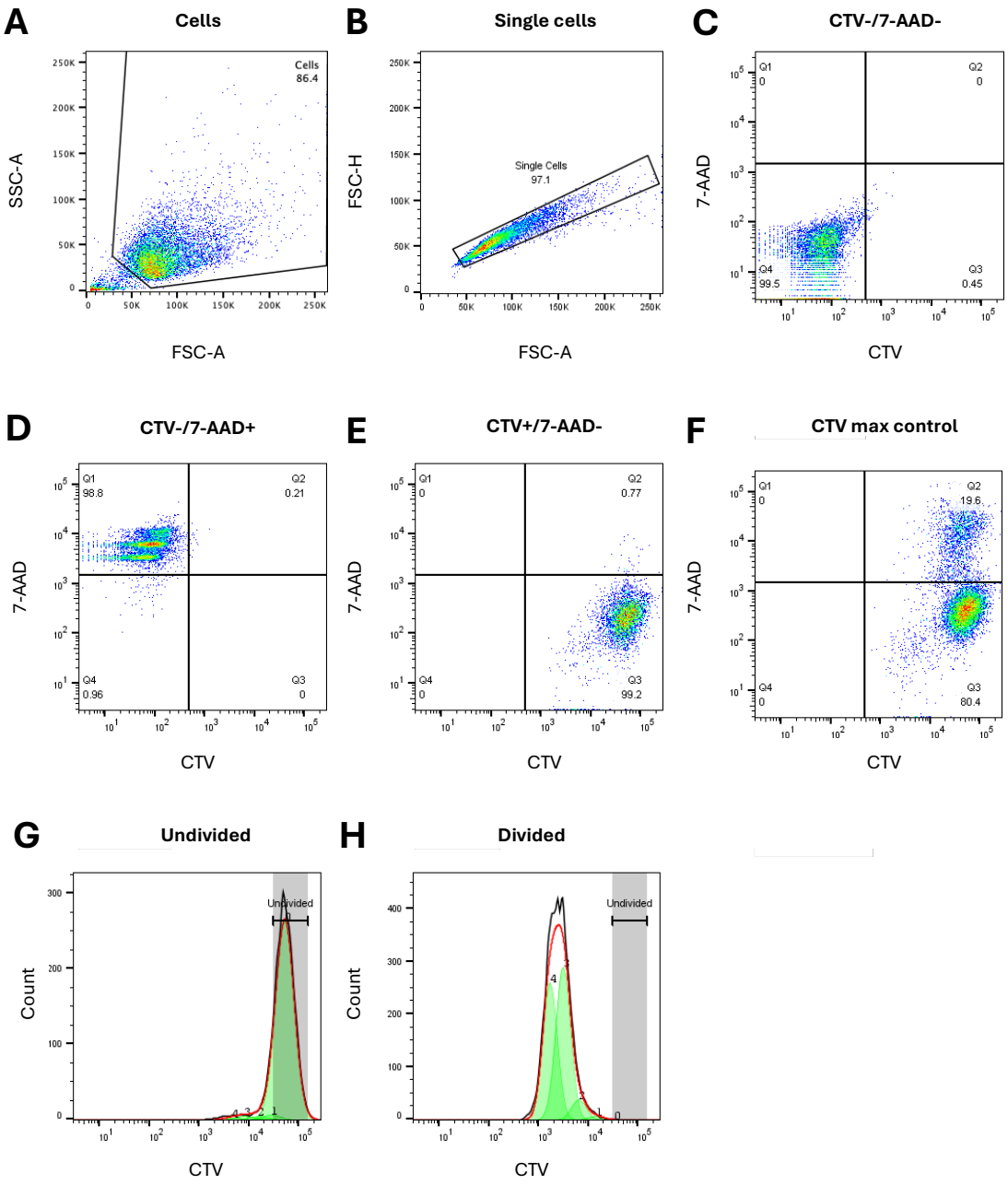


Figure 2-13 Gating strategy for CTV analysis

The FACS results were defined for the cell population area, excluding debris (A). Single cells were selected (B). The unstained sample was categorised as CTV negative and 7-AAD negative (C). The sample stained only with 7-AAD was identified as 7-AAD positive and CTV negative, indicating the presence of dead cells (D). The CTV max control without 7-AAD staining was identified as CTV positive and 7-AAD negative (E). The CTV max control with 7-AAD staining was analysed (F). Q3 cells were selected as viable cells, with the CTV histogram of the CTV max sample determined to represent undivided control cells (G). The CTV histogram of the vehicle was identified as representing divided cells (H).

2.2.4.4 Cell sorting

The transfected DLD1 and SW620 cells containing EGFP and BFP were harvested using trypsinisation and diluted in 10 mL of 1X PBS. The cells were filtered through a 70 µm cell strainer. One million cells were resuspended in 1 mL of 1X PBS in a sterile 5 mL round-bottom tube. With the assistance of Jennifer Cassels, a lab technician, the transfected cells were sorted for high fluorescence expression using a BD FACSARIA™ cell sorter, utilising the FITC channel for EGFP and the BV421 channel for BFP. DLD1 and SW620 cells from normal cultures served as a reference control. The high fluorescing EGFP and BFP CRC cells were collected and subsequently seeded in T25 flasks. After expansion, the fluorescence in the cells was detected using a BD FACS Canto II, with channels set at 488-530/30 for EGFP and 405-450/50 for BFP. The results were analysed using proliferation modelling on FlowJo™ v10 Software, as shown in Figure 2-14.

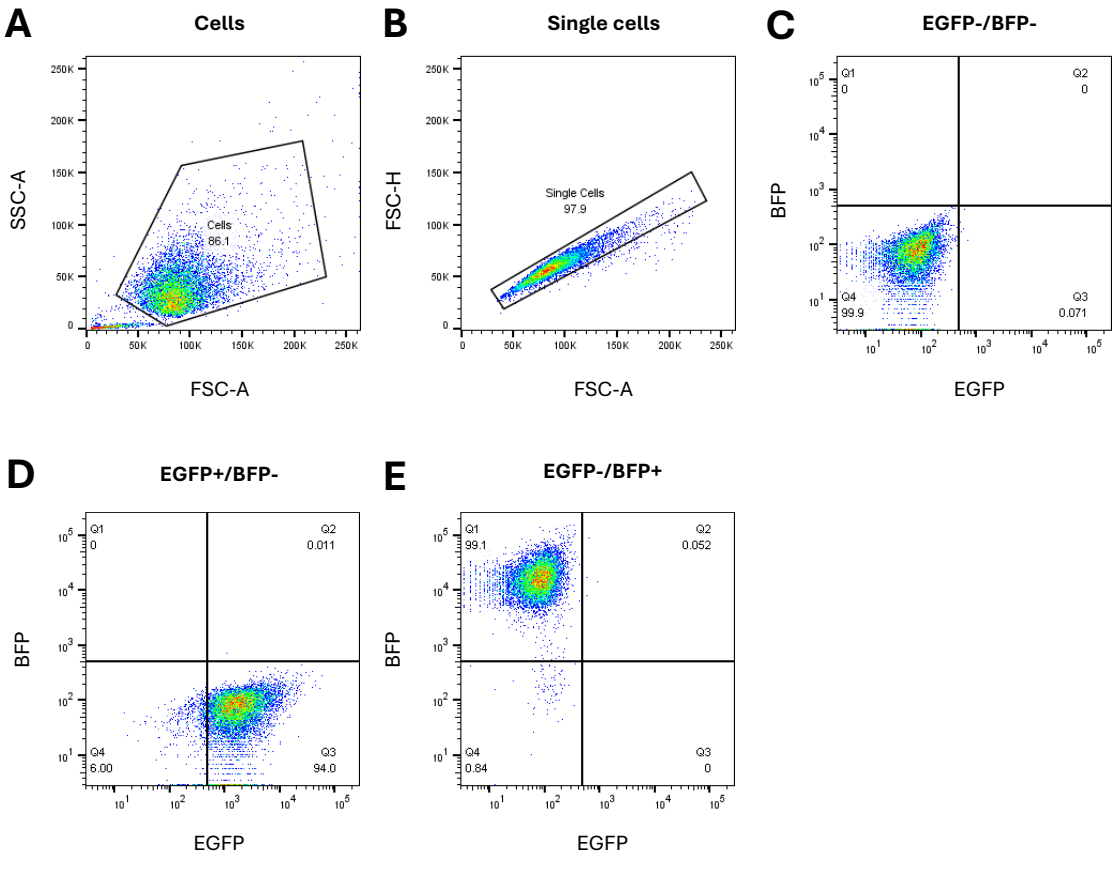


Figure 2-14 Gating strategy for sorting EGFP and BFP cells

The FACs results were defined for the cell population area, excluding debris (A). Single cells were selected (B). The non-transfected cells were gated as both EGFP negative and BFP negative (C).

The EGFP cells were detected as EGFP positive and BFP negative (D). The BFP cells were detected as BFP positive and EGFP negative (E).

2.2.5 Molecular biology - RNA expression analysis

2.2.5.1 RNA extraction

Five CRC cell lines (DLD1, HCT116, SW620, SW837, and COLO205) in normal cultures were collected at 5 x 10^5 cells per 1.5 mL tube. RNA extraction was performed using the Qiagen RNeasy Plus Micro kit according to the manufacturer's instructions. Briefly, the cell pellet was resuspended in 350 µL of Buffer RLT and homogenised by vortexing. Next, the sample was lysed with 350 µL of 70% ethanol before being transferred to the spin column and centrifuged at 12,000 x g for 1 min. Then, 700 µL of Buffer RW1 and 500 µL of Buffer RPE were added to the tube, followed by centrifugation at the same speed before adding 500 µL of 80% ethanol, and then centrifuging at 12,000 x g for 2 mins. The spin column lid was opened, and it was spun at full speed for 5 mins to dry the membrane. The column was placed in a 1.5 mL tube, 14 µL of RNase-free water was added, and it was centrifuged at full speed for 1 min to elute the RNA. Subsequently, 1 µL of RNA was quantified using the Nanodrop Spectrophotometer ND-1000 (Thermo Fisher Scientific Inc., US). RNA concentration (ng/µL) and purity at the ratios A260/A280 and A260/230 were determined. All RNA samples were stored at -80°C.

2.2.5.2 cDNA synthesis

The High-Capacity cDNA Reverse Transcription Kit was used to convert RNA into cDNA via reverse transcription. 2 µg of RNA, diluted in 13.2 µL of RNase-free water, were mixed with 6.8 µL of a mastermix solution that included 2 µL of Random primer, 2 µL of 10X RT buffer, 0.8 µL of 100 mM dNTPs, 1 µL of RNase inhibitor (20 U/µL), and 1 µL of RT enzyme, resulting in a total volume of 20 µL. Next, the reverse transcription was conducted using the ProFlex PCR System (Applied Biosystems, UK), with the reaction conditions set at 25°C for 10 mins, 37°C for 120 mins, and 85°C for 5 mins. An additional 60 µL of RNase-free water was used to dilute the cDNA to a concentration of 25 ng/µL. All cDNA samples were stored at -20°C.

2.2.5.3 Quantitative Real-time Polymerase Chain Reaction (qPCR)

The SYBR Green mix was prepared at 5.4 µL per sample, consisting of 5 µL of PowerTrack™ SYBR Green and 0.4 µL of 10 µM paired primers, as detailed in Table 2-12. The DNA template mix was created as a 4.6 µL solution containing 0.4 µL of cDNA (10 ng cDNA), 0.25 µL of 40X sample buffer, and 3.95 µL of RNase-free water. Subsequently, the SYBR Green mix and the DNA template mix were combined in a 384-well plate, with 10 µL per well. The plate was sealed and centrifuged at 400 x g for 2 mins. Then, cDNA amplification was detected using the QuantStudio™ 7 Pro Real-Time PCR System (Thermo Fisher Scientific Inc., US), which included one cycle of uracil-DNA glycosylase (UDG) activation at 50°C for 2 mins, one cycle of initialisation at 95°C for 2 mins, and 40 cycles of denaturation at 95°C for 15 secs, followed by annealing/extension at 60°C for 1 min. During amplification, SYBR Green binds to double-stranded DNA and emits fluorescence. An increased proportion of fluorescence correlates with the quantity of DNA. The cycle number that reaches the fluorescence threshold corresponds to Ct values. The results were calculated from the equations below.

$$\Delta\mathrm{Ct} = \mathrm{Ct\ of\ gene\ of\ interest} - \mathrm{Ct\ of\ house\ keeping\ gene}$$

$$\Delta\Delta\mathrm{Ct} = \Delta\mathrm{Ct\ of\ treatment} - \Delta\mathrm{Ct\ of\ control}$$

$$\mathrm{n-fold\ enrichment} = 2^{-\Delta\Delta\mathrm{Ct}}$$

2.2.6 Western blotting - Protein expression analysis

2.2.6.1 Protein lysis

We harvested cell pellets from five CRC cell lines in standard cultures (DLD1, HCT116, SW620, SW837, and COLO205), as well as under drug treatment and stimulation conditions. DLD1 and SW620 were treated with 20 µM of 5-FU, 2-3 µM of LDN-212854, and a drug combination for 4, 24, and 48 hours. Additionally, they were stimulated with 20 ng/mL of BMP4 and EGF for 4 hours and treated with 10 and 30 µM of Bortezomib for 24 hours. After trypsinisation, the cells were washed twice with 1X PBS or 1X HBSS, and 200 µL of protein lysis buffer containing protease and phosphatase inhibitors was added to at least 4 million cells. The pellets were vortexed thoroughly before and after being incubated on ice for

30 mins, followed by centrifugation at 16,000 x g for 20 mins. Protein lysate was collected from the supernatant into new 1.5 mL tubes and stored at -80°C.

2.2.6.2 Protein quantification

Protein quantification was performed using a Bio-Rad Bradford assay. Briefly, the BSA standard was diluted to 1 mg/mL (1:1 dilution in ddH_2O) and subsequently mixed with 1 mL of Quick Start™ Bradford 1X Dye Reagent, utilising the specified volumes from Table 2-26 to establish a range of 0-20 µg/µL protein standards (with colours ranging from brown to blue). 2 µL of protein lysate were added to a 1 mL aliquot of the 1X dye, ensuring they were within the range of the protein standard. 100 µL of protein standards and samples were loaded into a flat-bottom 96-well microplate in duplicate, as shown in Figure 2-15. The plate was examined for protein quantification by measuring the optical density (OD) at 595 nm using the SpectraMax microplate reader with Softmax Pro software (Molecular Devices, US). The amounts of protein lysate were calculated from the mean concentration values (µg/µL). The protein lysates were typically prepared for loading at a concentration of 20 µg/10 µL, totalling 100 µg in 50 µL (by adding 10 µL of 5X working sample buffer) for one aliquot. The stock of protein-loading samples was stored at -20°C.

Table 2-26 BSA standard dilution

Tube	BSA required concentration (µg/µL)	BSA required volume (µL)
1	0	0
2	2.5	2.5
3	5	5
4	7.5	7.5
5	10	10
6	15	15
7	20	20

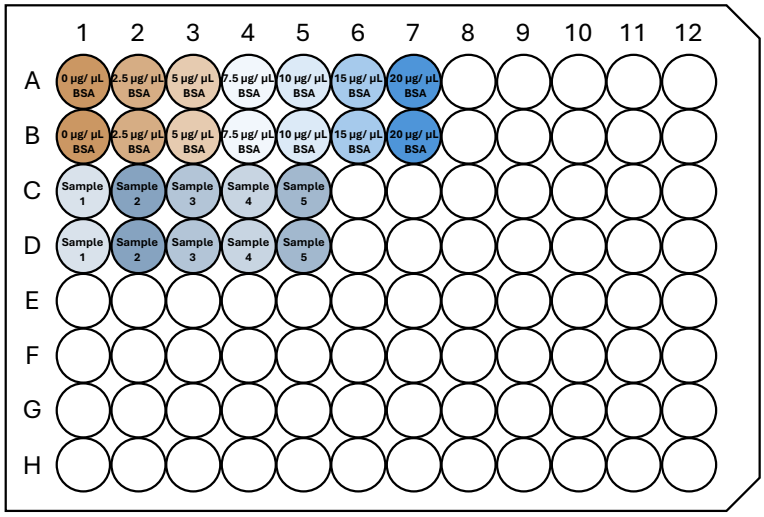


Figure 2-15 Preparation of protein standards and samples for protein quantification

The BSA standard was diluted to a concentration range of 0 to 20 µg/µL, and the protein samples were also diluted with Bradford reagent. Duplicate aliquots of 100 µL of protein standards and samples in a 96-well plate were analysed using a plate reader.

2.2.6.3 SDS-PAGE Western Blotting

Hand-casting 0.75 mm acrylamide gels involved preparing the running gel at a specific percentage between 7.5 and 12, based on the target protein sizes. This was combined with a 5% stacking gel, as detailed in Table 2-17. Next, a pair of gels was arranged in the running tank (Bio-Rad, US) and filled with 1X SDS running buffer. The protein loading samples were thawed on ice and boiled at 95°C for 5 mins. Each gel lane was filled with either 5 µL of protein ladder, 10 µL of protein loading samples (20-30 µg), or 10 µL of 1X sample buffer for any empty lanes. The gels were run at 80V for 20 mins and then at 180V for 30-40 mins using the PowerPac™ Basic Power Supply (Bio-Rad, US). Subsequently, proteins were transferred from the gel to a nitrocellulose membrane using either semi-dry or wet transferring techniques. For semi-dry transferring, four pieces of Whatman filter paper, one piece of nitrocellulose membrane, the gel, and four additional pieces of Whatman filter paper were soaked in 1X transfer buffer and layered on the electrode plate of the Trans-Blot® SD Semi-Dry Transfer Cell according to the order shown in Figure 2-16A. The electric current was set to a constant amperage (0.04A/gel) for one hour using the PowerPac™ HC High-Current Power Supply. In contrast, wet transferring involved soaking sponge pads, Whatman filter papers, and a nitrocellulose membrane in 1X transfer buffer for 5 mins. Then, two sponge pads, one piece of Whatman filter paper, the gel, one piece of nitrocellulose

membrane, one additional Whatman filter paper, and two more sponge pads were arranged on the cathode core of the Mini Blot Module (Invitrogen, US) according to the order of assembly in Figure 2-16B. Two sealed Mini Blot Modules were placed in the tank, which was filled with 1X transfer buffer. The electric current was maintained at a constant 30 volts for one hour using the PowerPac™ HC High-Current Power Supply.

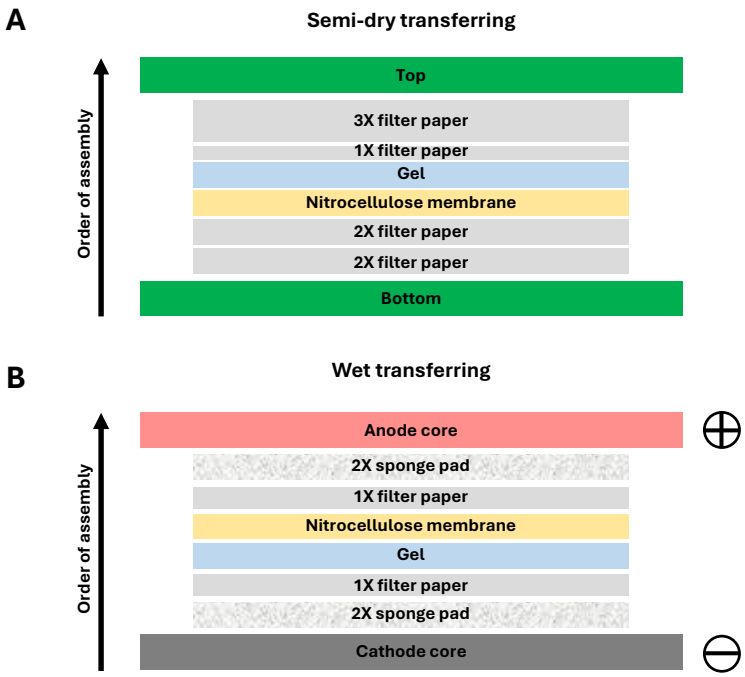


Figure 2-16 Order of assembly for semi-dry and wet transfer in western blotting

The assembly order of materials, including the sponge pad, filter paper, gel, and nitrocellulose membrane, was conducted for semi-dry transfer (A) and wet transfer in western blotting (B).

Once the transfer was completed, the membrane was removed and washed with distilled water to eliminate methanol before being stained with 5 mL of Ponceau S solution for one min. The protein bands appeared red and were then washed with distilled water to enhance their clarity. The membrane was subsequently incubated in 1X TBS on a shaker to eliminate the staining. Afterwards, the membrane was blocked with 12 mL of 5X BSA or 5% milk solution for one hour to prevent non-specific binding. The blot was incubated with the primary antibody, as detailed in Table 2-19, diluted as outlined in 10 mL at 4°C overnight. Next, the membrane was washed 3-5 times with 1X TBSN or TBST for 10 mins each before being conjugated with goat-anti-mouse or anti-rabbit HRP at a concentration of 1:10,000 for 1-2 hours. The membrane was washed 4 times with 1X TBSN or TBST

for 10 mins each, rinsed once with 1X TBS, and then washed twice with 1X TBS for 10 mins each. Lastly, the membrane was treated with Immobilon® Forte Western HRP substrate or SuperSignal™ West Femto Maximum Sensitivity Substrate for 5 mins, and the target protein bands were detected using the Licor Odyssey FC Imaging System (LI-COR Biotech, US) by visualising protein ladders at channels 600 and 700 for 2 mins each, and the protein at channel Chemi for at least 5 mins. Densitometry was analysed using Image Studio 6.0.

2.2.7 *In vitro* 3D Spheroids studies

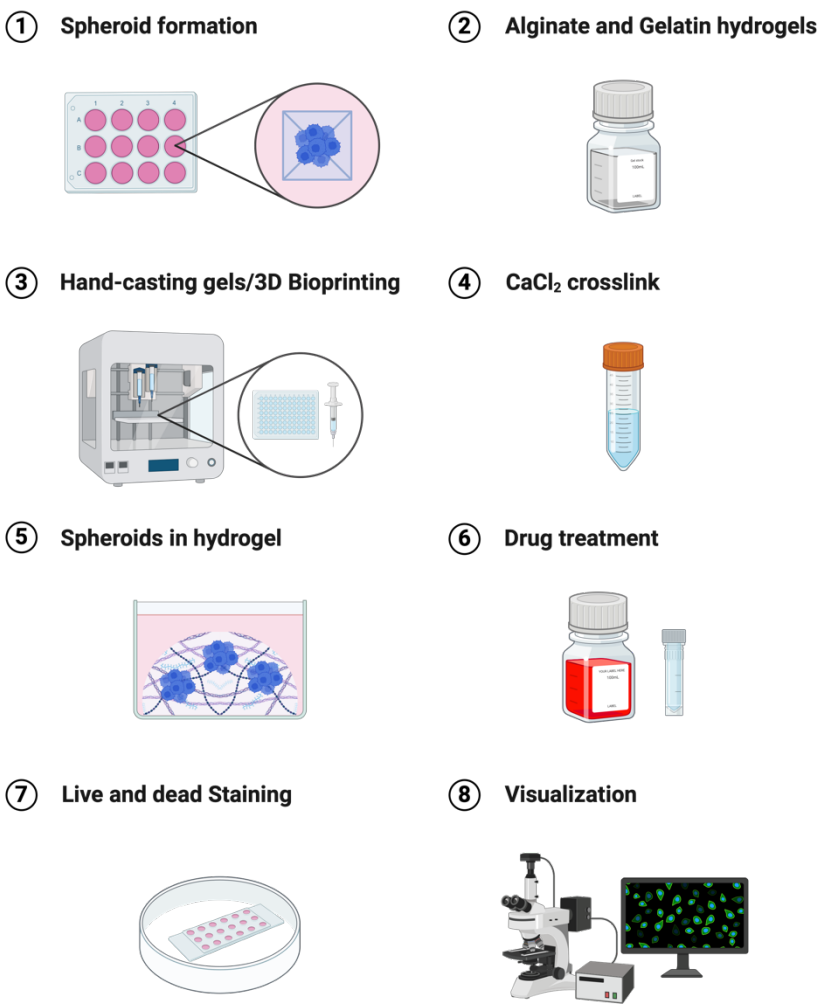


Figure 2-17 Workflow for developing an *in vitro* 3D culture model for drug screening in metastatic colorectal cancer

CRC spheroids were generated using a 3D spheroid culture plate, such as the AggreWell™ plate. Alginate and Gelatin hydrogels were utilised to support an *in vitro* 3D culture. Hand-casting or 3D bioprinting was conducted for the casting of hydrogels. $CaCl_2$ was employed to crosslink the hydrogels. The spheroids were cultured in the hydrogels and treated under drug conditions. Subsequently, the spheroids were stained with live and dead stains and analysed for cell viability.

2.2.7.1 Generating 3D Spheroids by AggreWell™ plates

The 24-well plate (AggreWell™ 400) was pre-treated with 500 µL of an anti-adherence solution and spun at 1,300 x g for 5 mins. Next, each well was rinsed with 500 µL of 1X PBS before adding 200 µL of fresh warm medium. DLD1 and SW620 were harvested and seeded at densities of 50 and 100 cells per microwell, by resuspending them at 6 x 10^4 and 1.2 x 10^5 cells/mL, respectively. A cell suspension was prepared in a total volume of 9 mL, which was sufficient for dispensing 300 µL of cells into each well of the plate using the reverse pipetting technique. The plate was then centrifuged at 100 x g for 3 mins to settle the cells in the microwells. The cells were cultured in a normal culturing medium, totaling 2 mL per well, at 37°C and 5% CO_2 for 7 days. An additional 500 µL of fresh medium was gently added to each well every 2-3 days. On days 3, 5, and 7, spheroids were stained with Live/Dead staining to assess cell viability and determine the optimal duration for spheroid formation.

To harvest the spheroids, a pipette (P1000) was used to gently aspirate and dispense to dislodge the spheroids, which were then collected onto a 37 µm Reversible Strainer placed ontop of a 50 mL tube. The spheroids remained on the filter and were washed gently 3-5 times with 1X PBS through the strainer. The strainer was inverted and placed onto a new 50 mL tube that had been pre-rinsed with 1X PBS containing 2% FBS. Subsequently, the spheroids were eluted with 3-6 mL of medium. 50 µL of spheroid suspension were aliquoted into a 96-well plate, counted under a light microscope at 10X magnification, and images were acquired using a ZEISS Axiocam 705 mono microscope with ZEN 3.5 blue edition. The spheroid yield was calculated using the equation below.

$$\text{Total number of spheroids} = \frac{\text{no. of counted spheroids}}{50\ \mu\text{L}} \text{ x volume of spheroid suspension } (\mu\text{L})$$

In subsequent experiments, we discovered that SW620 required a minimum of five days for spheroid formation to occur and that, at a density of 100 cells per microwell, it improved the ability to generate a spheroid.

2.2.7.2 Alginate and Gelatin hydrogel and calcium chloride preparation

The hydrogel, consisting of 2% alginate and 8% gelatin (2% Alg and 8% Gel), was previously developed by Dr Lauren Hope to create a 3D *in vitro* model for bone marrow culture. The gels were cross-linked with calcium chloride ($CaCl_2$), resulting in a range of stiffness and a highly porous network. This setup facilitates a physiological microenvironment for culturing cells under both normal and disease conditions.

To prepare a 1.5X stock solution of 2% alginate and 8% gelatin hydrogel in 20 mL of 1X PBS, sodium alginate and gelatin type A powders were weighed at 0.6 g and 2.4 g, respectively. Two aliquots of gel powders, 50 mL of 1X PBS, an autoclaved 100 mL bottle, a magnetic stir bar, and a spatula were sterilised under UV light for 30 mins. Next, the sodium alginate powder was added to 20 mL of 1X PBS in the bottle using the magnetic stir bar. The bottle was sealed with parafilm and stirred on a hotplate at 37°C at 500-700 rpm. The alginate dissolved after 1-2 hours of mixing, this was followed by the gentle addition of gelatin powder to the bottle under sterile conditions. The gel mixture was stirred gradually to prevent clumping against the sides of the bottle and was then incubated at 37°C overnight. The gel stock was stored at 4°C and warmed to 37°C for a minimum of 1 hour before use (overnight incubation is recommended).

A 1 M calcium chloride stock solution was prepared by dissolving 21.908 g of calcium chloride hexahydrate crystals in 100 mL of ddH_2O, which was stored at room temperature. A 150 mM calcium chloride working solution, consisting of 7.5 mL of 1M $CaCl_2$ and 42.5 mL of ddH_2O, was sterilised before use by filtering it through a 0.2 µm filter.

2.2.7.3 The hand-casting hydrogels

Initial studies involved hand casting hydrogels. All necessary materials, including 1 mL syringes, 1.5 mL tubes, 21 G short blunt needles, a multichannel pipette, and a spatula, were sterilised under UV light for 30 mins. The warm, dissolved 1.5X hydrogel was transferred into four 1.5 mL tubes using 1 mL syringes, each containing 1 mL. Next, 250 µL of warm normal medium was added to each gel aliquot tube. The aliquots were mixed using a pipette and incubated at 37°C until use. The SW620 spheroids, cultured five days prior, were harvested and aliquoted

to yield approximately 15,000 spheroids. The spheroids were then centrifuged at 300 x g for 5 mins and resuspended in a total of 1 mL of fresh medium. 250 µL of the spheroid suspension was added to 1.25 mL of the gel aliquot, resulting in a prepared spheroid suspension of 4,000 spheroids per 1.5 mL. The hydrogel/spheroid mix was aspirated using a syringe and sealed with a needle. One drop of the hydrogel/spheroid mix was then dropped into the centre of each well of a 96-well plate containing 200 µL of 150 mM $CaCl_2$ as quickly as possible to avoid solidification. The gels were incubated for 8 mins and then gently washed twice with the medium by tilting the plate to prevent the gels from being aspirated off. The plate was incubated at 37°C for 24 hours before initiating the drug treatment.

2.2.7.4 3D bioprinting hydrogels

The materials for 3D printing included three 1 mL syringes, two 2 mL tubes, two 3 mL cartridges in a Petri dish, two pistons in a Petri dish, and two 21 G short blunt needles in a Petri dish, a multichannel pipette, and a spatula, all sterilised under UV light for 30 mins. Subsequently, the warm dissolved 1.5X hydrogel was aspirated into two 2 mL tubes using 1 mL syringes, each containing 1.2 mL. 300 µL of the warm normal medium were added to each gel aliquot tube, mixed, and incubated at 37°C until use. Each well of a 96-well plate was filled with 200 µL of 150 mM $CaCl_2$ and kept at room temperature. The SW620-BFP spheroids, cultured for five days before 3D bioprinting, were harvested. Afterwards, SW620 spheroids were aliquoted to yield approximately 3,600 spheroids, spun at 300 x g for 5 mins, and then resuspended in a total of 600 µL of fresh medium. 300 µL of the spheroid suspension were added to 1.5 mL of the gel aliquot, resulting in a prepared spheroid suspension at 1,000 spheroids/mL, totaling 3.6 mL of 1X hydrogel. Two tubes of the hydrogel were loaded immediately into the cartridges using a syringe. The cartridge was closed with a piston at the end, and a spatula was used to push the piston in to eliminate air inside the cartridge. Finally, a needle was added to the tip of the cartridge.

The 3D bioprinter (BIO X6, CELLINK A Bico Company, Sweden) was sterilised, and the printing head was maintained at 37°C throughout the printing process. The cartridge was loaded into the printing head, and it contained liquid hydrogel until printing commenced. The 3D printing model was designed and set as a droplet

using DNA Studio V3.3.2+7 (CELLINKA Bico Company, Sweden). The printing pressure ranged from 30 to 40 kPa, and the speed was between 1.05 and 1.35 secs. This setup produces a uniform size of numerous gels and minimises human error. A droplet of the hydrogel was printed into each well of a 96-well plate containing 200 µL of 150 mM $CaCl_2$. After the gels were incubated for 8 mins, they were gently washed twice with the medium. The plate was incubated at 37°C for 24 hours before starting the drug treatment.

2.2.7.5 Drug treatment of SW620 spheroids in hydrogels

After the hydrogels were cast for 24 hours, the medium was gently removed and replaced with 200 µL of treatment conditions, which included normal medium, a 0.1% DMSO solution as a vehicle control, 20 µM of 5-FU, 2 µM of LDN-212854, and a combination of the two drugs (20 µM of 5-FU and 2 µM of LDN-212854) for 7 days. The treatment conditions were refreshed every 2 days. The gels were stained with Live/Dead staining to examine the cell viability of spheroids in the hydrogels on days 0, 1, 3, 5, and 7 after casting, as well as on days 4 and 7 of the drug treatment.

2.2.7.6 Live/Dead staining and cell viability analysis of spheroids

Regarding cell viability, as indicated by intracellular esterase activity and plasma membrane integrity, Calcein AM, a 4 mM stock solution, is a cell-permeant dye recognised for its ability to signal widespread intracellular esterase activity in live cells by exhibiting green fluorescence. Ethidium homodimer-1 (EthD-1), a 2 mM stock solution, is a nucleic acid-binding dye that penetrates damaged cell membranes and produces bright red fluorescence in dead cells.

To assess cell viability and determine the optimal duration for spheroid formation using the AggreWell™ plate, spheroids in each well were stained with 500 µL of a dye mixture containing 1 µM Calcein AM and 2 µM EthD-1 at 37°C for 30 to 40 mins in the dark. Following this, nuclei were stained with 1.6 µM Hoechst 33342 for 10 to 15 mins on days 3, 5, and 7 in culture. Subsequently, the stained spheroids were harvested, seeded onto a 96-well plate, and examined under a light microscope at 10X magnification. Images were captured using a ZEISS Axiocam 705 mono microscope with ZEN 3.5 blue edition, utilising the Calce channel for Calcein AM and the EtHD1 channel for EthD-1.

To investigate the cell viability of spheroids in the hydrogels, the medium was aspirated from the gel, and it was stained with 200 µL of a dye mixture containing 2 µM Calcein AM and 4 µM EthD-1 at 37°C for 20 to 40 mins. Afterwards, the stained gel was removed and placed on a slide using a spatula. Images were captured using a ZEISS Axiocam 705 mono, utilising channel Syt40 for BFP and an EVOS M7000 microscope (Thermo Fisher Scientific Inc., US), utilising channel DAPI for BFP, channel GFP for Calcein AM, and channel for EthD-1.

Fiji, also known as ImageJ2, is a scientific image analysis software used to evaluate images of cells and spheroids in terms of morphology and fluorescence-probing studies (218). To analyse the cell viability of spheroids, we established a threshold to focus on the area of interest (the spheroid) in an 8-bit image. Subsequently, the original image was divided into green, red, and blue channels, representing live, dead, and BFP categories, respectively. The spheroid was selected in the same region across all channels and was then assessed for spheroid area (size), minimum and maximum values, mean grey value, shape descriptors, and integrated density, Figure 2-18. The image scales were adjusted according to the microscope models. For a ZEISS Axiocam 705 mono microscope: 5X magnification at 0.7246 pixels/µm, 10X magnification at 1.4493 pixels/µm, and 20X magnification at 1.9324 pixels/µm. For an EVOS M7000 microscope: 10X magnification at 1.6075 pixels/µm and 20X magnification at 3.215 pixels/µm. The fluorescent intensity of each spheroid was determined using the equation below.

$$\text{Corrected total cell fluorescence (CTCF)} = \text{Integrated density} - (\text{Area spheroids x Mean gray of background})$$

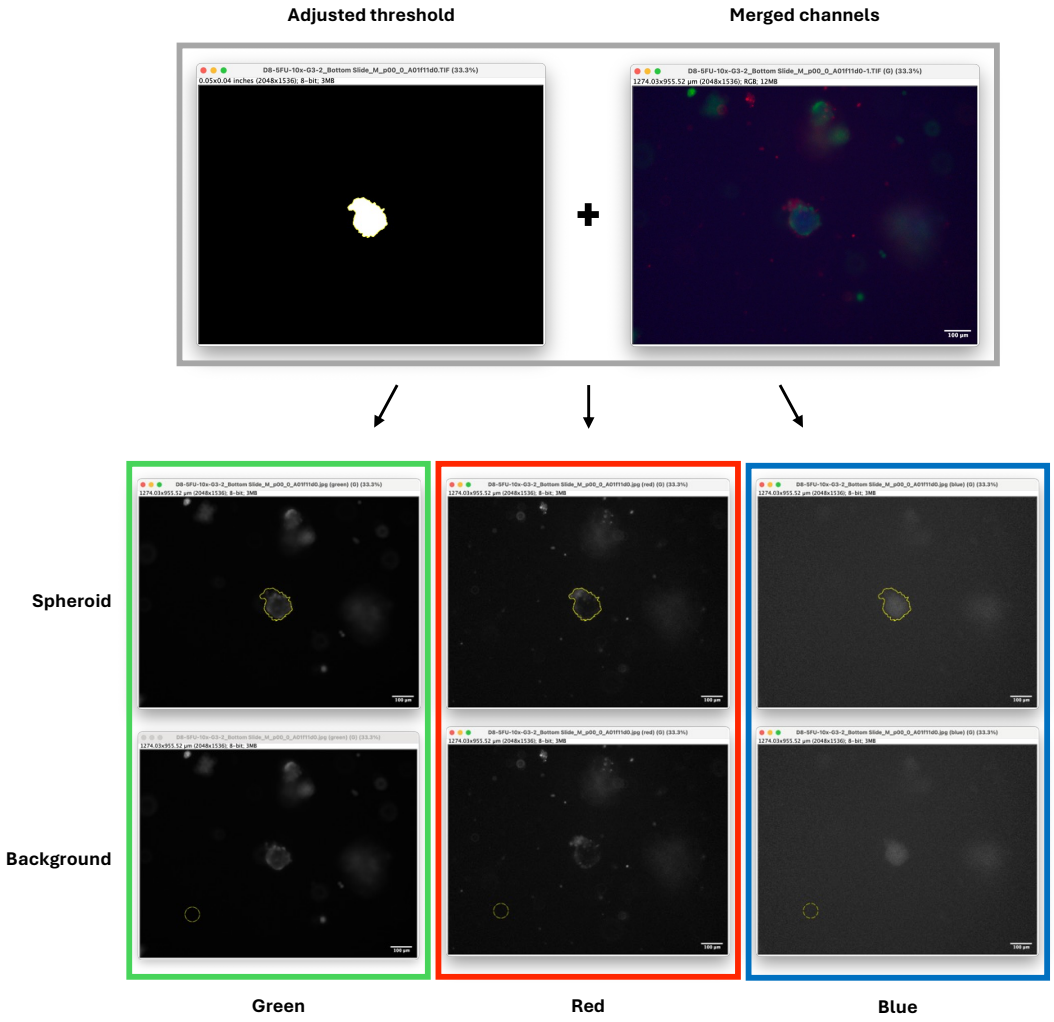


Figure 2-18 Image processing for spheroid analysis using Fiji

An image of spheroids in hydrogels was converted to an 8-bit format and adjusted to the threshold to focus on the area of interest. Subsequently, the original image was divided into three channels: green, red, and blue. For analysis, each channel was selected from the same area of interest and background.

2.2.8 Statistical analysis

All data are presented as the average of at least three replicates, with standard deviation represented by error bars unless otherwise stated. This excludes data obtained from *in silico* analysis using the DepMap and Stemformatics databases, as well as prognostic significance and transcriptomic analysis using patient tissues, which were performed with RStudio. Microsoft Excel was utilised to process raw data, and GraphPad Prism 10 was employed for statistical analysis. For data with a normal distribution, a T-test assessed significant differences between two independent groups. A one-way ANOVA with multiple comparisons (Dunnett's multiple comparisons test) was used to compare drug treatment conditions with

vehicle control. Conversely, the Mann-Whitney U test and Wilcoxon test compared two independent groups of non-normally distributed data, and the Kruskal-Wallis test with Dunn's multiple comparisons post-test compared multiple groups in nonparametric analyses. The significance levels were indicated by asterisks as follows: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

Chapter 3 Evaluation of intrinsic expression patterns of the TGF-β and BMP signalling pathways in CRC cell lines and cells within the bone marrow.

3.1 Introduction

CRC is an aggressive disease characterised by heterogeneity and various factors involved in tumourigenesis (1-3). TGF-β and BMP signalling, crucial pathways in CRC, play an essential role in maintaining tissue homeostasis within the intestinal crypts (11-15). The dysregulation of these pathways is highlighted in the development of CRC (2, 17, 18, 124, 125). Inactivation of signalling during the early stages of cancer is associated with a reduction in tumour suppression. Conversely, tumour promotion in advanced stages of cancer is facilitated by alterations in TGF-β and BMP signalling (162). The crosstalk between signalling pathways, including TGF-β and BMP signalling, serves as an important regulator in both tumour cells and their microenvironment, contributing to invasive features and leading to metastasis (2, 17, 18, 124, 125). However, a better understanding of the dysregulation of these signalling pathways during carcinogenesis, particularly in promoting CRC metastasis, requires further elucidation.

Metastatic CRC presents a significant challenge due to its complex therapeutic management and is associated with treatment resistance, leading to poorer survival outcomes (7, 131). There is a limited information on the understanding the underlying altered mechanisms of biological processes in tumour cells and their microenvironment at the target sites (7, 131, 132). Of the metastatic CRC sites, the spread to the bone is associated with a shorter survival time compared to the peritoneum, liver, and lung (129). Although bone metastasis from CRC is rare, there is a lack of knowledge about how CRC can survive and grow in the bone. Studies into the mechanism are of importance and could enhance the development of treatment strategies (8-10).

TGF-β and BMP signalling are well-established as key pathways in colon crypts, involved in cell growth, apoptosis, and differentiation (11-15). Furthermore, recognised mutations in CRC, such as those in *TGFBR2* and *SMAD4* (17, 124, 164-166), along with dysregulation of BMP2, BMP4, and BMP9 (167-171), and mutations

in *BMPR1A* (22, 23), have been reported to contribute to CRC development. However, their role in CRC with bone metastasis has not been extensively studied. In the BM, it is known that the high cell-to-cell interaction and dynamic signalling pathways within the BM niche, especially the BMP signalling pathway, contribute to bone homeostasis (150-152). Greater emphasis on bone metastasis has been studied in other primary cancers, such as breast and prostate cancer, as both are more frequently associated with invasive behaviour into the bone. The adaptive microenvironmental features of the BM niche are proposed to arise in response to systemic stimulation from tumour cells, resulting in a premetastatic niche for colonisation and serving as a protective environment during the dormancy stage and reactivation in the later stage (219). These alterations may also occur in the bone with metastatic CRC.

The molecular signature represents a promising approach to elucidated the relationship between the primary tumour and its tissue sites of metastasis, providing a information as preliminary data for future investigations (219, 220). However, there are limited biological resources available for studying bone metastasis in CRC. At present, bioinformatic analysis of public omics databases serves as a valuable tool. Consequently, this study aimed to identify dysregulation of the gene expression profile using available online databases, with a focus on the TGF-β and BMP signalling pathways in CRC cells and bone marrow cells.

3.2 *In silico* analysis of TGF-β and BMP pathways in CRC cell lines using the gene expression database from the Cancer Dependency Map (DepMap) portal

3.2.1 Differential expression of TGF-β and BMP signalling was observed across CRC cell lines.

CRC cells, may exhibit a specific molecular signature characterised by heterogeneity and an aggressive phenotype, providing a fundamental understanding of the altered mechanisms of CRC and metastasis. To investigate the intrinsic expression patterns, particularly within the TGF-β and BMP signalling pathways across the CRC cell lineage, genes in the TGF-β, Activin, BMP, and SMAD signalling pathways (Table 2-1) were analysed using mRNA expression levels from the DepMap dataset (Section 2.2.1.1). The CRC cell lines (Table 2-3) available for use in the subsequent investigation were analysed to determine their expression

patterns. The cell lines used were DLD-1 and HCT116 derived from primary colon cancer, SW837 derived from primary rectal cancer and SW620 which is a metastatic colon cancer cell line. Consequently, the expression levels of each signalling member were compared across the 4 different CRC cell lines and visualised in a heatmap of TGF-β (Figure 3-1A), Activin (Figure 3-1B), BMP (Figure 3-1C), and SMAD (Figure 3-1D) signalling pathways.

The overall expression pattern of TGF-β signalling indicated an upregulation of signalling in the CRC cell lines; however, a downregulation of *TGFB2* was observed in HCT116 and SW620 cells. The differential expression pattern of Activin signalling is observed in CRC, where it is upregulated in most Activin receptors. The expression of BMP signalling exhibited varying levels in both ligands and receptors. Interestingly, an upregulation of B*MP4*, *BMPR1A*, and *BMPR2* was identified in all 4 CRC cell lines. As part of signal transduction, SMAD signalling showed upregulation of R-SMAD genes (*SMAD 1, 2, 3*, and *5*) and the I-SMAD gene (*SMAD7*) in all of the CRC cell lines, while the remaining R-SMAD and I-SMAD genes (*SMAD9* and *SMAD6*, respectively) were upregulated in some of the cell lines. Notably, the co-SMAD gene (*SMAD4*) was downregulated in SW620 cells.

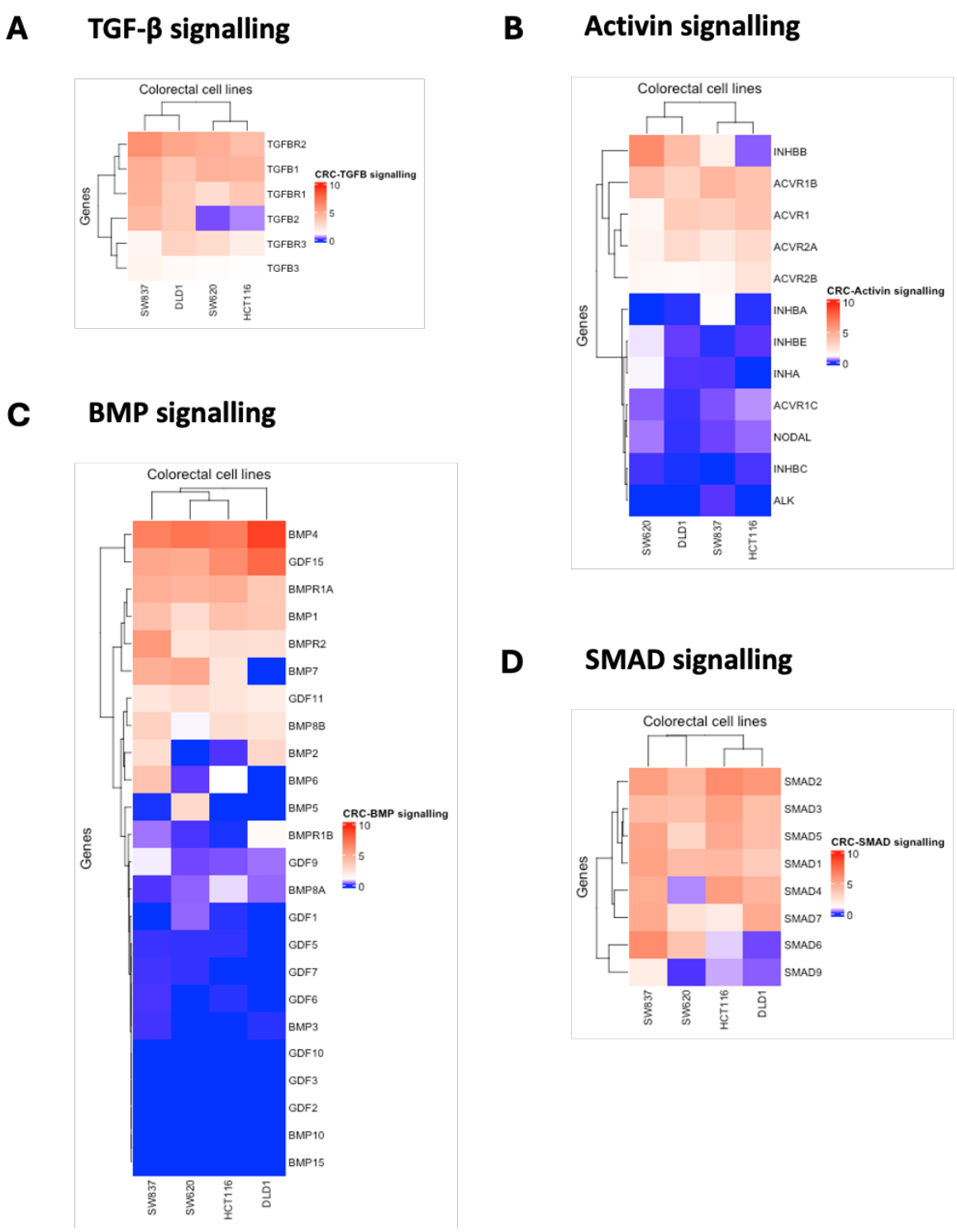


Figure 3-1 Heatmaps of the TGF-β superfamily in colorectal cancer cell lines

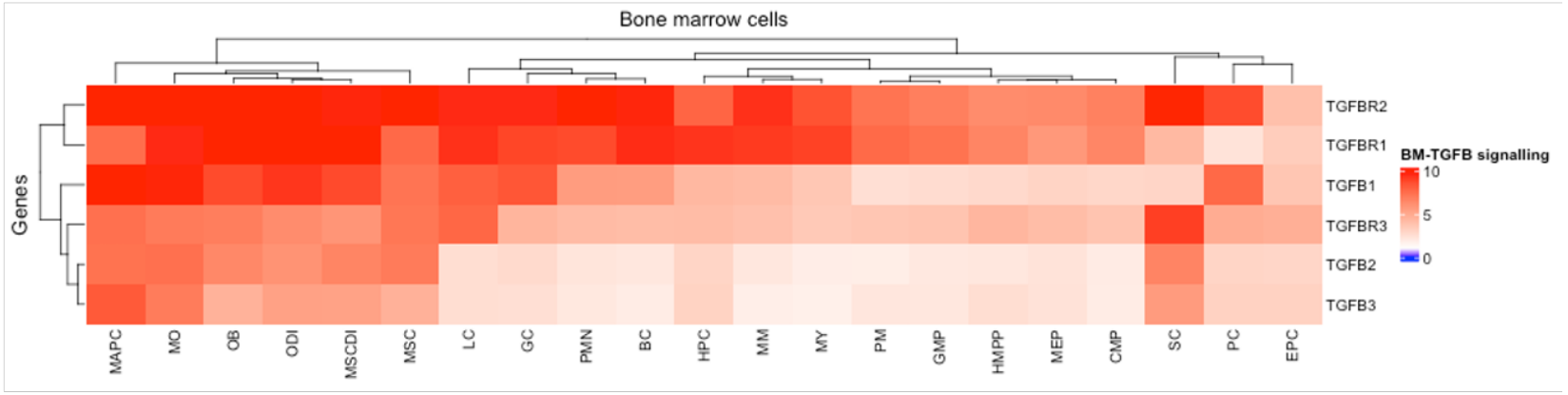
Four heatmaps, comprising TGF-β signalling (A), Activin signalling (B), BMP signalling (C), and SMAD signalling (D), displayed differential expression of TGF-β and BMP members in four colorectal cancer cell lines (DLD1: Primary Colon Adenocarcinoma, CMS1; HCT116: Primary Colon Adenocarcinoma, CMS4; SW620: Metastatic Colon Adenocarcinoma, CMS4; SW837: Primary Rectal Adenocarcinoma, CMS4). The unit of gene expression is represented in transcription per million (TPM). Higher expression is indicated in red, while lower expression is indicated in blue.

3.3 *In silico* analysis of TGF-β and BMP signalling pathways in disease-free bone marrow using Stemformatics

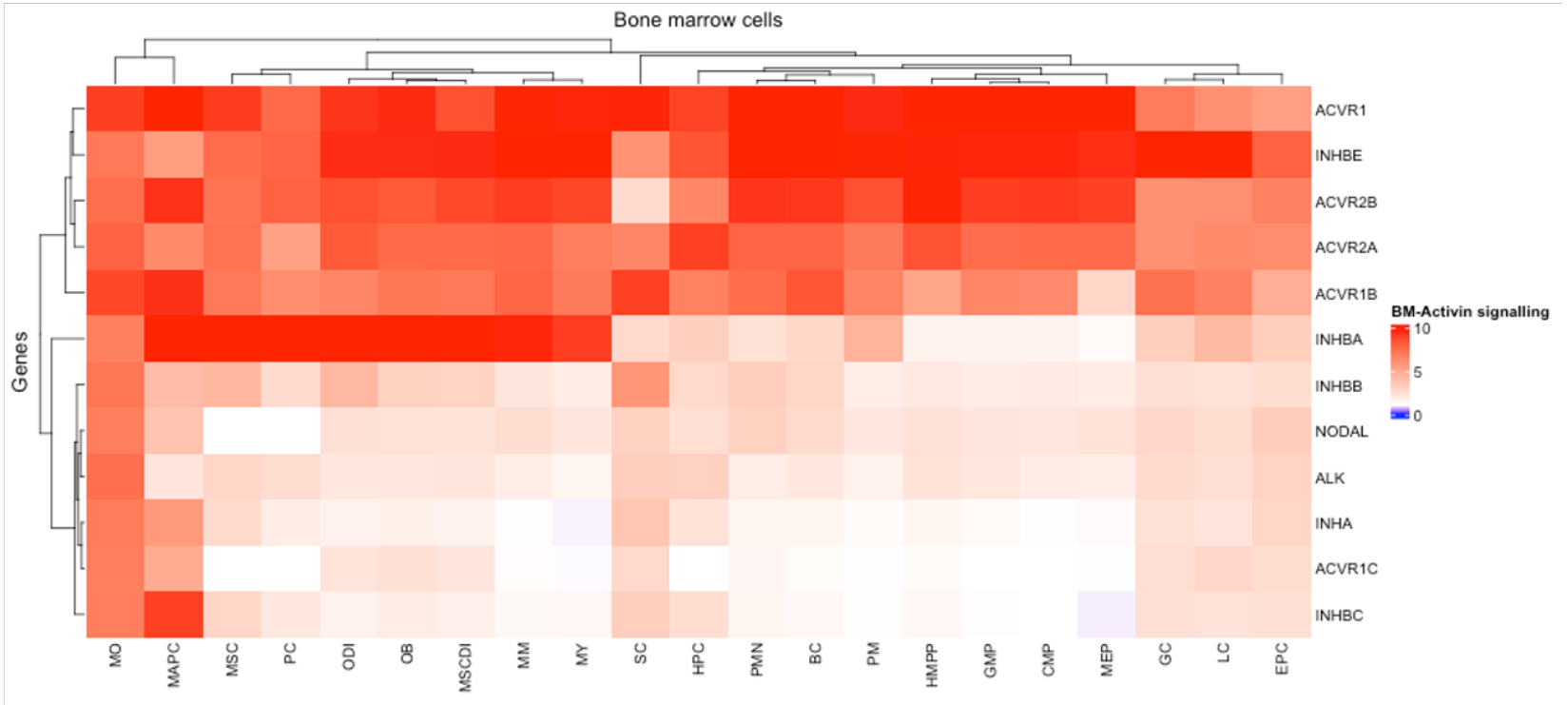
3.3.1 High expression of TGF-β and BMP signalling was observed across disease-free bone marrow

It is essential to gain a deeper understanding of bone marrow as the targeted site of metastasis. Unfortunately, DepMap provides limited expression data for bone. However, more mRNA expression data for BM cells are available in the Stemformatics database. Therefore, datasets of disease-free BM cells (Table 2-2) were accessed for gene expression analysis, as described in section 2.2.1.2. Normalised expression levels of TGF-ß and BMP signalling from microarray experiments involving cell lines and adult human samples were extracted, and the means of gene expression for each BM cell were calculated. The heatmaps demonstrated that BM cells expressed high levels of many genes in the TGF-ß superfamily (Figure 3-2). For example, *TGFR2*, *ACVR1*, *BMP1*, *BMPR1A*, *BMPR2*, and *SMAD2* were found to have the highest expression levels in all BM cells, as shown in the results for TGF-ß signalling (Figure 3-2A), Activin signalling (Figure 3-2B), BMP signalling (Figure 3-2C), and SMAD signalling (Figure 3-2D). In addition, upregulation of SMAD4 was observed across BM cells.

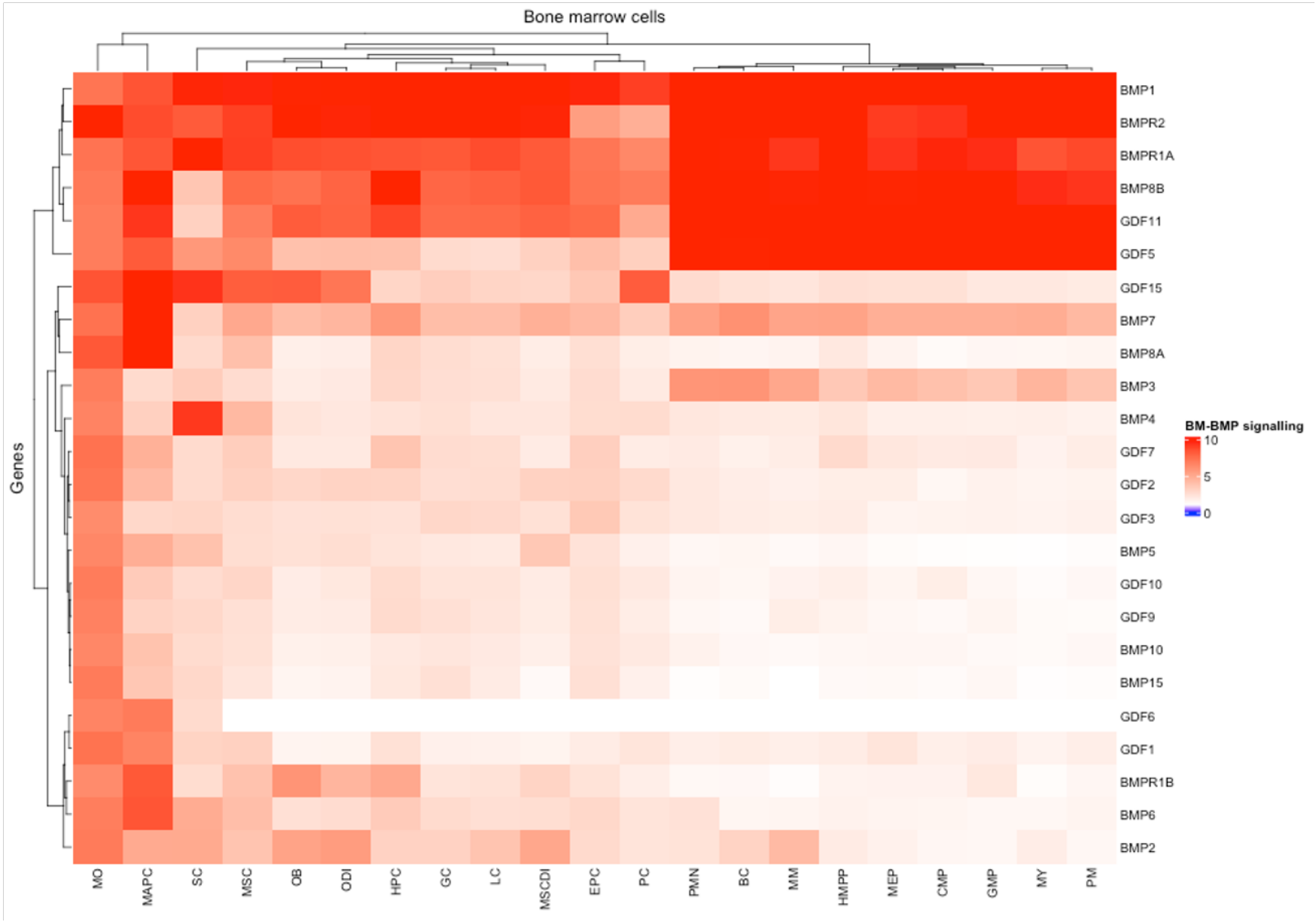
A TGF-β signalling



B Activin signalling



C **BMP signalling**



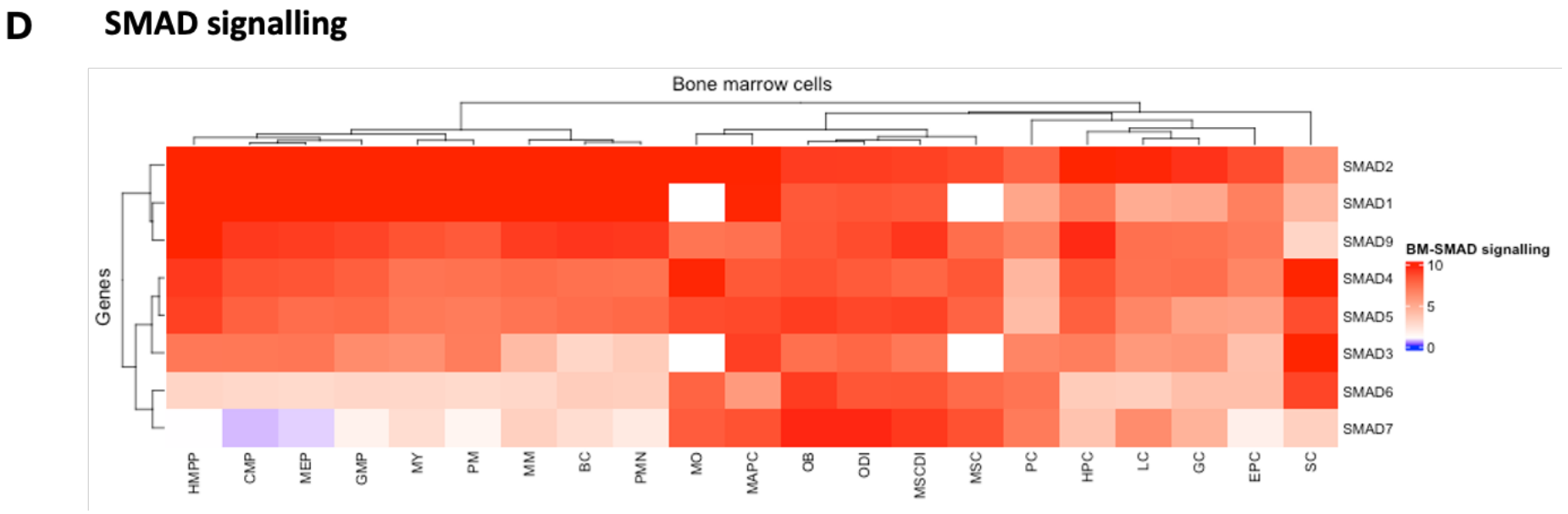


Figure 3-2 Heatmaps of the TGF-β superfamily in disease-free BM cell types

Four heatmaps, comprising TGF-β signalling (A), Activin signalling (B), BMP signalling (C), and SMAD signalling (D), displayed overall high expression levels of TGF-β and BMP members across disease-free BM cell types: Band cell (BC), Common myeloid progenitor (CMP), Erythropoietic cell (EPC), Granulocyte (GC), Granulocyte-macrophage progenitor cell (GMP), Haematopoietic precursor cell (HPC), Leukocyte (LC), Multipotent adult progenitor cell (MAPC), Megakaryocyte-erythroid progenitor cell (MEP), Metamyelocyte (MM), Macrophage (MO), Haematopoietic multipotent progenitor cell (MPP), Mesenchymal stromal cell (MSC), MSC differentiation intermediate (MSCDI), Myelocyte (MY), Osteoblast (OB), Osteogenic differentiation intermediate (ODI), Pericyte cell (PC), Promyelocyte (PM), Polymorphonuclear cell (PMN), and Stromal cell (SC). The unit of gene expression is represented as average values of log-normalised expression from each BM cell type. Higher expression is indicated in red, whereas lower expression is indicated in blue.

3.4 RNA and protein expression analysis of TGF-β and BMP signalling in CRC cell lines.

3.4.1 Downregulation of SMAD4 in metastatic CRC cell lines

The bioinformatic results from various cells in bone marrow indicated that TGF-ß and BMP signalling are activated under disease-free conditions. In contrast, CRC cell lines displayed a different regulatory pattern for the signalling pathways. Nevertheless, an upregulation of signal activation was observed in all CRC cells. Prior to identifying the relevant components of the TGF-ß and BMP signalling pathways in CRC associated with unfavourable survival outcomes, qPCR was utilised to validate the mRNA expression of the key components in the signalling pathway across five CRC cell lines, including 4 CRC cell lines with the DepMap data, and COLO205 cells (a metastatic colon cancer cell line), as described in section 2.2.5. As a result, a similar expression pattern was observed between the bioinformatic (Figure 3-3A) and qPCR analyses (Figure 3-3B). This confirms that the upregulation of TGF-ß and BMP upstream signalling is evident in CRC, with variable expression levels of the signal ligands. *SMAD9* exhibited reduced expression in some CRC cells (DLD1, SW620, and COLO205), and *SMAD4* was downregulated in both metastatic CRC cells (SW620 and COLO205).

Western blot analysis was performed to investigate protein levels and phosphorylation within the SMAD signal transduction pathway in CRC cell lines, as described in section 2.2.6. The results are presented in Figure 3-3C. In response to TGF-ß activation, SMAD2 and SMAD3 were found to be phosphorylated in CRC. SMAD1/5/8, as activated by BMP signalling, was predominantly phosphorylated in the SW620 cell line compared to other CRC cell lines. SMAD4 protein was not detected in metastatic cell lines (SW620 and COLO205), and a decrease in SMAD4 was observed in the SW837 cell line.

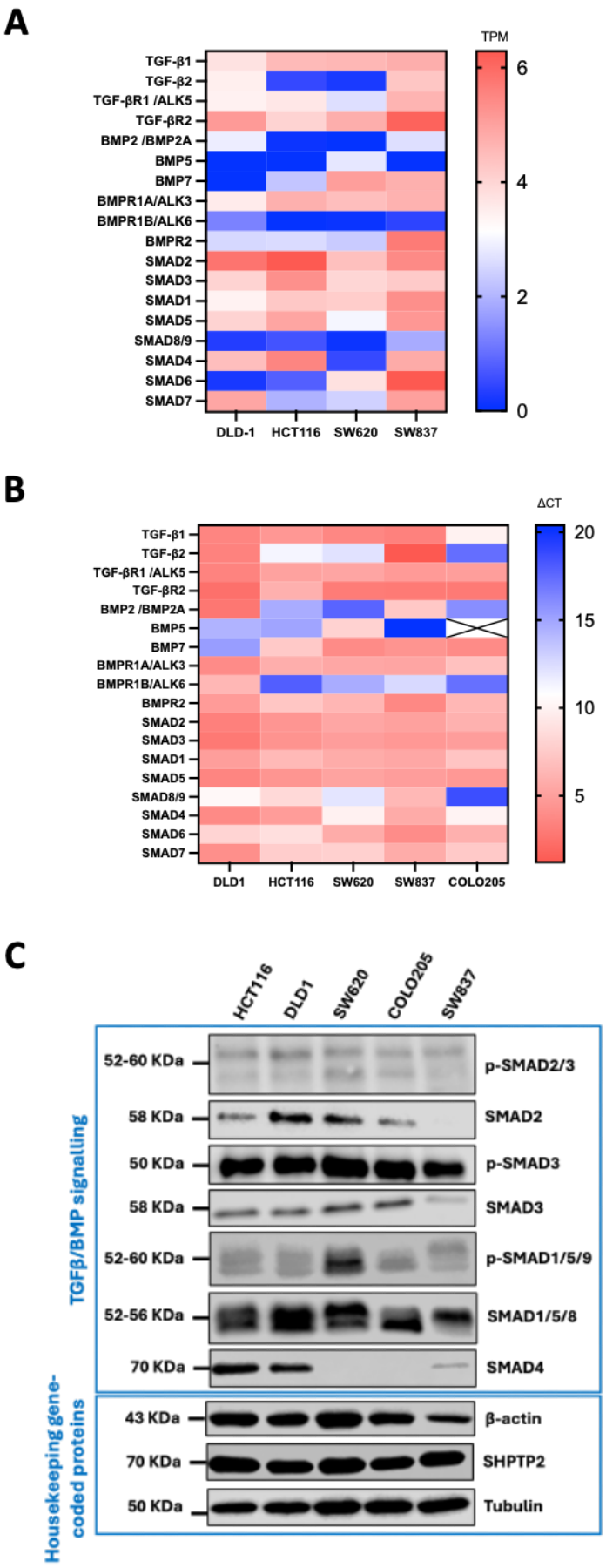

A
TPM
6
4
2
0
TGF-β1
TGF-β2
TGF-βR1 /ALK5
TGF-βR2
BMP2 /BMP2A
BMP5
BMP7
BMPR1A/ALK3
BMPR1B/ALK6
BMPR2
SMAD2
SMAD3
SMAD1
SMAD5
SMAD8/9
SMAD4
SMAD6
SMAD7
DLD-1
HCT116
SW620
SW837
B
ΔCT
20
15
10
5
TGF-β1
TGF-β2
TGF-βR1 /ALK5
TGF-βR2
BMP2 /BMP2A
BMP5
BMP7
BMPR1A/ALK3
BMPR1B/ALK6
BMPR2
SMAD2
SMAD3
SMAD1
SMAD5
SMAD8/9
SMAD4
SMAD6
SMAD7
DLD1
HCT116
SW620
SW837
COLO205
C
HCT116
DLD1
SW620
COLO205
SW837
TGFβ/BMP signalling
52-60 KDa
p-SMAD2/3
58 KDa
SMAD2
50 KDa
p-SMAD3
58 KDa
SMAD3
52-60 KDa
p-SMAD1/5/9
52-56 KDa
SMAD1/5/8
70 KDa
SMAD4
Housekeeping gene-coded proteins
43 KDa
β-actin
70 KDa
SHPTP2
50 KDa
Tubulin

Figure 3-3 RNA and protein expression of the TGF-β and BMP signalling pathways in colorectal cancer cell lines

A heatmap displays the gene expression levels of TGF-β and BMP signalling pathways using data from DepMap (A). A heatmap was generated to illustrate the validation results of mRNA expression in five CRC cell lines through qPCR analysis (B; n = 1). A Western blot result shows protein levels and phosphorylation of the SMAD signal transduction pathway in the CRC cell lines (C; n =1). The five CRC cell lines include DLD1 (Primary Colon Adenocarcinoma, CMS1), HCT116 (Primary Colon Adenocarcinoma, CMS4), SW620 (Metastatic Colon Adenocarcinoma, CMS4), COLO205 (Metastatic Colon Adenocarcinoma) and SW837 (Primary Rectal Adenocarcinoma, CMS4). In the DepMap heatmap, the unit of gene expression is represented in transcription per million (TPM). Higher expression is indicated in red, while lower expression is indicated in blue. In a heatmap of the qPCR analysis, the unit of gene expression is represented in ΔCt, Lower ΔCt refers to higher gene expression and is shown in red, while higher ΔCt refers to lower gene expression and is shown in blue.

3.5 Discussion

The role of the TGF-ß and BMP signalling pathways in regulating tissue homeostasis and the dysregulation of these pathways has been highlighted as a key factor in driving cancer development and progression. Changes in biological processes regulated by these signalling pathways occur not only within tumour cells but also to the surrounding components of the microenvironment, including the distant target tissue site of metastasis (162, 163). A shift in the function of TGF-ß and BMP signalling between tumour suppression and tumour promotion is suggested in CRC (16-18). An increasing risk of CRC has been reported to be associated with mutations in the TGF-ß and BMP signalling pathways in both sporadic and hereditary CRC, leading to the development of colorectal polyps (premalignant lesions) into progressive phenotypes of the cancer (2, 3, 19-21). Mutations in *TGFBR2* and *SMAD4* are primarily recognised as crucial factors in CRC, as well as in JPS, which also involves a mutation in *SMAD4* (21, 35, 221). However, other signalling components remain relevant in their signal transduction, and the dysfunction remains to be elucidated. For instance, the *BMPR1A* mutation is also implicated in both premalignant and malignant stages of CRC (22, 23).

The CMS classification in CRC is based on a large-scale data analysis of 18 distinct public gene expression datasets from 4,151 patients provides a system for grouping patients based on their tumour molecular signatures; CMS1, CMS2, CMS3 and CMS4. CMS4 patients are defined as having TGF-ß activation, stromal infiltration, EMT, and angiogenesis. Additionally, they were predicted to experience the worst survival outcomes compared to other CMS groups (57). Thus, the findings based on *in silico* analysis presented in this chapter enhance the understanding of the

baseline gene expression levels of each signalling component of the TGF-β superfamily, which are altered during the progression of CRC across several CRC cell lines. Activation of the upstream signalling receptors, such as TGFBR2, ACVR1B, and BMPR1A, in addition differential activation of signalling ligands, was observed in the CRC cell lines. Consequently, most of the signal transduction genes were upregulated, including SMAD1, SMAD2, SMAD3, and SMAD5. Interestingly, SMAD4 was found to have a reduced level in metastatic CRC cells. The expression of key intracellular signal components at both the RNA and protein levels, as well as their phosphorylation status, was confirmed to be consistent through both qPCR and Western blot analysis. The phosphorylation of SMAD1/5/8 was observed in SW620 cells; however, a loss of SMAD4 was detected.

A lack of understanding of bone marrow as a target site for CRC metastasis remains a challenge for investigation. Despite the rare incidence of CRC spreading to the bone marrow, the most severe survival outcomes have been predicted, and novel treatment approaches would be beneficial (7, 131, 132). To gain a fundamental understanding of TGF-β and BMP signalling in the BM niche, the intrinsic gene expression levels of each signalling component were examined using public transcriptomic data. This revealed that the highly active regulation of these signalling pathways in various cell types maintains homeostasis within this microenvironment. Proteomic analysis of osteoblasts and MSCs derived from bone marrow indicated minor age- and sex-dependent differences in the proteomic profiles of these two cell types. Furthermore, osteoblast-specific proteins have been shown to be involved in intracellular signalling (Wnt/Notch/BMP pathways (222). While HPCs with ageing reveal reduced functionality and altered myeloid differentiation, they cause changes in the bone marrow niche. Ageing of stem cells has been reported to result in a high turnover of haematopoiesis, which can lead to conditions such as anaemia and haematologic malignancies (223).

Limitations persist when studying patient tissues due to the challenges associated with collecting specimens of CRC metastasis, particularly in cases where it spreads to the bone. Only a limited number of studies on transcriptomic analyses on patient samples have been conducted in relation to bone metastasis from breast cancer and non-small cell lung cancer (NSCLC), as well as investigations into osteosarcoma (224-226). Recently, a single-cell profiling of the bone metastasis microenvironment from multiple cancer types, including CRC, identified three

immune archetypes: monocyte-enriched (Mono), macrophage/osteoclast-enriched (Mφ-OC), and regulatory/exhausted T cell-enriched (Treg-Tex). These findings reveal the metastatic heterogeneity of bone marrow across different cancers, which are engaged in distinct mechanisms of immune suppression and bone remodelling (227). Consequently, developing a novel *in vitro* co-culture model of metastatic colorectal cancer (CRC) with bone marrow cells would serve as an alternative and valuable tool for elucidating the underlying specific mechanisms of biological processes in CRC metastasis. As a relevant example, bone metastasis models involving breast cancer are gaining increasing attention, providing new insights into how the bone marrow creates a specific microenvironment to support cancer cell colonisation and progression, while also facilitating the development of cancer therapy (219, 228).

In conclusion, this chapter showed that differential expression levels of TGF-β and BMP signalling pathways were observed across CRC cell lines. While a high activation of the upstream signalling axis was present in all cells, increased phosphorylation of SMAD1/5/8 was noted, contrasting with the absence of SMAD4 protein, which was detected in a metastatic CRC cell line. It was suggested that the upregulation of TGF-β and BMP signalling pathways in the BM cells maintains homeostasis in the BM niche. A co-culture model between metastatic CRC and BM cells would be beneficial for future investigations.

Chapter 4 SMAD4 expression in colorectal adenomatous polyps and its correlation with clinical outcomes

4.1 Introduction

Adenomatous polyps represent the majority of premalignant lesions that can develop into CRC (4-6). Colonoscopy, a standard screening method for CRC, has the capability to remove premalignant lesions and assist in preventing the onset of CRC (40, 46). High-risk patients with advanced colorectal polyps, as outlined in the BSG2020 guideline, such as adenomatous polyps measuring at least 10 mm, those with high-grade dysplasia, or those presenting five or more premalignant polyps, are advised to undergo surveillance colonoscopy at regular intervals. This guidelines can significantly predict a higher rate of overall metachronous polyps in high-risk patients by using the characteristics of the index polyps as an indicator (6). However, almost half of the patients who undergo surveillance colonoscopy do not presented with metachronous lesions. Less than 1% of patients in a recent study developed CRC, indicating that many patients may be being over-monitored by this invasive procedure (49).

In terms of advancements in molecular biology technologies, identifying tumour biomarkers and prognostic factors aims to distinguish specific tumour characteristics. This provides greater benefits in clinical management, enhances precision medicine, and guides future research directions. Assessing the presence of proteins within pathological tissues using IHC technology, which illustrates their role in pathogenesis, is valuable in clinical applications such as characterising subtypes, offering prognostic information, stratifying patients for treatment, and predicting therapeutic responses (229). Additionally, transcriptomics analysis is suggested to provide essential information for cancer subtypes and reflect key elements of tumour heterogeneity (230). Combining IHC and transcriptomics analyses offers an alternative approach to clarifying tumour heterogeneity during tumourigenesis, providing improved accuracy and reliability (229, 230). Therefore, utilising these approaches earlier in the identification of colorectal adenomatous polyps offers additional benefits in histopathological applications and may influence clinical management due to the current limitations of prognostic factors in premalignant lesions.

The accumulation of mutations and dysregulation within the cell signalling network enables the transformation of adenomatous polyps into malignant tumours via the adenoma-carcinoma pathway (91). *SMAD4* is a key gene identified as mutated during CRC carcinogenesis, promoting tumour growth and progression (2, 3, 19-21). A germline mutation of *SMAD4* has been identified in JPS (231-235). Additionally, somatic inactivation of *SMAD4* has been observed in CRC, pancreatic cancer, and other cancer types. These defects lead to a loss of epithelial SMAD4 expression, which can be detected through IHC analysis, providing a valuable tool for pathological assessment (232, 236-238). Therefore, a more comprehensive understanding of the clinically relevant dysregulation of SMAD4 expression phenotypes is essential for further research, particularly in larger cohorts across CRC tumourigenesis. This chapter aims to investigate SMAD4 expression in colorectal adenomatous polyps through IHC analysis and its relationship with clinical and pathological data, particularly across histological subtypes, as well as to conduct transcriptomic analysis of these premalignant lesions in prognostically relevant subgroups.

4.2 SMAD4 antibody specificity in CRC cell lines

The specificity of the SMAD4 antibody was examined and optimised on CRC cell lines and control samples using Western blot and IHC assays, as described in Section 2.2.2.1. Consequently, Western blot results performed on recombinant proteins demonstrated that the SMAD4 antibody bound to the target SMAD4 protein at the predicted size of 70 KDa (Figure 4-1A). Additionally, the Western blot results of the CRC cell lines showed that HCT116 and DLD1 cells exhibited higher SMAD4 expression, followed by SW837 cells. In contrast, SMAD4 expression was not detected in SW620 and COLO205 cells (Figure 4-1B). Furthermore, the IHC staining results showed that SW480 and SW620 exhibited negative staining for SMAD4, whereas DLD1 and HeLa expressed SMAD4 as a brown colour in the cytoplasm and nuclei (Figure 4-1C).

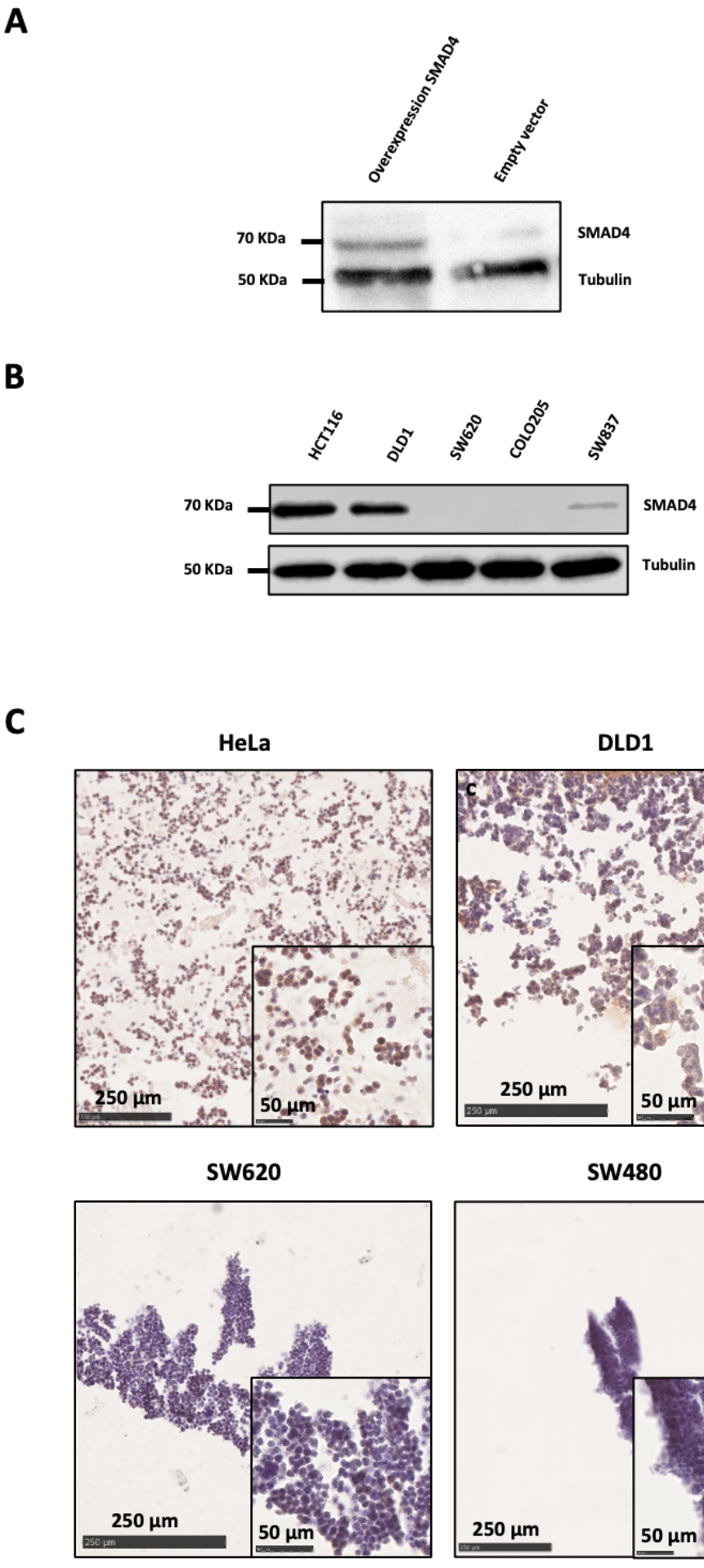

A
Overexpression SMAD4
Empty vector
70 KDa
50 KDa
SMAD4
Tubulin
B
HCT116
DLD1
SW620
COLO205
SW837
70 KDa
50 KDa
SMAD4
Tubulin
C
HeLa
DLD1
250 µm
50 µm
250 µm
50 µm
SW620
SW480
250 µm
50 µm
250 µm
50 µm

Figure 4-1 Specificity of SMAD4 antibody in CRC cell lines

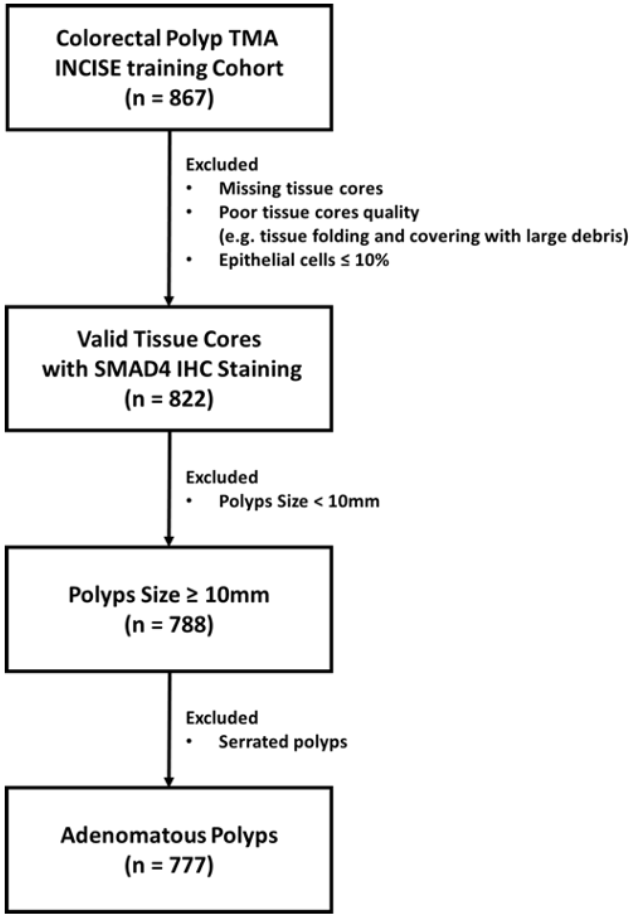
Western blot results showed antibody specificity in recombinant proteins (A). SMAD4 expression was detected in HCT116, DLD1, and SW837 cells. Conversely, SMAD4 protein was absent in SW620 and COLO205 cells (B). The IHC results showed that HeLa and DLD1 cells expressed SMAD4. In contrast, no SMAD4 expression was observed in SW620 and SW480 (C). The scale bar = 50 and 200 μm in Figure C.

4.3 Clinicopathological characteristics of the INCISE cohort

Subsequently, to assess the prognostic significance of SMAD4 expression in patients with pre-cancerous colorectal adenomas, tissue microarrays (TMAs) from the INCISE training cohort (n = 867) were stained for SMAD4 protein, as described in Section 2.2.2.2. After excluding samples based on tissue quality control criteria, tissue cores from 822 patients remained. Larger polyps measuring 10 mm, representing 788 cases, excluded the serrated cases and were selected for analysing of adenomatous polyps (n = 777 cases) to evaluate the prognostic significance, as described in Section 2.2.2.5 (Figure 4-2A).

The characteristics of patients in this study showed that 71% of the patients are male (n = 552). Patients aged over 65 years accounted for 38.9% (n = 302). 12.5% of polyps were detected in the right colon (n = 97), 72.8% in the left colon (n = 565), and 14.7% in the rectum (n = 114). Adenoma with high dysplasia accounted for 16.2% (n = 126). 33.8% of patients had one polyp at the initial colonoscopy (n = 263), 53.8% had 2-4 polyps (n = 418), and 12.4% had at least five polyps (n = 96). 44.5% of patients had polyps with tubular histology (n = 346), 48.4% had tubulovillous histology (n = 376), and 7.1% had villous histology (n = 55). 65.9% of patients were stratified as high risk according to the BSG2020 guidelines (n = 512), and 55.2% had metachronous polyps and CRC detected at follow-up (n = 429) (Figure 4-2B).

A


Colorectal Polyp TMA
INCISE training Cohort
(n = 867)
Excluded
• Missing tissue cores
• Poor tissue cores quality
(e.g. tissue folding and covering with large debris)
• Epithelial cells ≤ 10%
Valid Tissue Cores
with SMAD4 IHC Staining
(n = 822)
Excluded
• Polyps Size < 10mm
Polyps Size ≥ 10mm
(n = 788)
Excluded
• Serrated polyps
Adenomatous Polyps
(n = 777)

B

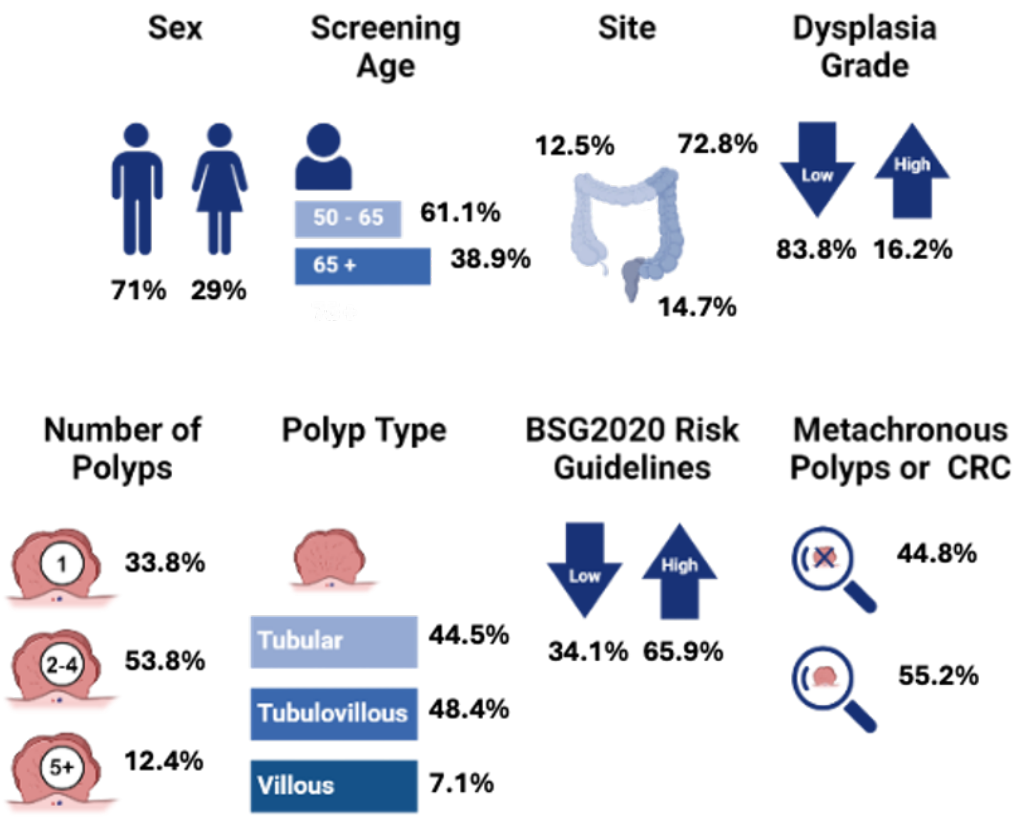

Sex
71%
29%
Screening Age
50 - 65
61.1%
65 +
38.9%
Site
12.5%
72.8%
14.7%
Dysplasia Grade
Low
High
83.8%
16.2%
Number of Polyps
1
33.8%
2-4
53.8%
5+
12.4%
Polyp Type
Tubular
44.5%
Tubulovillous
48.4%
Villous
7.1%
BSG2020 Risk Guidelines
Low
High
34.1%
65.9%
Metachronous Polyps or CRC
44.8%
55.2%

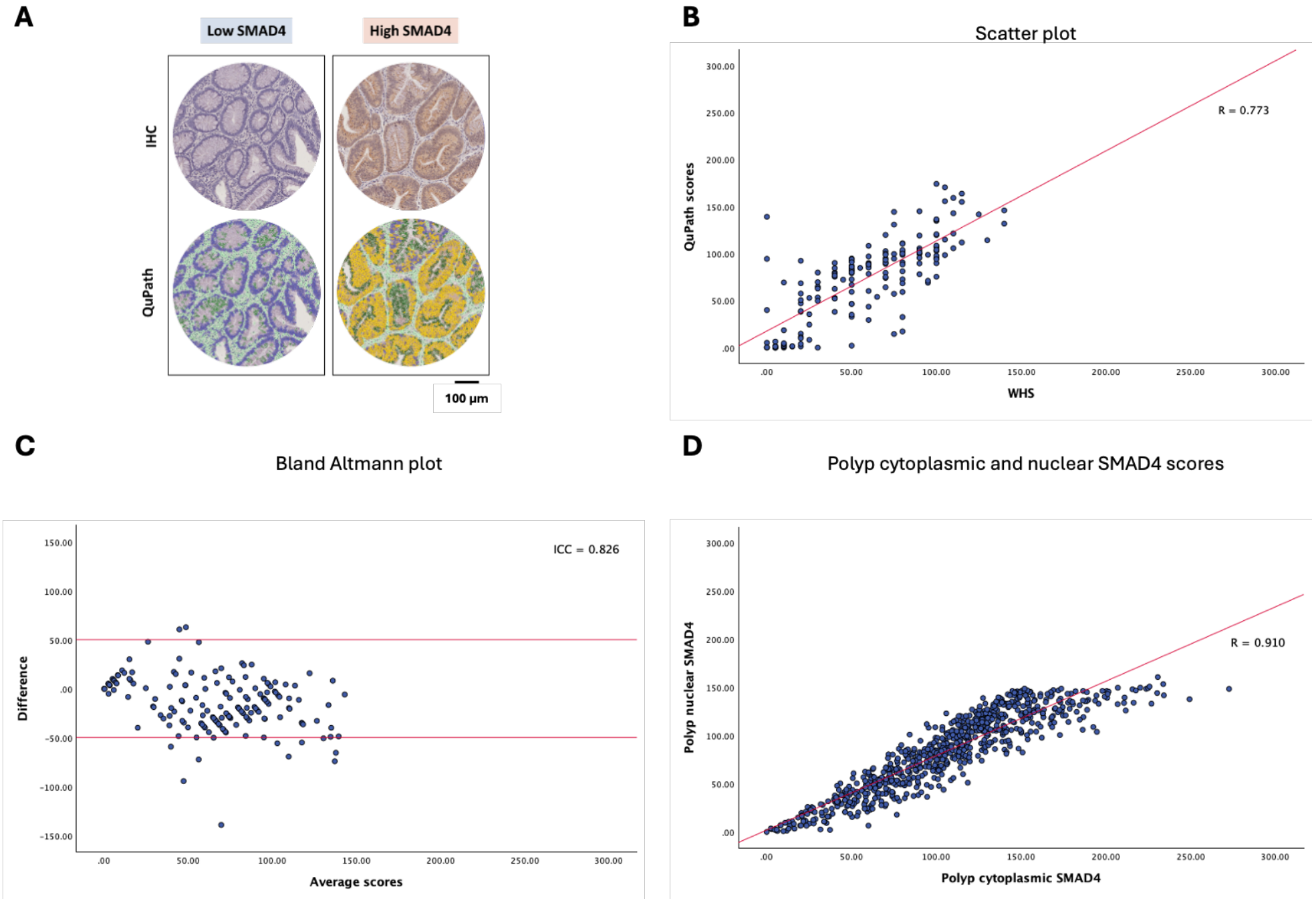
Figure 4-2 Colorectal polyps and clinicopathological features of the INCISE cohort

The Consort diagram shows how colorectal polyp samples were included in this study (A). The tissue microarrays (TMAs) of colorectal polyps from the INCISE training cohort were stained for SMAD4 expression using Immunohistochemistry (IHC). Only adenomatous polyps measuring ≥ 10 mm were analysed in relation to clinical outcomes. The demographic chart displays the characteristics of this cohort (B).

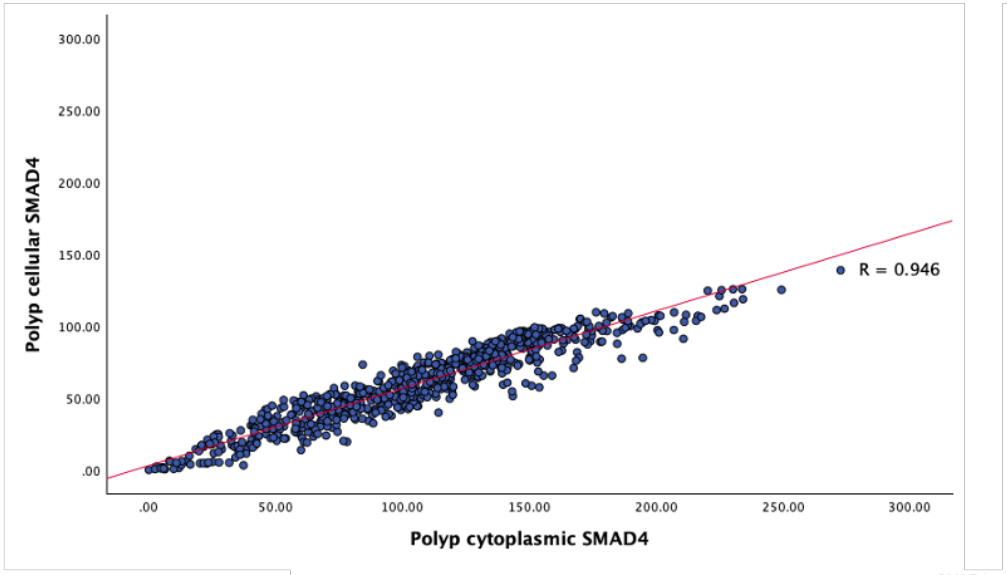
4.4 Expression of SMAD4 in colorectal adenomatous polyps and its association with clinical outcomes

After IHC staining for SMAD4 expression was performed, the staining intensity of SMAD4 in both cytoplasmic and nuclear compartments of epithelial cells was evaluated using manual WHS method on 10% of TMA cores, as described in Section 2.2.2.3. In parallel, digital weighted histoscoring was also assessed for all TMA cores using QuPath in this study, as described in Section 2.2.2.4. Representative IHC staining patterns of low and high SMAD4 in colorectal tissues showed an intense brown colour, indicating high SMAD4 expression, whereas an absent or weak brown colour was observed in tissues with low SMAD4 levels. Additionally, cell classification by QuPath were masked according to intensity and low SMAD4 expression in epithelial cells was represented by blue, whilst higher SMAD4 expression was indicated by yellow (Figure 4-3A).

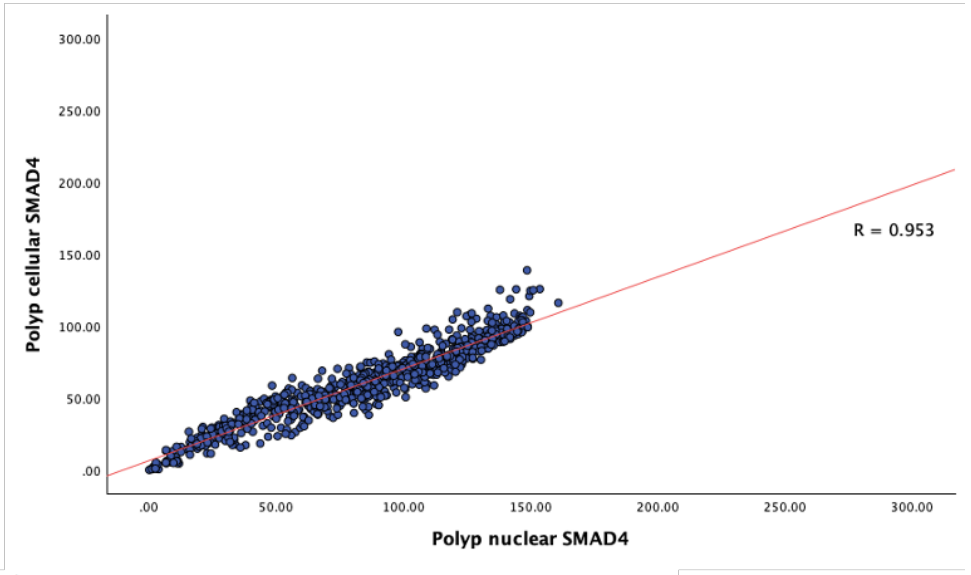
The QuPath score for SMAD4 expression in polyps was validated as a reliability measure with the WHS. The scatter plot comparing WHS and QuPath scores showed a positive correlation, with a correlation coefficient of 0.773 (Figure 4-3B). Additionally, the Bland-Altman plot indicated no significant difference between the two scores, confirming a strong positive correlation with an ICC of 0.826 (Figure 4-3C). Subsequently, average QuPath scores from basal and luminal polyps of 777 cases were obtained and assessed for the correlation between cytoplasmic, nuclear, and cellular scores of SMAD4 expression. A strong correlation was observed among these three scores (including cytoplasm and nucleus, R = 0.910; cytoplasm and cellular, R = 0.946; nucleus and cellular, R = 0.953; Figure 4-3D to F, respectively). Therefore, the cellular SMAD4 expression scores were used to determine cut points for stratifying patients with polyps via maximally selected rank statistics into low and high SMAD4 expression groups, with a cut point of 57.88 as shown in Figure 4-3G


A
Low SMAD4
High SMAD4
IHC
QuPath
100 µm
B
Scatter plot
R = 0.773
QuPath scores
WHS
C
Bland Altmann plot
ICC = 0.826
Difference
Average scores
D
Polyp cytoplasmic and nuclear SMAD4 scores
R = 0.910
Polyp nuclear SMAD4
Polyp cytoplasmic SMAD4

E


Polyp cytoplasmic and cellular SMAD4 scores
Polyp cellular SMAD4
Polyp cytoplasmic SMAD4
R = 0.946
300.00
250.00
200.00
150.00
100.00
50.00
.00

F


Polyp nuclear and cellular SMAD4 scores
Polyp cellular SMAD4
Polyp nuclear SMAD4
R = 0.953
300.00
250.00
200.00
150.00
100.00
50.00
.00

G

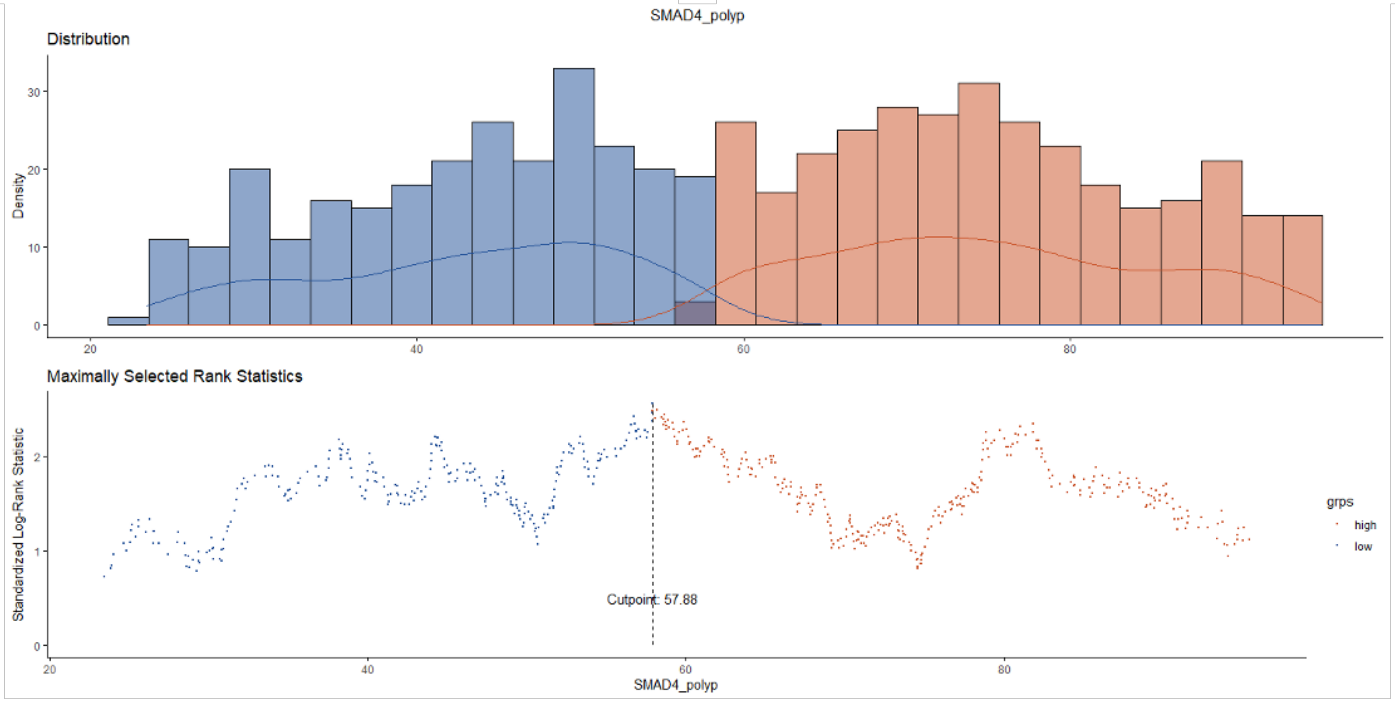

Distribution
Density
Maximally Selected Rank Statistics
Standardized Log-Rank Statistic
Cutpoint: 57.88
SMAD4_polyp
grps
high
low

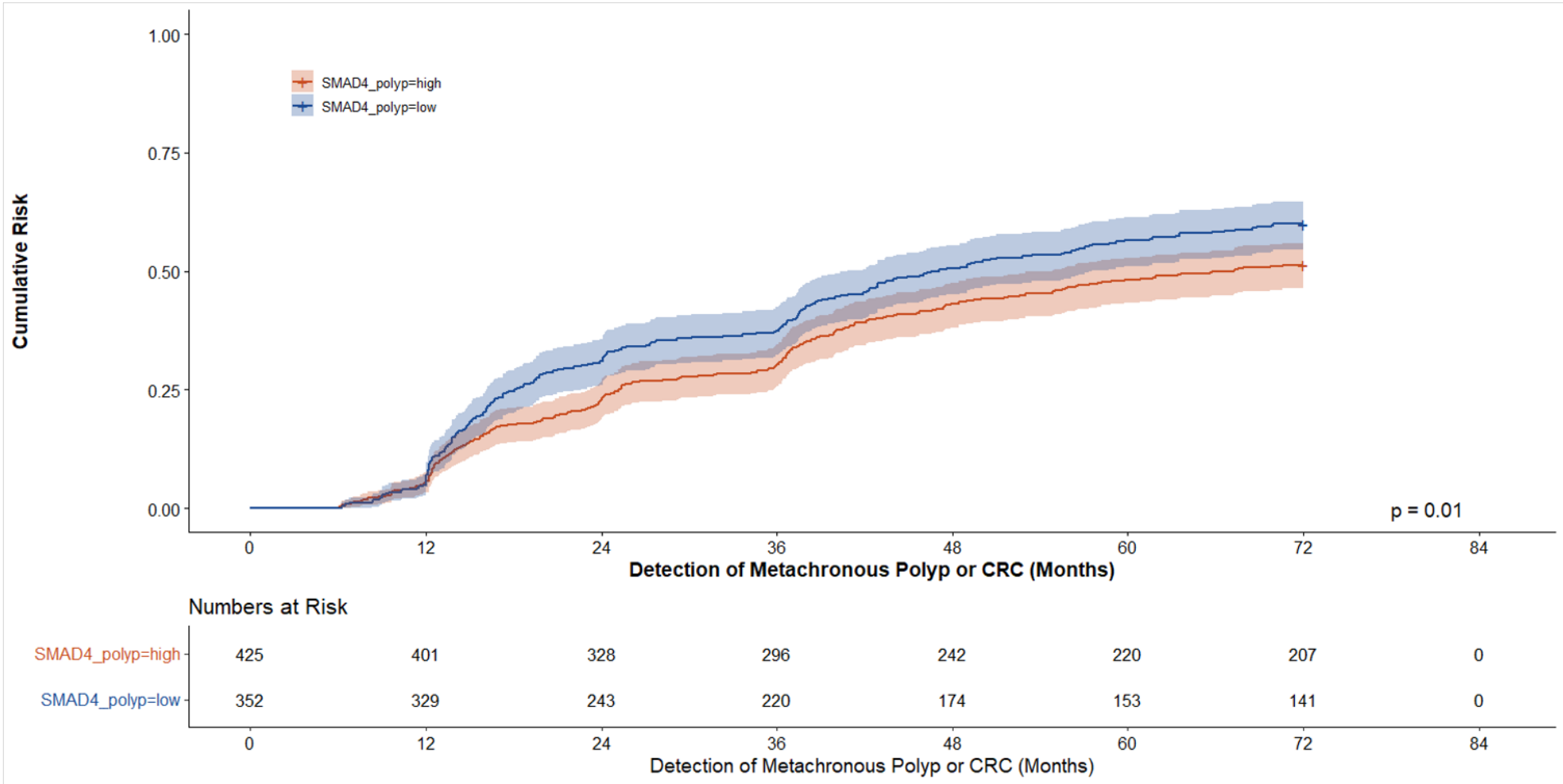
Figure 4-3 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in colorectal adenomatous polyps

Representative images of colorectal polyps showing low SMAD4 expression (left) and high SMAD4 expression (right) from IHC staining (top) and classified by QuPath (bottom) (A). A scatter plot of WHS and QuPath scores demonstrated a correlation between these two scores in detecting SMAD4 expression (R = 0.773) (B). A Bland-Altman plot indicated no significant difference between the scores and showed a positive correlation (ICC = 0.826) (C). Scatter plots of QuPath scores revealed a strong correlation between cytoplasmic and nuclear scores (R = 0.910) (D), cytoplasmic and cellular scores (R = 0.946) (E), and nuclear and cellular scores (R = 0.953) (F). The histogram of an optimal cut point was determined by the survminer package using maximally selected rank statistics in R, as shown at 57.88 from this colorectal polyp cohort (G). The scale bar = 100 μm in Figures A.

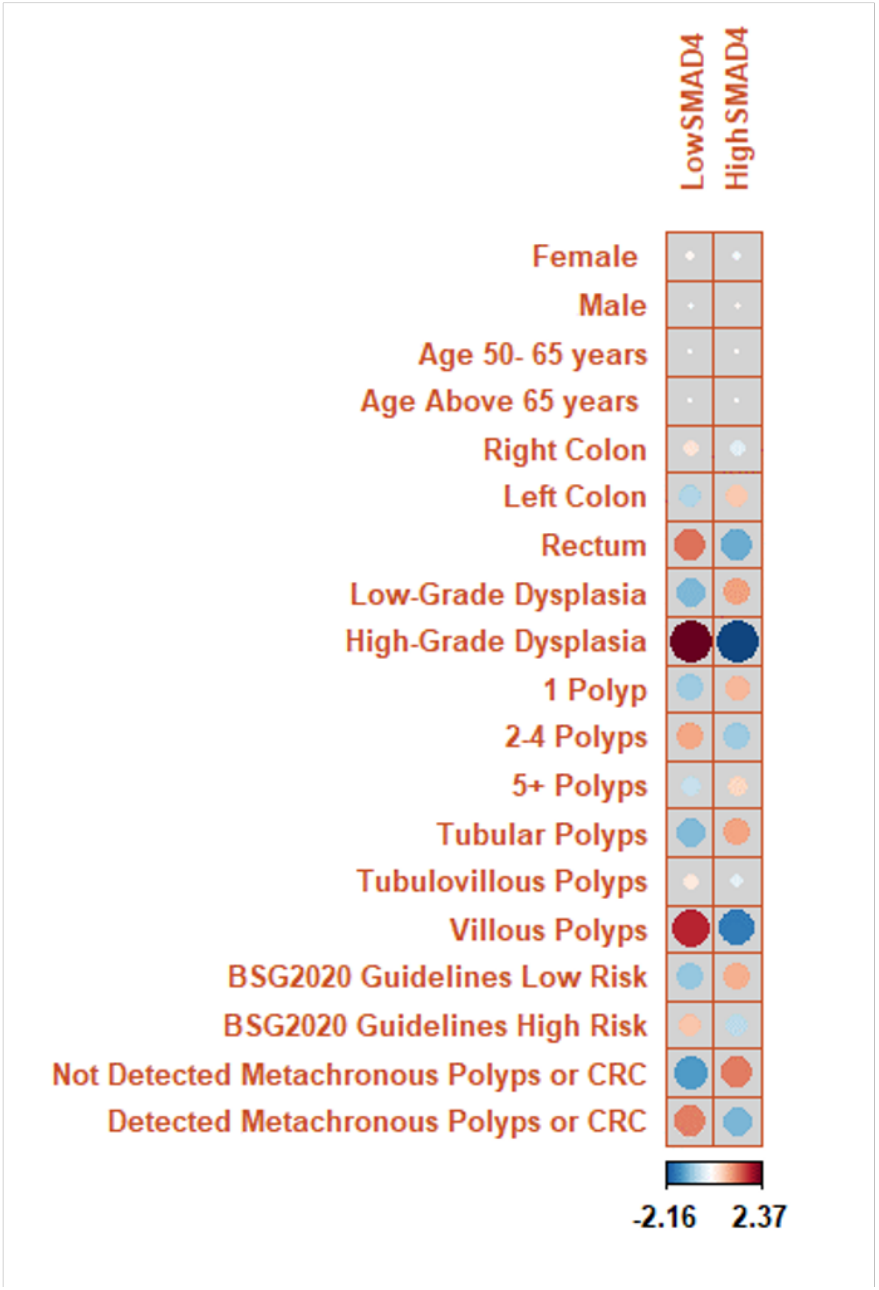
Patients with colorectal adenomatous polyps were divided into two groups based on SMAD4 expression levels, consisting of 352 cases with low SMAD4 expression and 425 cases with high SMAD4 expression. Kaplan–Meier (KM) survival analysis was conducted to evaluate the relationship between SMAD4 expression status and the time until detection of metachronous polyps or CRC development following the initial colonoscopy. Overall, the results indicated that low SMAD4 expression in the epithelial cells of the polyps was significantly associated with an increased rate of metachronous polyp or CRC development (Log-rank, p = 0.01) (Figure 4-4A). The interval from polyp detection index polyp to metachronous polyp detection was predicted to be shorter in patients with low SMAD4 expression, approximately 45.9 (43.3-48.5) months, compared to 50.5 (48.2-52.8) months in patients with high SMAD4 expression. Subsequently, SMAD4 expression in polyps and its association with clinicopathological features were analysed. The Chi-squared table results demonstrated that low SMAD4 expression was linked to higher dysplasia grades (p < 0.001), histological subtypes (p = 0.018), and the occurrence of metachronous polyps or CRC development (p = 0.026) (Figure 4-4B and Table 4-1). Furthermore, we compared levels of SMAD4 expression between low- and high-risk groups. No difference in expression was observed in this premalignant cohort (Wilcoxon test, p = 0.11) (Figure 4-4C). Cox regression analysis showed sex, age, site of polyps, number of polyps, BSG 2020 risk score, detection of metachronous polyps or CRC status, and SMAD4 expression were significant in univariate analysis (p < 0.05). High SMAD4 expression was independently linked to a lower risk of developing metachronous polyps or CRC (HR = 0.757, 95% CI = 0.627-0.916, p = 0.004). Multivariate analysis revealed that, unlike other independent prognostic factors, sex was a significant predictor of

outcomes, with male patients having a higher risk in this cohort (HR = 1.474, 95% CI = 1.173-1.852, $p < 0.001$) (Table 4-1).

A


1.00
0.75
0.50
0.25
0.00
Cumulative Risk
SMAD4_polyp=high
SMAD4_polyp=low
p = 0.01
0
12
24
36
48
60
72
84
Detection of Metachronous Polyp or CRC (Months)
Numbers at Risk
SMAD4_polyp=high 425 401 328 296 242 220 207 0
SMAD4_polyp=low 352 329 243 220 174 153 141 0
0
12
24
36
48
60
72
84
Detection of Metachronous Polyp or CRC (Months)

B


LowSMAD4
HighSMAD4
Female
Male
Age 50- 65 years
Age Above 65 years
Right Colon
Left Colon
Rectum
Low-Grade Dysplasia
High-Grade Dysplasia
1 Polyp
2-4 Polyps
5+ Polyps
Tubular Polyps
Tubulovillous Polyps
Villous Polyps
BSG2020 Guidelines Low Risk
BSG2020 Guidelines High Risk
Not Detected Metachronous Polyps or CRC
Detected Metachronous Polyps or CRC
-2.16
2.37

C

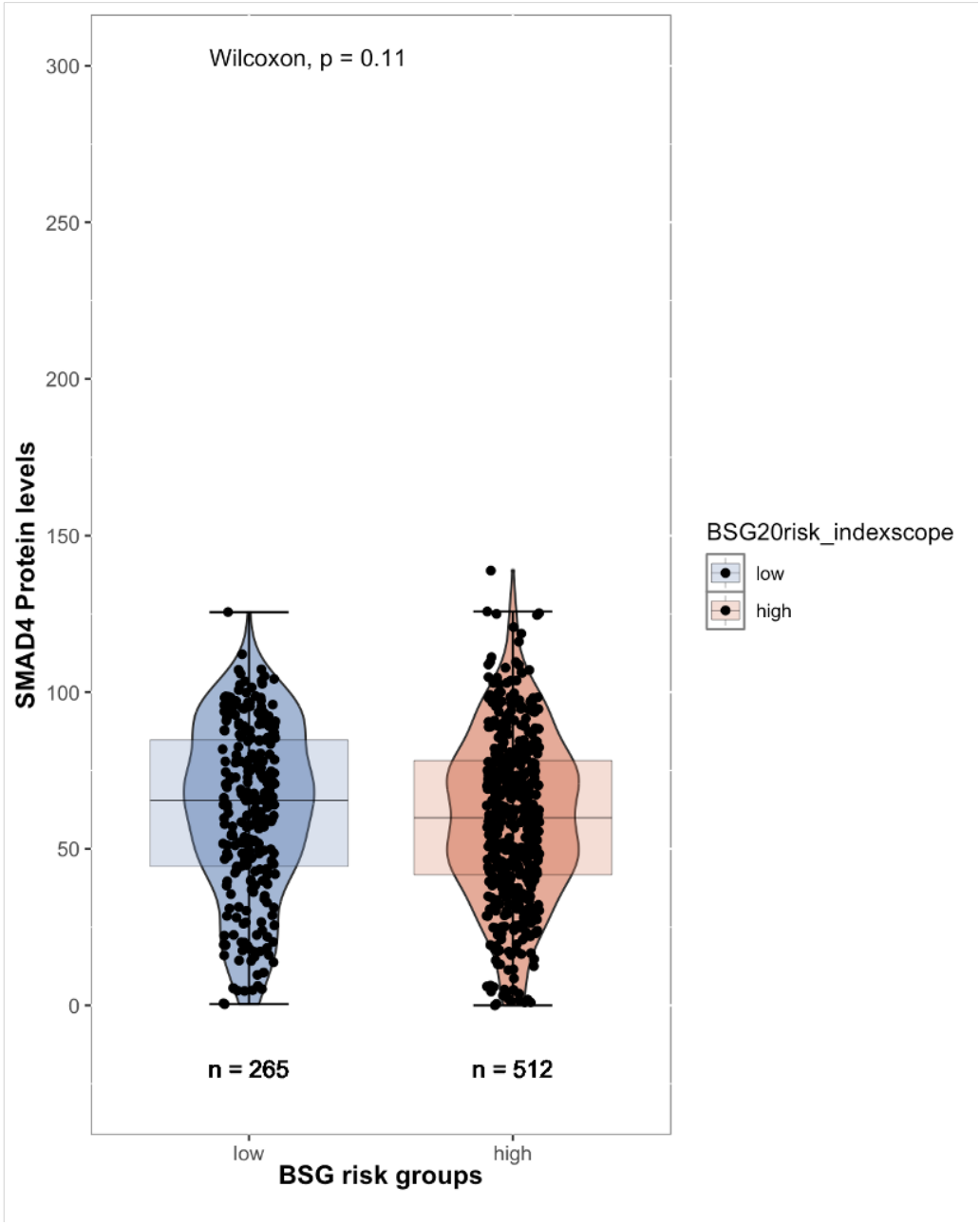

Wilcoxon, p = 0.11
SMAD4 Protein levels
300
250
200
150
100
50
0
n = 265
n = 512
low
high
BSG risk groups
BSG20risk_indexscope
low
high

Figure 4-4 SMAD4 expression in relation to the detection of metachronous polyps or CRC development and clinical outcomes

The Kaplan-Meier curve for SMAD4 expression in colorectal adenomatous polyps showed that low levels of SMAD4 are associated with a higher rate of developing metachronous polyps or CRC (Log-rank, p = 0.01) (A). The CORR plot illustrated the relationship between SMAD4 expression levels and the clinicopathological characteristics of patients with colorectal polyps (B). A boxplot revealed no significant difference in SMAD4 expression levels between low- and high-BSG risk groups (Wilcoxon test, p = 0.11) (C).

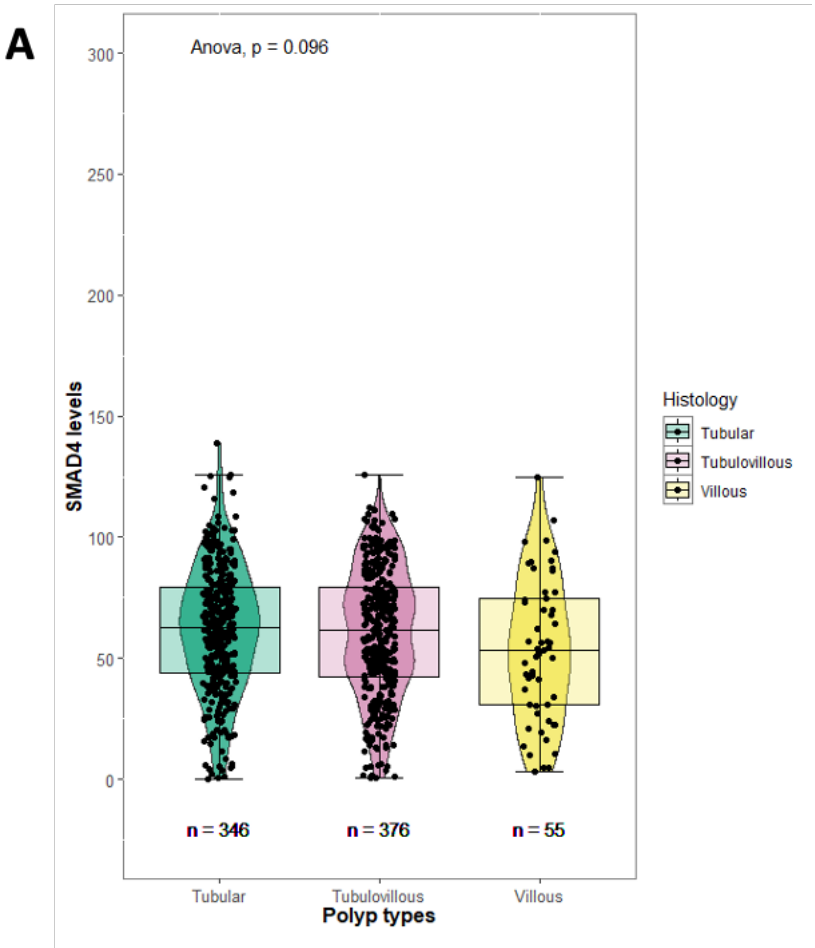
Table 4-1 The relationship of SMAD4 expression in colorectal adenomatous polyps and clinicopathological characteristics

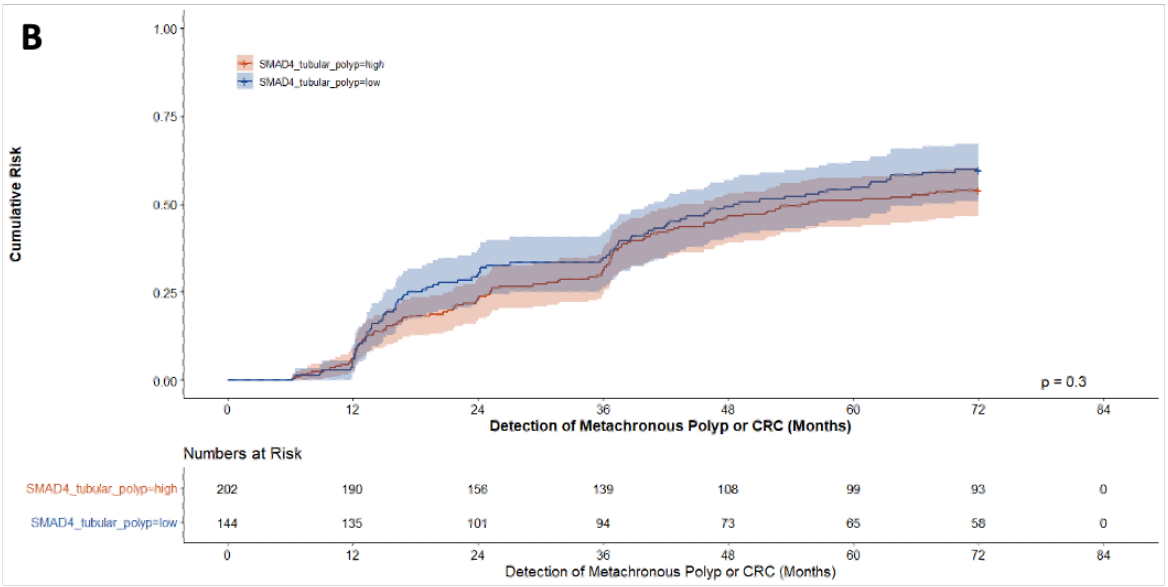
	SMAD4 expression in polyps			Univariate analysis		Multivariate analysis	
	Low n (%)	High n (%)	*p*	HR (95% CI)	*p*	HR (95% CI)	*p*
Sex							
Female	103 (29.3%)	122 (28.7%)	0.865	1		1	
Male	249 (70.7%)	303 (71.3%)		1.800 (1.437-2.254)	<0.001	1.474 (1.173-1.852)	<0.001
Age							
50- 65 years	215 (61.1%)	260 (61.2%)	0.978	1		1	
≥ 65 years	137 (38.9%)	165 (38.8%)		1.334 (1.101-1.617)	0.003	1.161 (0.954-1.412)	0.136
Site of most advanced polyp							
Right colon	46 (13.1%)	51 (12.0%)	0.128	1		1	
Left colon	245 (69.6%)	320 (75.5%)		0.587 (0.450-768)	<0.001	0.777 (0.594-1.017)	0.066
Rectum	61 (17.3%)	53 (12.5%)		0.913 (0.652-1.277)	0.593	1.158 (0.826-1.625)	0.394
Dysplasia grade							
Low grade	277 (78.7%)	374 (88.0%)	<0.001	1		-	-
High grade	75 (21.3%)	51 (12.0%)		1.166 (0.905-1.502)	0.235	-	-
Number of polyps							
1 polyp	110 (31.3%)	153 (36.0%)	0.188	1		1	
2-4 polyps	202 (57.4%)	216 (50.8%)		1.636 (1.310-2.042)	<0.001	0.418 (0.057-3.043)	0.389
≥ 5 polyps	40 (11.4%)	56 (13.2%)		2.893 (2.142-3.906)	<0.001	0.576 (0.078-4.270)	0.590
Histology							
Tubular polyps	144 (40.9%)	202 (47.5%)	0.018	1		-	-
Villous polyps	34 (9.7%)	21 (4.9%)		0.939 (0.635-1.387)	0.752	-	-
Tubulovillous polyps	174 (49.4%)	202 (47.5%)		1.012 (0.832-1.231)	0.904	-	-
BSG2020 risk score							
Low BSG risk	110 (31.3%)	155 (36.5%)	0.126	1		1	
High BSG risk	242 (68.8%)	270 (63.5%)		1.806 (1.458-2.235)	<0.001	2.994 (0.412-21.749)	0.278
Metachronous polyps or CRC detection							
No future polyp or CRC	141 (40.1%)	207 (48.7%)	0.026	1		1	
Developed a metachronous polyp or CRC	211 (59.9%)	218 (51.3%)		73.657 (33.775-160.630)	<0.001	205477.863 (0-6.276E+23))	0.573
SMAD4 expression in polyps							
Low SMAD4	-	-	-	1		1	
High SMAD4	-	-	-	0.757 (0.627-0.916)	0.004	0.907 (0.749-1.099)	0.321

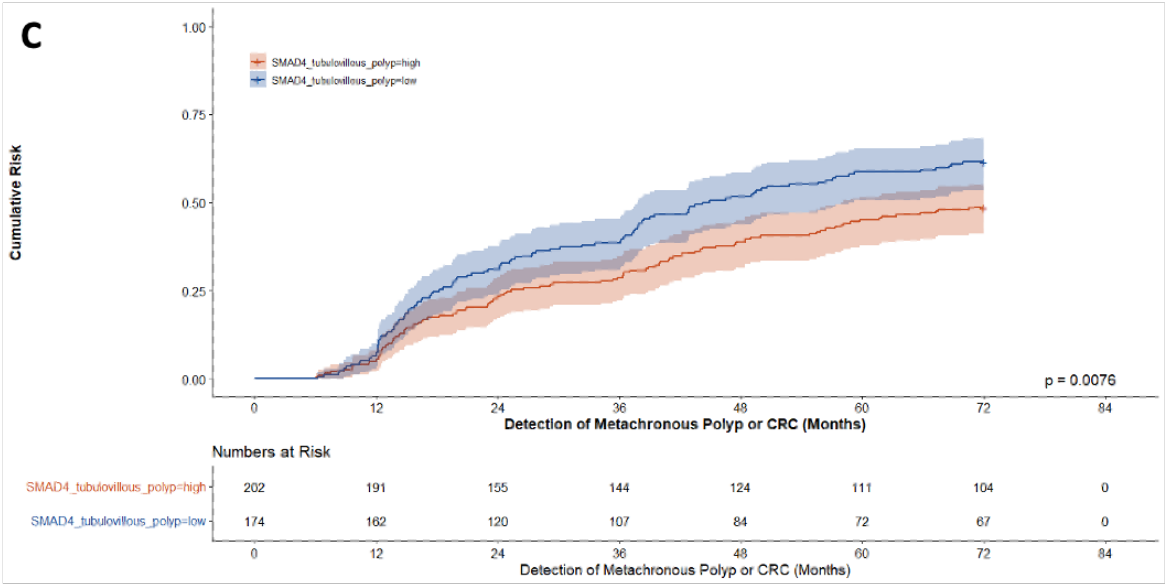
4.5 SMAD4 expression in tubulovillous polyps predicts the development of metachronous polyps and contributes to transcriptomic alterations.

4.5.1 Expression of SMAD4 in tubulovillous polyps and its association with clinical outcomes

Regarding the relationship between SMAD4 levels and the histological subtypes of colorectal adenomatous polyps, we further stratified patients according to the histological subtypes of their polyps: tubular (n = 346), tubulovillous (n = 376), and villous polyps (n = 55). When comparing SMAD4 expression levels across these subtypes, the results showed no significant differences among the three groups (ANOVA, p = 0.096) (Figure 4-5A). Nonetheless, KM survival analysis was conducted for each histological subtype and indicated that low SMAD4 expression in tubulovillous polyps was linked to a higher incidence of metachronous polyp or CRC development (Log-rank, p = 0.0076) (Figure 4-5C). In contrast, a non-significant association was observed in polyps with tubular (Log-rank, p = 0.3) (Figure 4-5B) or villous (Log-rank, p = 0.82) (Figure 4-5D) histology.




B
SMAD4_tubular_polyp=high
SMAD4_tubular_polyp=low
Cumulative Risk
1.00
0.75
0.50
0.25
0.00
p = 0.3
0
12
24
36
48
60
72
84
Detection of Metachronous Polyp or CRC (Months)
Numbers at Risk
SMAD4_tubular_polyp=high 202 190 158 139 108 99 93 0
SMAD4_tubular_polyp=low 144 135 101 94 73 65 58 0
Detection of Metachronous Polyp or CRC (Months)


C
SMAD4_tubulovillous_polyp=high
SMAD4_tubulovillous_polyp=low
Cumulative Risk
1.00
0.75
0.50
0.25
0.00
p = 0.0076
0
12
24
36
48
60
72
84
Detection of Metachronous Polyp or CRC (Months)
Numbers at Risk
SMAD4_tubulovillous_polyp=high 202 191 155 144 124 111 104 0
SMAD4_tubulovillous_polyp=low 174 162 120 107 84 72 67 0
Detection of Metachronous Polyp or CRC (Months)

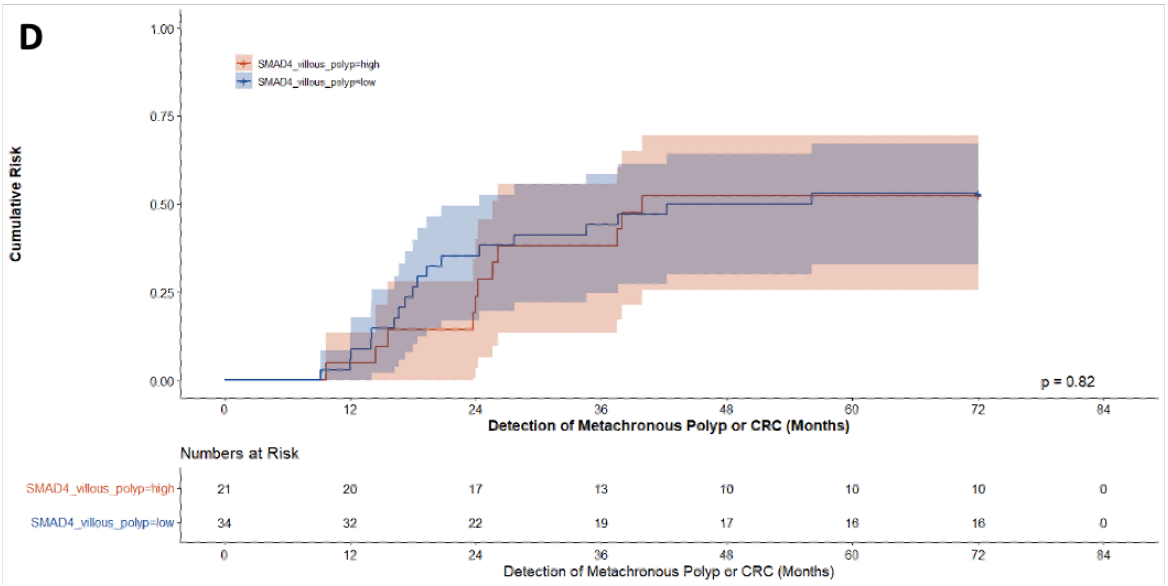

D
SMAD4_villous_polyp=high
SMAD4_villous_polyp=low
Cumulative Risk
1.00
0.75
0.50
0.25
0.00
p = 0.82
0
12
24
36
48
60
72
84
Detection of Metachronous Polyp or CRC (Months)
Numbers at Risk
SMAD4_villous_polyp=high 21 20 17 13 10 10 10 0
SMAD4_villous_polyp=low 34 32 22 19 17 16 16 0
Detection of Metachronous Polyp or CRC (Months)

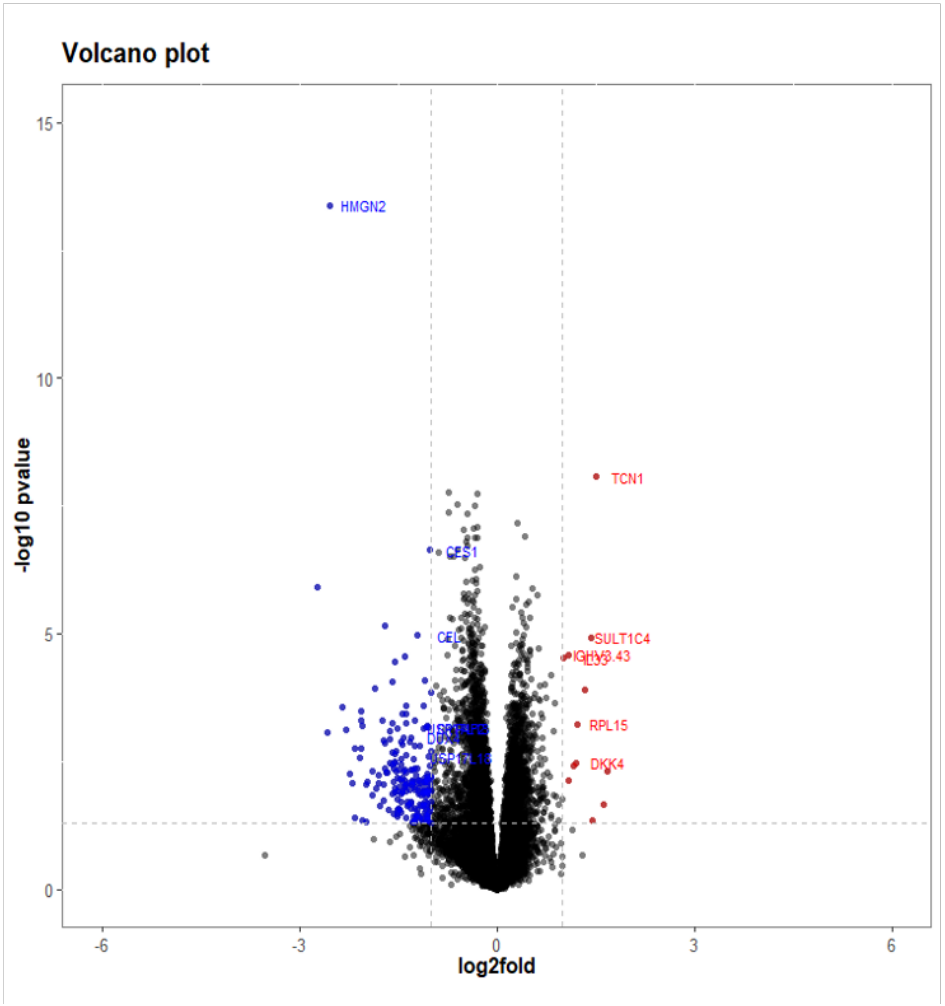
Figure 4-5 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal adenomatous polyps, stratified by histological subtypes

A boxplot showed no significant difference in SMAD4 expression levels among the three histological subtypes (ANOVA, p = 0.096) (A). Kaplan-Meier curve of SMAD4 expression in tubular polyps showed that low SMAD4 expression was not associated with an increasing rate of metachronous polyp development (Log-rank, p = 0.3) (B). Kaplan-Meier curve of SMAD4 expression in tubulovillous polyps showed that low SMAD4 expression was associated with an increasing rate of metachronous polyp development (Log-rank, p = 0.0076) (C). Kaplan-Meier curve of SMAD4 expression in villous polyps showed that low SMAD4 expression was not associated with an increasing rate of metachronous polyp development (Log-rank, p = 0.82) (D).

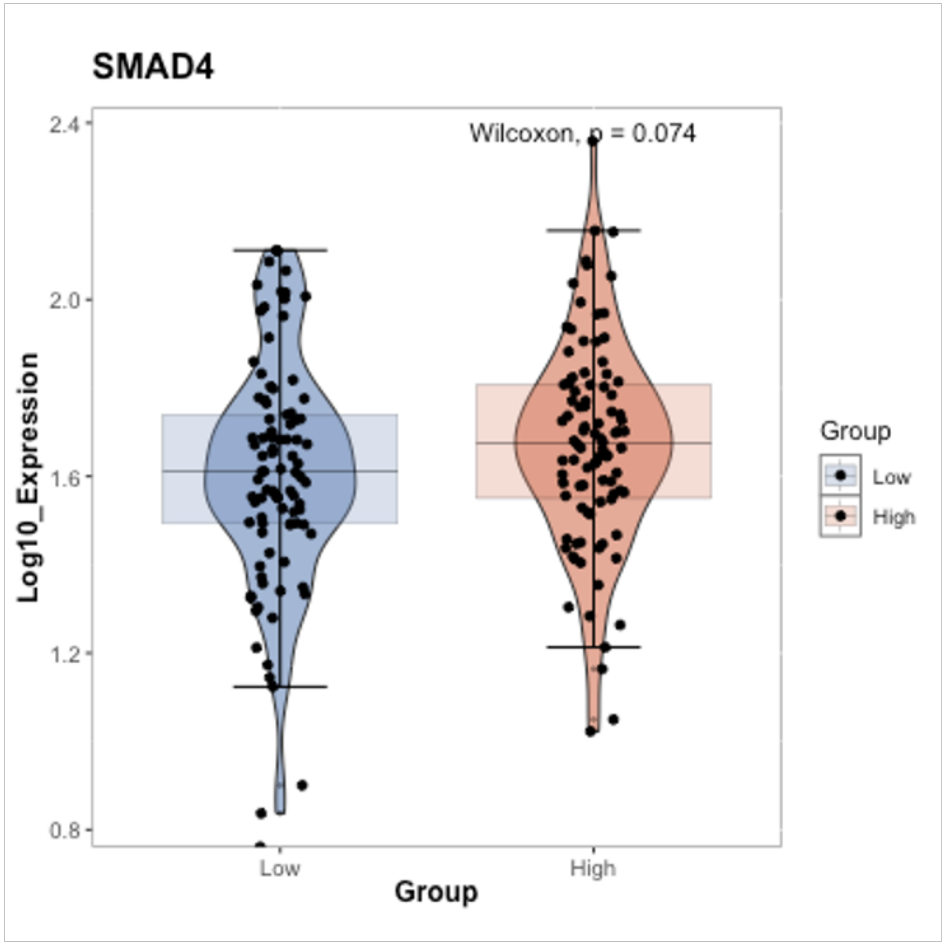
4.5.2 Transcriptomic analysis of low and high SMAD4 expression in tubulovillous polyps

As previous results suggested that decreased SMAD4 expression in premalignant adenomatous polyps is associated with a higher likelihood of metachronous polyp development, especially in patients with index tubulovillous polyps, we initially selected tubulovillous polyp cases from the first and fourth quartiles, representing samples with very low and high SMAD4 expression, respectively. These were used to compare differences in transcriptomic analysis, as described in Section 2.2.2.6. There were 93 cases in the low SMAD4 group and 94 cases in the high SMAD4 group. The volcano plot illustrated differential gene expression between the two groups. Twelve genes were upregulated in the high SMAD4 polyps, while 196 genes were downregulated ($p \leq 0.05$) (Figure 4-6A). Some of the upregulated genes in the low SMAD4 polyps are involved in protein deubiquitination. Levels of SMAD4 mRNA expression were compared between low and high SMAD4 protein-expressing polyps; however, no significant difference was observed (Wilcoxon test, p = 0.074) (Figure 4-6B). GSEA analysis was performed, revealing that no gene sets were significantly enriched at a False Discovery Rate (FDR or q-val) < 25%. However, five gene sets showed enrichment in tyrosine metabolism (Normalised Enrichment Score (NES) = -1.705, nominal p-value = 0.008), PPAR signalling (NES = -1.607, nominal p-value = 0.022), Arginine and Proline metabolism (NES = -1.566, nominal p-value = 0.034), Leukocyte transendothelial migration (NES = -1.474, nominal p-value = 0.036), and Basal cell carcinoma (NES = -1.451, nominal p-value = 0.044) (Figure 4-6C), using Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway as a reference. These results demonstrate that some signalling pathways are dysregulated in tubulovillous polyps with low SMAD4 expression compared to those with high SMAD4 expression.

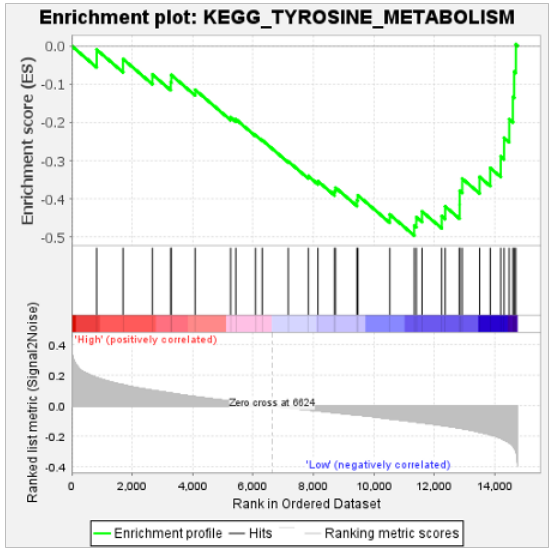
A


Volcano plot
-log10 pvalue
log2fold
15
10
5
0
-6
-3
0
3
6
HMGN2
TCN1
CES1
SULT1C4
RPL15
DKK4

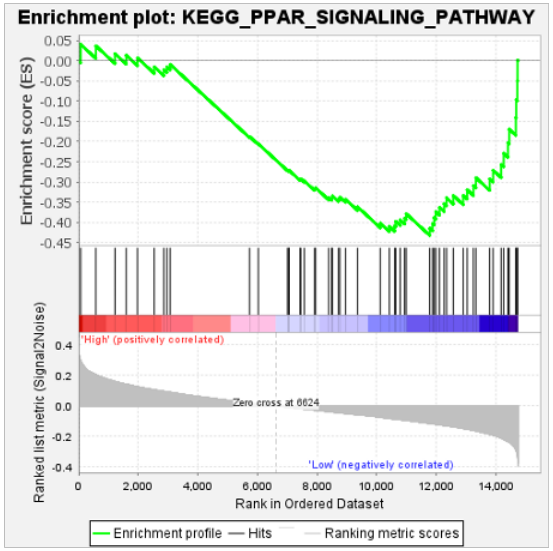
B


SMAD4
Wilcoxon, p = 0.074
Log10_Expression
2.4
2.0
1.6
1.2
0.8
Low
High
Group
Group
Low
High

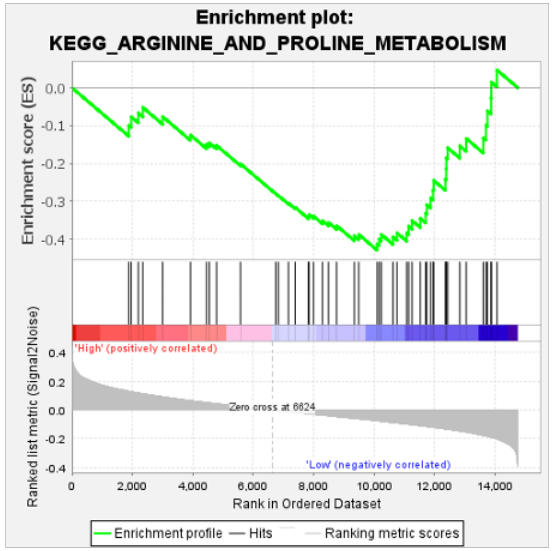
C



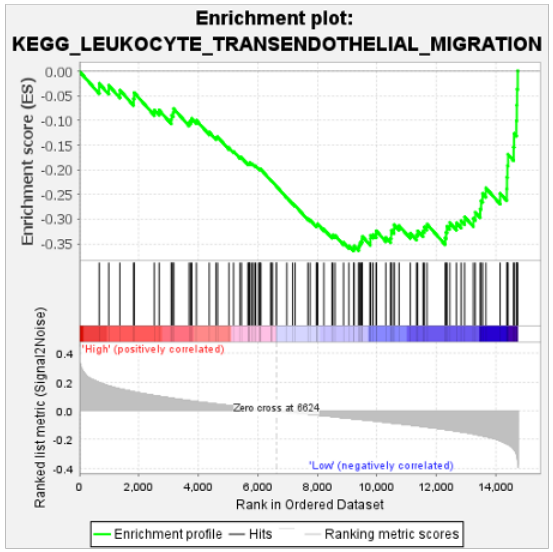
NES = -1.7047154
NOM p-value = 0.00814664



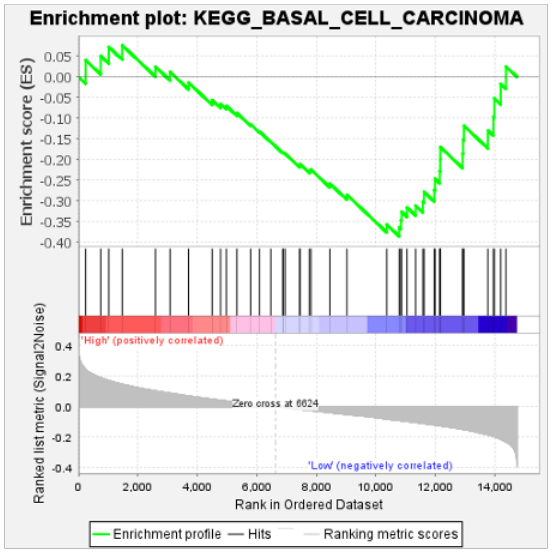
NES = -1.6068149
NOM p-value = 0.02244898



NES = -1.5663675
NOM p-value = 0.034343433



NES = -1.4741925
NOM p-value = 0.036144577



NES = -1.4506965
NOM p-value = 0.044088177

Figure 4-6 Transcriptomic analysis comparing low and high SMAD4 expression in tubulovillous polyps

A volcano plot displayed differential gene expression between the low SMAD4 group (n = 93) and the high SMAD4 group (n = 94), with 196 genes downregulated and 12 genes upregulated in the high SMAD4 group ($p \leq 0.05$) (A). A boxplot comparing SMAD4 gene expression levels between the low- and high-SMAD4 protein groups showed no significant difference in RNA levels (Wilcoxon test, $p = 0.074$) (B). GSEA analysis plots showed that gene sets including tyrosine metabolism (NES = -1.705, nominal p-value = 0.008), PPAR signalling (NES = -1.607, nominal p-value = 0.022), Arginine and Proline metabolism (NES = -1.566, nominal p-value = 0.034), Leukocyte transendothelial migration (NES = -1.474, nominal p-value = 0.036), and Basal cell carcinoma (NES = -1.451, nominal p-value = 0.044) (C), were enriched based on KEGG pathway reference; Normalised Enrichment Score (NES).

4.6 Discussion

Regarding the genetic mutation status of *SMAD4* in colorectal polyps associated with hereditary conditions, it is more frequently observed in JPS and is also linked to an increased risk of CRC (232, 235). Somatic loss of SMAD4 is also seen, indicating a reduction in epithelial SMAD4 expression in some patients with polyps, as identified by IHC (232). However, a better understanding of SMAD4 dysregulation in colorectal polyps and its pathological implications is needed, which could assist in clinical applications. In this chapter, the SMAD4 antibody demonstrated specificity in targeting the SMAD4 protein, as evidenced by its size and staining in both the cytoplasmic and nuclear compartments of epithelial cells, as observed via Western blot and IHC assays. The larger cohort of colorectal polyps, comprising 777 cases of adenomatous polyps, was examined for the prognostic relevance of SMAD4 expression and its clinical outcomes, including the detection of metachronous polyps or CRC development, as well as its relationship with clinical data.

Our findings demonstrated a reliable use of QuPath scores for evaluating epithelial SMAD4 expression in polyps. Consequently, patients with low SMAD4 expression in adenomatous polyps were linked to a higher incidence of developing metachronous polyps or CRC (Log-rank, $p = 0.01$). Although SMAD4 expression status was not associated with BSG2020 risk stratification, it has been shown that reduced SMAD4 expression correlates with higher dysplasia grades ($p < 0.001$), histological subtypes ($p = 0.018$), and the occurrence of metachronous polyps or CRC ($p = 0.016$). Therefore, this analysis could expand the clinical benefits of assessing SMAD4 expression in polyps, especially considering previous studies had limited number of patients, most of whom had JPS, and the focus was mainly on

genetic mutations (231-235). Additionally, this study is the first to investigate the prognostic significance of SMAD4 expression in adenomatous polyps, particularly regarding histological subtypes. Notably, the results showed that lower SMAD4 expression is associated with a higher likelihood of developing metachronous lesions in cases of tubulovillous polyps (Log-rank, p = 0.0076), whereas no significant differences were observed in tubular and villous polyps (Log-rank, p = 0.3 and 0.82, respectively).

Furthermore, transcriptomic analysis comparing low- and high-SMAD4 expression levels in tubulovillous polyps identified 208 genes with differential expression between the groups. Several genes involved in protein deubiquitination were upregulated in polyps with low SMAD4 protein levels. GSEA analysis revealed enrichment in tyrosine metabolism, PPAR signalling, arginine and proline metabolism, leukocyte transendothelial migration, and basal cell carcinoma in polyps with low SMAD4 expression, although these findings were not highly statistically significant. Nonetheless, these results implied that some mechanisms underlying early CRC tumourigenesis may be altered and linked to SMAD4 dysregulation in this adenomatous polyp subtype. A study of the transcriptome of 39 sessile serrated adenoma/polyps compared to their adjacent normal epithelial cells found that this polyp subtype is associated with MSI-H CRC and decreased *CDX2* expression (239). In comparison, the transcriptome profile of 32 paired colorectal adenomas and normal mucosa from another study indicated that the transformation process involves remodelling of the Wnt pathway, with over- and under-expression in 78 known components of this signalling cascade, and 19 Wnt targets apparently linked to the upregulation of *KIAA1199* (240). These may be involved in various mechanisms due to the heterogeneity observed in colorectal polyps.

Despite a more substantial implied clinical impact from this transcriptomic analysis with supervised clustering, which focused on comparisons of IHC-SMAD4 phenotypes in tubulovillous subtypes of colorectal polyps, it may limit biological insight and reflect fewer variations, along with the less aggressive nature of these lesions compared to CRC (230). Several transcriptomic analyses studies were conducted on adenomatous polyps in relation to their paired normal tissues (239-241). A more comprehensive investigation of colorectal adenomas was performed in 31 patients using integrated genomics and methylomics analyses,

comparing those with and without malignancy. More mutations, altered expression, and hypermethylation were observed in the cancer-adjacent polyps compared to the cancer-free polyps. For instance, somatic mutations exclusive to the cancer-adjacent polyps included *TP53*, *FBXW7, PIK3CA*, *KIAA1804*, and *SMAD2*. Significant expression changes were noted in *GREM1*, *IGF2*, *CTGF*, and *PLAU*. *FES* and *HES1* exhibited alterations in their expression and methylation levels. Lastly, *ERBB3* and *E2F8* showed changes across all analytical platforms (241). Furthermore, an integrated genomic and transcriptomic analysis was conducted on 26 patients with colorectal adenomas and 28 patients with colorectal cancer, revealing molecular characteristics and correlations between these groups, and suggesting that a 519-gene panel might be utilised for early monitoring of CRC progression (242). Thus, alternative approaches, such as multi-omic analysis or image-based classification (81), may offer more advantages in further investigating the molecular mechanisms underlying the transformation processes of colorectal adenoma to cancer and developing applications for clinical utilisation.

In conclusion, this chapter demonstrated that assessing SMAD4 protein expression through QuPath scoring provides a consistent alternative tool for predicting the development of metachronous polyps or CRC. Among histology subtypes, low SMAD4 expression was associated with a higher rate of subsequent lesion development, particularly in patients with tubulovillous polyps. Transcriptomic analysis comparing low and high levels of SMAD4 expression in tubulovillous polyps revealed differential gene expression between these groups. Further investigation involving CRC tissues would deepen our understanding and enhance strategies for clinical management.

Chapter 5 SMAD4 expression in colorectal cancer and its correlation with clinical outcomes

5.1 Introduction

CRC is recognised as a prevalent cancer leading to a high incidence of cancer-related deaths (31). Several factors, both hereditary and non-hereditary, influence the development of colorectal carcinogenesis (34, 35). The accumulation of various genetic and epigenetic mutations, along with the dysregulation of multiple signalling pathways, can initiate an early precursor, such as premalignant polyps, which then progress into malignant lesions and eventually metastasise (2, 3, 19-21). TGF-ß/BMP signalling pathways are emphasised as being involved in either suppressing or promoting carcinoma progression across different cancer stages (162). SMAD4, a central signal transduction protein in these pathways, was observed to have mutations in 10% of CRC patients (79). Additionally, the reduction of SMAD4 protein levels has been found in patients with colorectal polyps (232) and cancer (237). Patients with *SMAD4* mutations, as well as those with the loss of SMAD4 expression, were predicted to be associated with the late stages of CRC and had a worse survival rate (166, 237, 243, 244).

Aside from the tumour itself, the tumour microenvironment influences tumour progression and metastasis (245). The locations of the tumours that arise contribute to specific biological features and are linked to different clinical significance. Right-sided tumours tended to display microsatellite instability immune phenotypes (CMS1) and were associated with poorer outcomes, whereas left-sided tumours were related to classical types (CMS2) (26, 55). Nevertheless, 23% of CRC patients with CMS4 were suggested to be in advanced stages (III and IV), exhibiting prominent TGF-ß activation, stromal invasion, and angiogenesis. Furthermore, these patients have poorer relapse-free and overall survival rates (57).

Regarding tumour-associated stroma, a cellular process involved in EMT, it promotes the aggressive phenotype of tumours and is regulated by crosstalk among various signalling pathways, including TGF-ß/BMP signalling (126, 127). It has been used as a prognostic tool, with CRC patients showing higher tumour stroma at the invasive margin predicted to have poorer survival outcomes (65). However,

knowledge of these clinicopathological factors in CRC, as well as their relationship with the dysregulation of TGF-ß and BMP signalling pathways, particularly SMAD4, remains limited. Therefore, this chapter aimed to provide a clearer understanding of the clinical relevance associated with dysregulation of SMAD4 expression in CRC by analysing SMAD4 expression levels through IHC and exploring its relationship with clinicopathological factors, especially tumour location and TSP. Furthermore, transcriptomic analysis was examined in these CRC tumours in relation to SMAD4 levels and linked to prognostic significance.

5.2 Clinicopathological characteristics of the cancer cohorts

5.2.1 Clinicopathological characteristics of the early CRC stages cohort

Referring to the previous chapter, which examined SMAD4 expression in colorectal polyps, the same SMAD4 antibody and staining conditions were applied in CRC cohorts to assess the prognostic significance of SMAD4 expression in malignant lesions. The first cohort included TMA tissues from early-stage CRC patients (n = 159), as described in Section 2.1.2.2. After excluding cases with poor tissue quality, 129 cases with valid SMAD4 staining results were analysed (Figure 5-1A). The patients in this study consisted of 69.8% males (n = 90), and 52.7% were aged 65 years or older (n = 68). Tumours location consisted of; the right colon in 14.0% (n = 18), the left colon in 58.9% (n = 76), and the rectum in 27.1% (n = 35). The tumours were classified as TNM stage I in 88.4% (n = 114), TNM stage II no tumours and stage III in 11.6% (n = 15). Additionally, the lesions were diagnosed as well to moderately differentiated in 96.9% (n = 125) and poorly differentiated in 3.1% (n = 4) (Figure 5-1B).

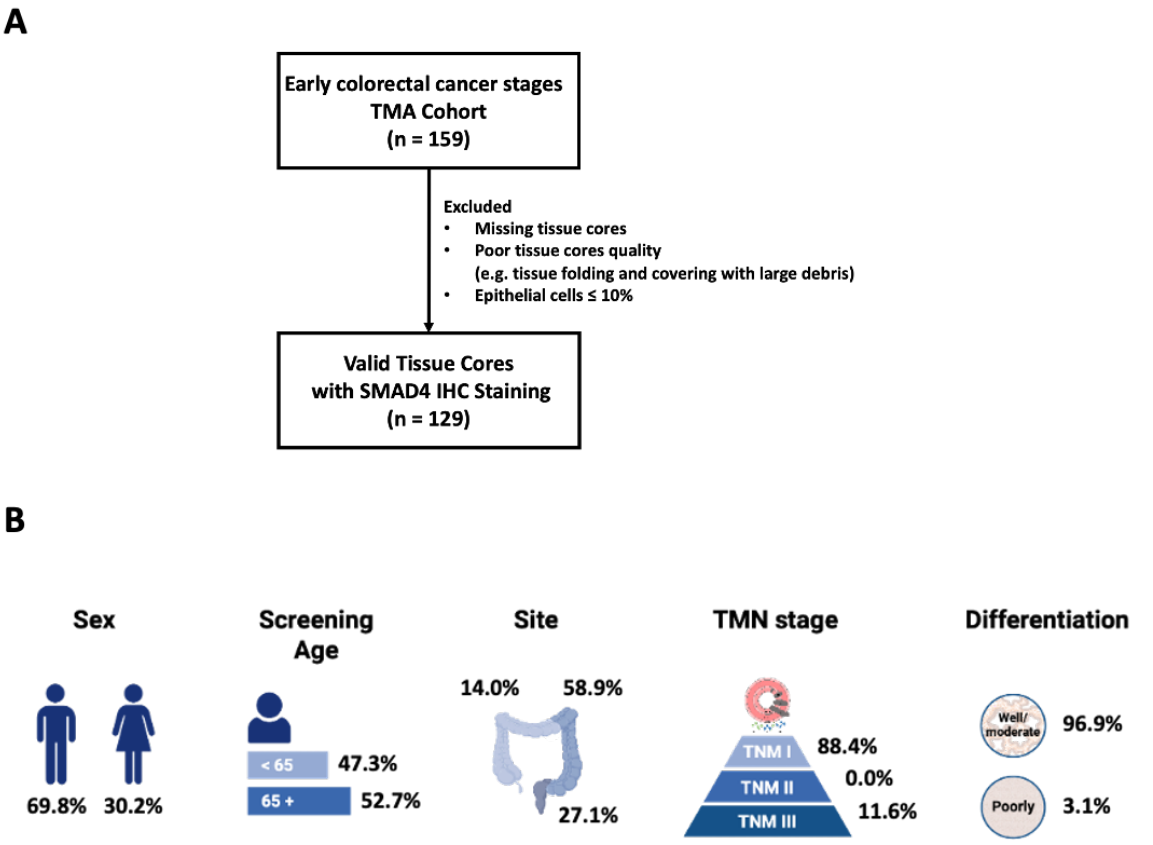


Figure 5-1 Colorectal cancer and clinicopathological features of the early-stage CRC cohort

The Consort diagram shows how CRC samples were included in this study (A). The tissue microarrays (TMAs) of CRC from the early-stage CRC cohort were stained for SMAD4 expression using Immunohistochemistry (IHC) and analysed in relation to clinical outcomes. The demographic chart displays the characteristics of this cohort (B).

5.2.2 Clinicopathological characteristics of the GRI cohort

The second cohort included TMA tissues from patients with CRC stages I-III (n = 787), excluding stage IV cases, as described in 2.1.2.3. Using the same exclusion criteria, samples with poorer tissue quality were excluded, and the remaining 727 cases were regarded as valid based on SMAD4 staining. Subsequently, eligible cases were excluded if mortality occurred within 30 days of surgery or if they had received neoadjuvant therapy. A total of 673 CRC carcinoma cases were analysed for the prognostic significance of SMAD4 expression. In these patients, TSP scores were available for 649 cases (Figure 5-2A).

The characteristics of the patients included 55.3% male patients (n = 372), and 44.7% female patients (n = 301), as well as 31.6% patients younger than 65 years old (n = 213), 32.1% aged between 65 and 74 years old (n = 216), and 36.3% aged 65 years or older (n = 244). Tumours were found in the right colon of 42.9% (n = 289), the left colon of 32.1% (n = 216), and the rectum of 25% (n = 168). Moreover, the tumours were classified as TNM stage I in 12.2% (n = 82), stage II in 47.5% (n = 320), and stage III in 40.3% (n = 271). Pathological diagnosis revealed that tumours were well to moderately differentiated in 88.8% (n = 594) and poorly differentiated in 11.2% (n = 75). MMR status with deficient MMR (dMMR) in 18.2% (n = 120), proficient MMR (pMMR) in 65.9% (n = 435), and loss of set (MLH1/PMS2 or MHS2/6) in 15.9% (n = 105). Recurrence parameters were considered in this study. Petersen index (PI) was low (PI 0-1) in 77.9% (n = 524) and high (PI 2-5) in 22.1% (n = 149). Low tumour budding was observed in 69.4% (n = 451), while 30.6% of cases exhibited high tumour budding (n = 199). DFS status included 32.8% with no recurrence (n = 215), 31.5% with recurrence (n = 206), and 35.7% who died but not from cancer (n = 234). In relation to the tumour microenvironment (TME), GMS refers to a local TME. Patients were scored as follows: GMS0, 16.2% (n = 105); GMS1, 63.3% (n = 410); and GMS2, 20.5% (n = 133). mGPS refers to a systemic effect of TME, including mGPS0 of 59.9% (n = 403), mGPS1 of 23.8% (n = 160), and mGPS2 of 16.3% (n = 110). In addition, the KMG was 83.8% (n = 543) and 16.2% (n = 105) in the low and high groups, respectively. A low TSP was presented in 77.8% (n = 505), and a higher TSP was presented in 22.2% (n = 144) (Figure 5-2B).

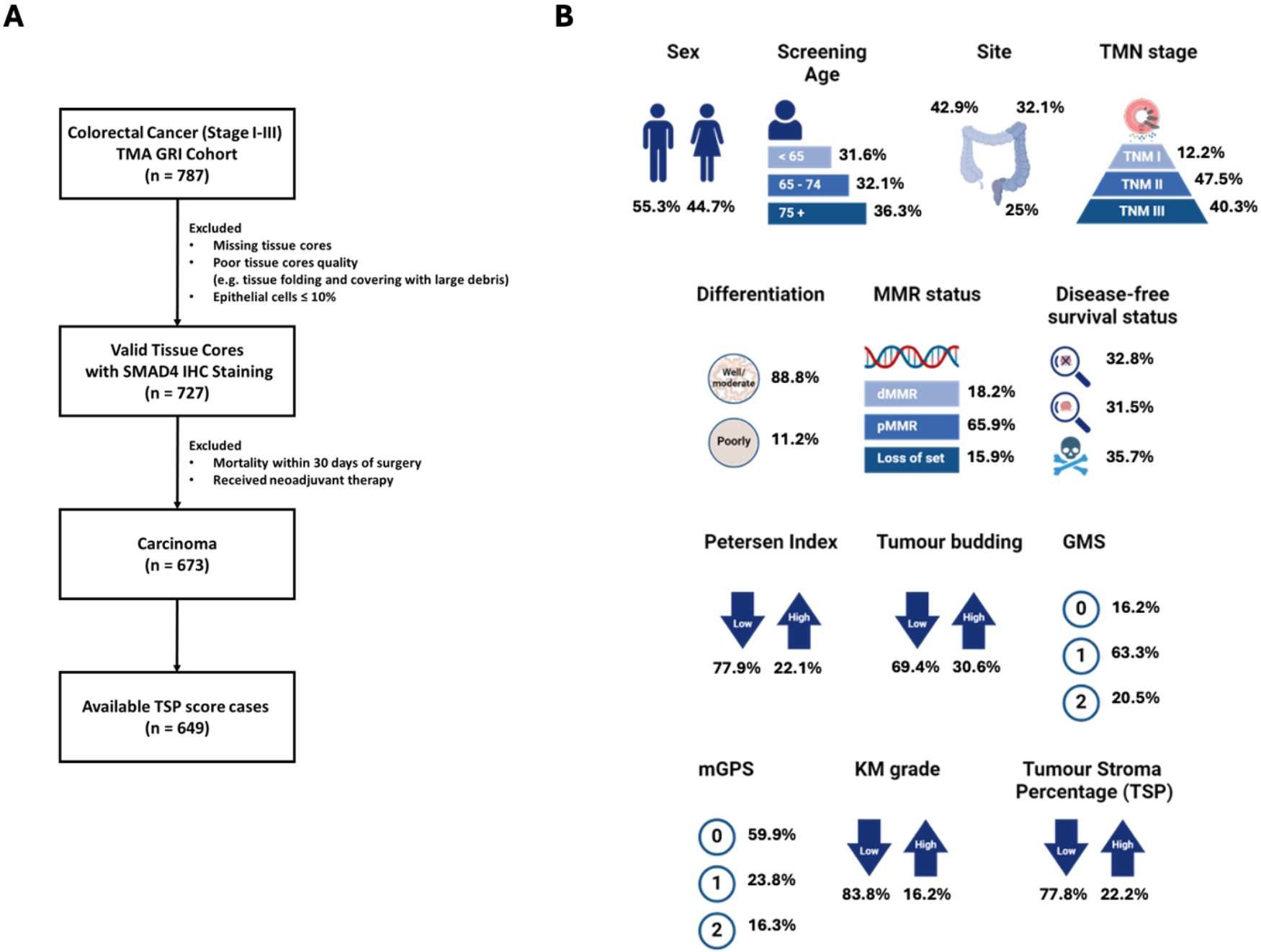

A
Colorectal Cancer (Stage I-III)
TMA GRI Cohort
(n = 787)
Excluded
• Missing tissue cores
• Poor tissue cores quality
(e.g. tissue folding and covering with large debris)
• Epithelial cells ≤ 10%
Valid Tissue Cores
with SMAD4 IHC Staining
(n = 727)
Excluded
• Mortality within 30 days of surgery
• Received neoadjuvant therapy
Carcinoma
(n = 673)
Available TSP score cases
(n = 649)
B
Sex
55.3% 44.7%
Screening Age
< 65 31.6%
65 - 74 32.1%
75 + 36.3%
Site
42.9% 32.1%
25%
TMN stage
TNM I 12.2%
TNM II 47.5%
TNM III 40.3%
Differentiation
Well/ moderate 88.8%
Poorly 11.2%
MMR status
dMMR 18.2%
pMMR 65.9%
Loss of set 15.9%
Disease-free survival status
32.8%
31.5%
35.7%
Petersen Index
Low High
77.9% 22.1%
Tumour budding
Low High
69.4% 30.6%
GMS
0 16.2%
1 63.3%
2 20.5%
mGPS
0 59.9%
1 23.8%
2 16.3%
KM grade
Low High
83.8% 16.2%
Tumour Stroma Percentage (TSP)
Low High
77.8% 22.2%

Figure 5-2 Colorectal cancer and clinicopathological features of the GRI cohort

The Consort diagram shows how CRC samples were included in this study (A). The tissue microarrays (TMAs) of CRC from the GRI cohort were stained for SMAD4 expression using Immunohistochemistry (IHC) and analysed in relation to clinical outcomes. The demographic chart displays the characteristics of this cohort (B).

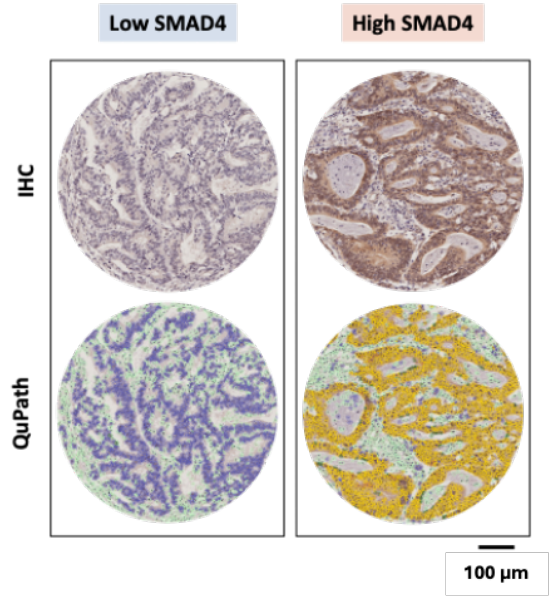
5.3 Expression of SMAD4 in CRC and its association with clinical outcomes and characteristics

5.3.1 Association between SMAD4 expression, clinical characteristics, and clinical outcomes in the early CRC stages cohort

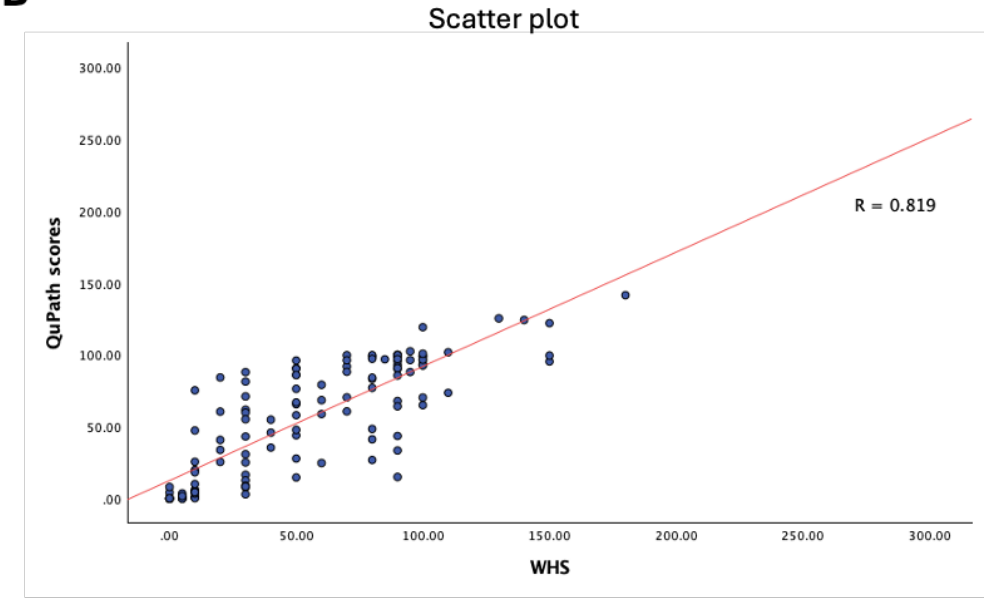
SMAD4 staining was performed by IHC as described in Section 2.2.2.2. The staining intensity of SMAD4 in the cytoplasmic and nuclear compartments of epithelial cells was then evaluated using 10% of TMA cores by WHS and all cores by QuPath, as shown in the representative images in Figure 5-3A. Similar to the previous SMAD4 expression patterns in Chapter 4, low SMAD4 expression in tumour epithelium exhibited a weak brown colour in the stained tissues and appeared as a blue area in QuPath. In contrast, higher SMAD4 expression levels appeared as darker brown and yellow areas, respectively. The reliability of using QuPath scores was assessed, and the scatter plot revealed a strong positive correlation with a correlation coefficient of 0.819 between WHS and QuPath scores (Figure 5-3B). The Bland-Altman plot also confirmed a positive correlation with an ICC of 0.901, indicating no significant difference between these two scores (Figure 5-3C). In this cohort, we obtained the QuPath scores separately for cytoplasmic and nuclear SMAD4 expression in the tumour of 129 cases. The correlation between QuPath scores of cytoplasmic and nuclear SMAD4 expression was evaluated, revealing a strong correlation (R = 0.806, Figure 5-3D). Cut points for cytoplasmic and nuclear SMAD4 expression scores were determined at 35.21 and 97, respectively, for dividing patients into low and high SMAD4 expression groups (Figure 5-3E and F).

Subsequently, we combined the scores from cytoplasmic and nuclear SMAD4 expression in tumours and classified patients into three groups: those with both low cytoplasmic and nuclear SMAD4 expression (n = 72), those with only one high SMAD4 expression (n = 40), and those with both high SMAD4 expression in tumour cytoplasm and nucleus (n = 17). Due to the limited survival data in this small CRC cohort, we then examined the relationship between SMAD4 expression in tumour epithelium and clinicopathological features, including sex, age, TNM stages, tumour site, and histological differentiation grade. The Chi-squared table results showed that SMAD4 expression levels were associated with tumour stages and locations (p = 0.002 and 0.05, respectively) (Table 5-1).

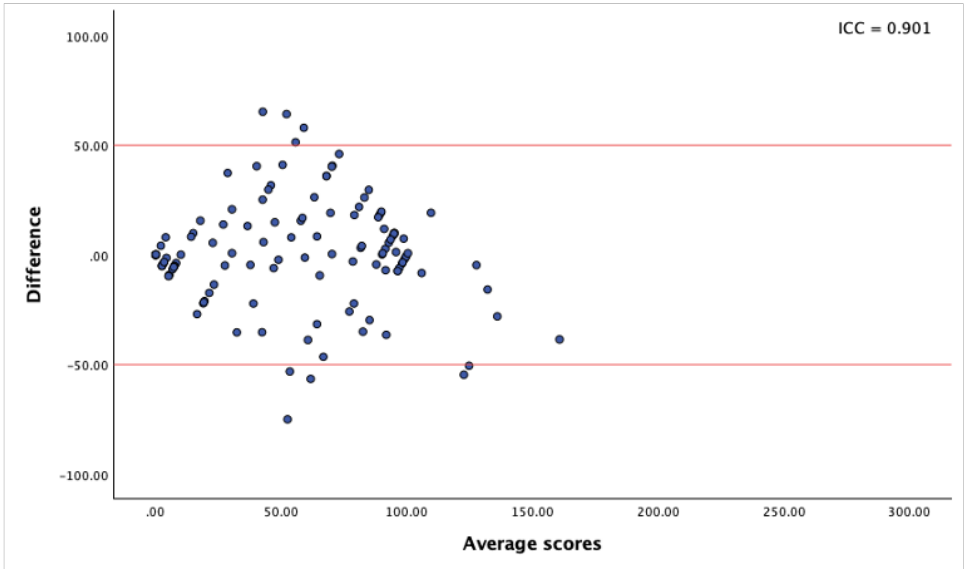
A


Low SMAD4
High SMAD4
IHC
QuPath
100 µm

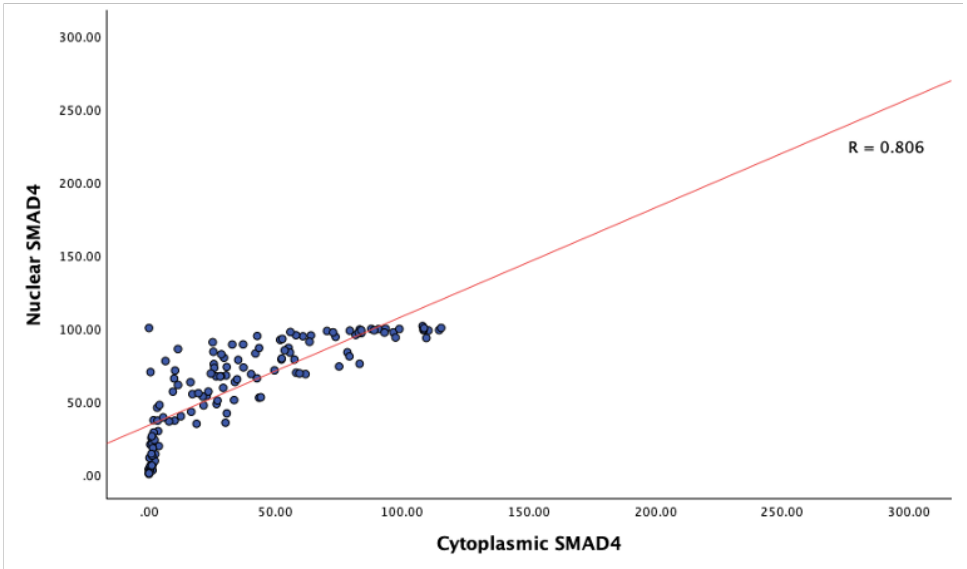
B


Scatter plot
QuPath scores
WHS
R = 0.819
300.00
250.00
200.00
150.00
100.00
50.00
.00

C


Bland Altmann plot
ICC = 0.901
Difference
Average scores
100.00
50.00
.00
-50.00
-100.00

D


Cytoplasmic and nuclear SMAD4 scores
Nuclear SMAD4
Cytoplasmic SMAD4
R = 0.806
300.00
250.00
200.00
150.00
100.00
50.00
.00

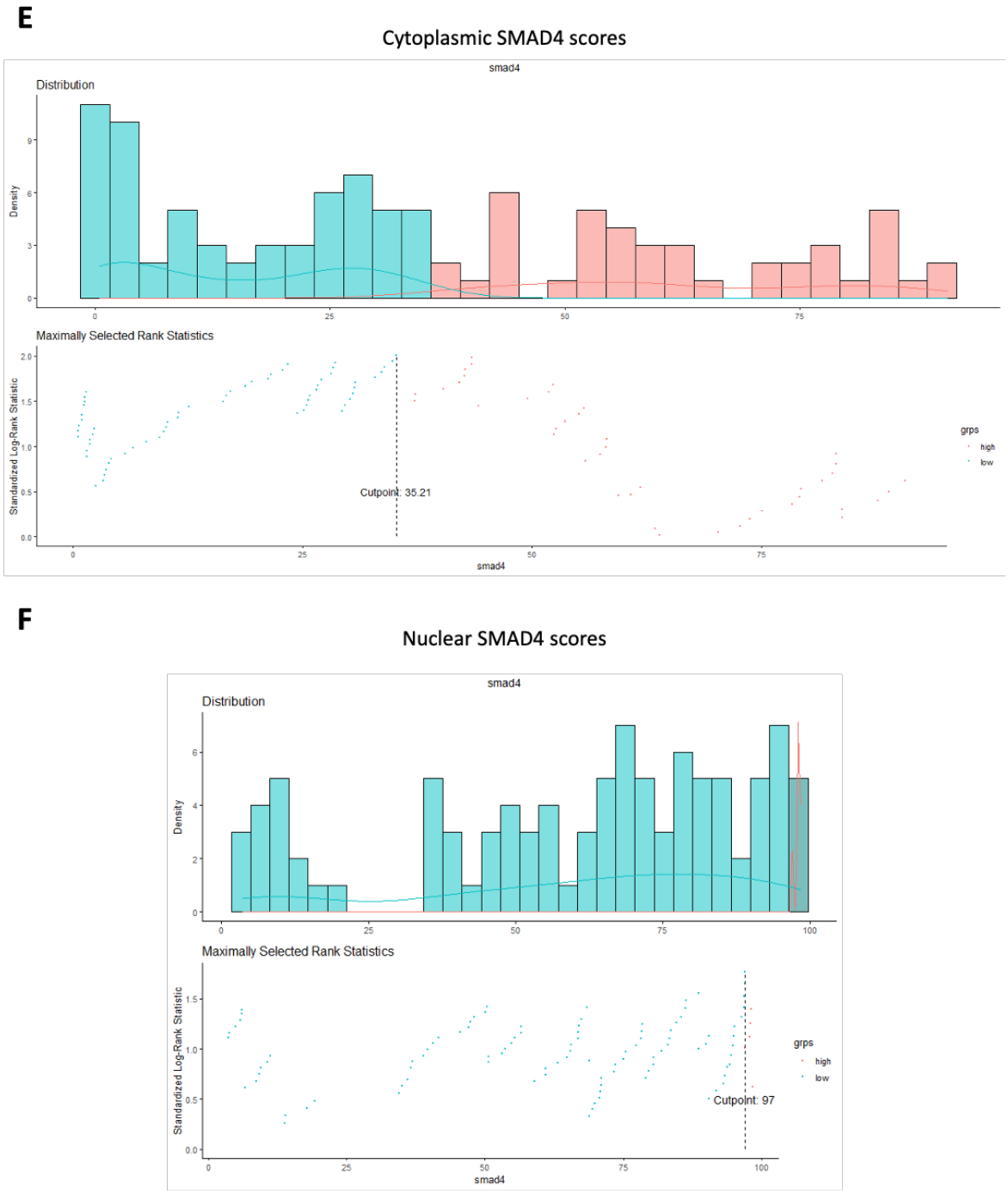


Figure 5-3 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in the early-stage CRC Cohort

Representative images of colorectal cancer showing low SMAD4 expression (left) and high SMAD4 expression (right) from IHC staining (top) and classified by QuPath (bottom) (A). A scatter plot of WHS and QuPath scores demonstrated a correlation between these two scores in detecting SMAD4 expression (R = 0.819) (B). A Bland-Altman plot indicated no significant difference between the scores and showed a positive correlation (ICC = 0.901) (C). Scatter plots of QuPath scores revealed a strong correlation between cytoplasmic and nuclear scores (R = 0.806) (D). The histogram of an optimal cut point was determined by the survminer package using maximally selected rank statistics in R, as shown at 35.21 and 97 for cytoplasmic and nuclear SMAD4 expression in this cohort (E and F). The scale bar = 100 μm in Figures A.

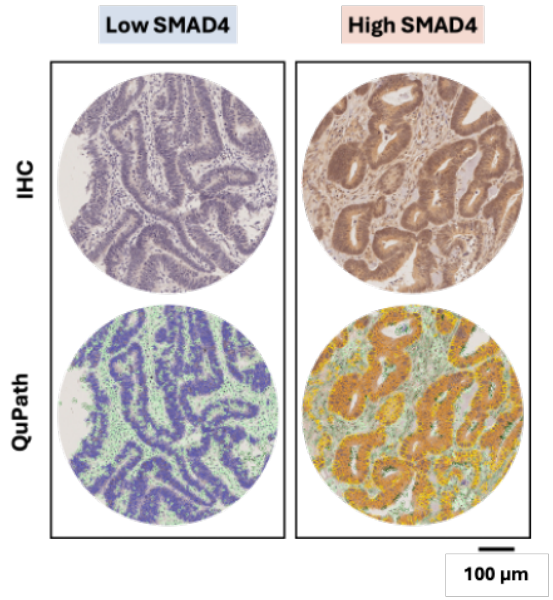
Table 5-1 The relationship of SMAD4 expression in early-stage CRC and clinicopathological characteristics

	Combined cytoplasmic and nuclear SMAD4 expression in early CRC stages cohort			
	Both low n = 72 (55.8%)	One high n = 40 (31%)	Both high n = 17 (13.2%)	p
Sex				
Male	53 (73.6%)	28 (70.0%)	9 (52.9%)	0.268
Female	19 (26.4%)	12 (30.0%)	8 (47.1%)	
Age				
< 65 years	36 (50.0%)	18 (45.0%)	7 (41.2%)	0.758
≥ 65 years	36 (50.0%)	22 (55.0%)	10 (58.8%)	
TNM stage				
TMN I	58 (80.6%)	39 (97.5%)	17 (100%)	0.002
TMN III	14 (19.4%)	1 (2.5%)	0 (0%)	
Tumour sites				
Rectum	18 (25.0%)	16 (40.0%)	1 (5.9%)	0.05
Left Colon	42 (58.3%)	21 (52.5%)	13 (76.5%)	
Right Colon	12 (16.7%)	3 (7.5%)	3 (17.6%)	
Differentiation				
Well/Moderate	70 (97.2%)	39 (97.5%)	16 (94.1%)	0.808
Poor	2 (2.8%)	1 (2.5%)	1 (5.9%)	

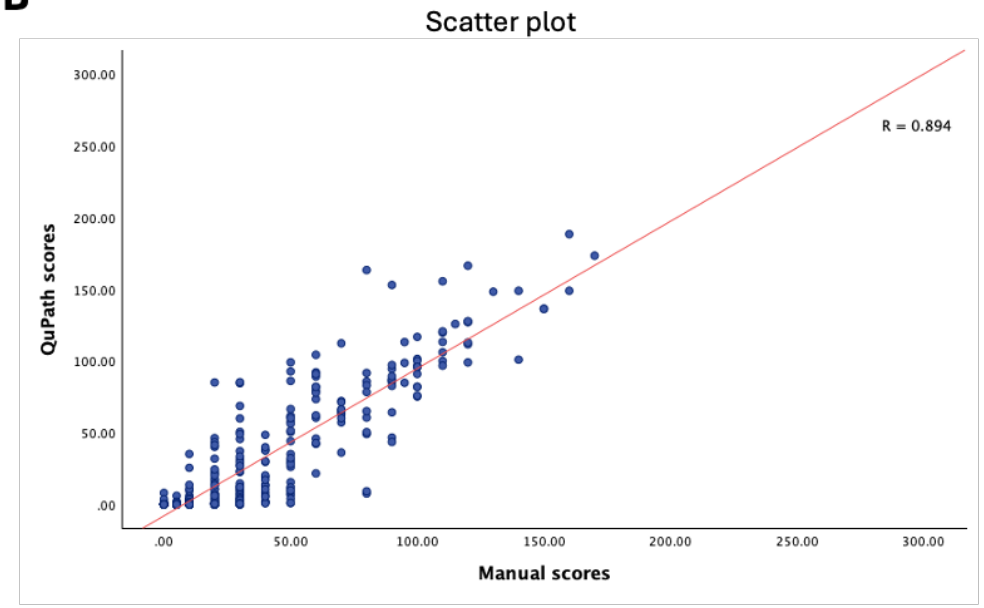
5.3.2 Association between SMAD4 expression, clinical characteristics, and clinical outcomes in the GRI cohort

TMAs from the GRI cohort were stained to evaluate SMAD4 expression. Using the same methods, WHS and QuPath scores were also obtained. Representative staining patterns demonstrating low and high SMAD4 expression in tumour epithelium from this cohort are shown (Figure 5-4A). A strong positive correlation was confirmed, with a correlation coefficient of 0.894 between WHS and QuPath scores (Figure 5-4B). The Bland-Altman plot also indicated a positive correlation, with an ICC of 0.931 (Figure 5-4C). The QuPath scores were analysed based on cytoplasmic and nuclear SMAD4 expression in the tumours of 673 cases. The correlation between cytoplasmic, nuclear, and cellular scores of SMAD4 expression was examined. Scatter plots revealed a strong correlation among these three scores (including cytoplasm and nucleus, R = 0.946; cytoplasm and cellular, R = 0.976; nucleus and cellular, R = 0.987; Figure 5-4D to F, respectively). Similar to the polyp cohort, cut points were established using the cellular SMAD4 expression scores at 27.27 (Figure 5-4G).

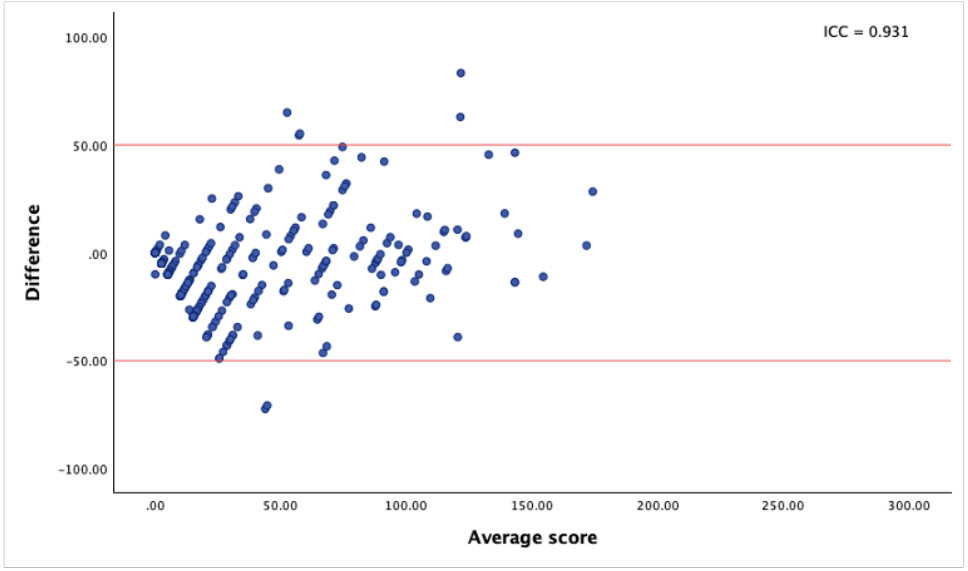
A


Low SMAD4
High SMAD4
IHC
QuPath
100 µm

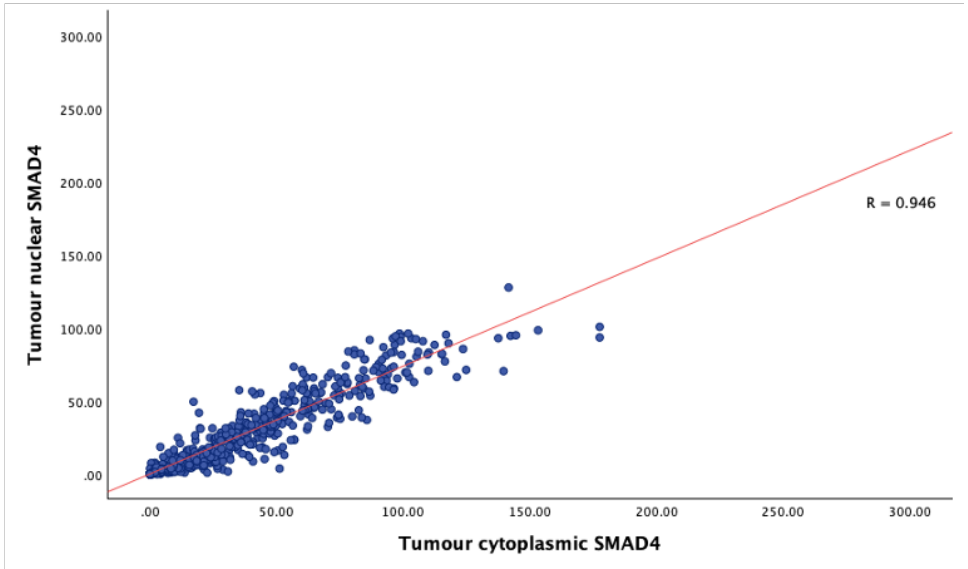
B


Scatter plot
R = 0.894
QuPath scores
Manual scores
300.00
250.00
200.00
150.00
100.00
50.00
.00

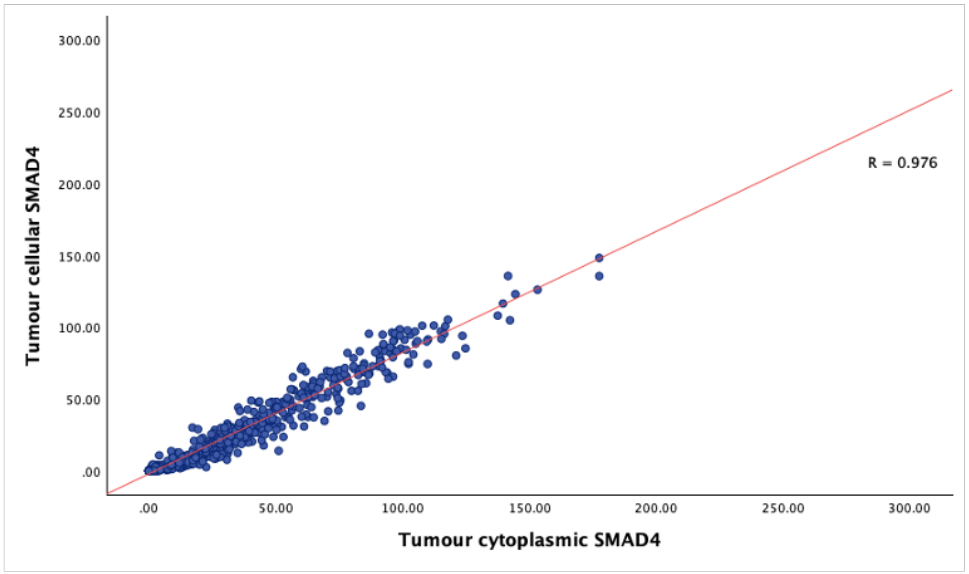
C


Bland Altmann plot
ICC = 0.931
Difference
Average score
100.00
50.00
.00
-50.00
-100.00

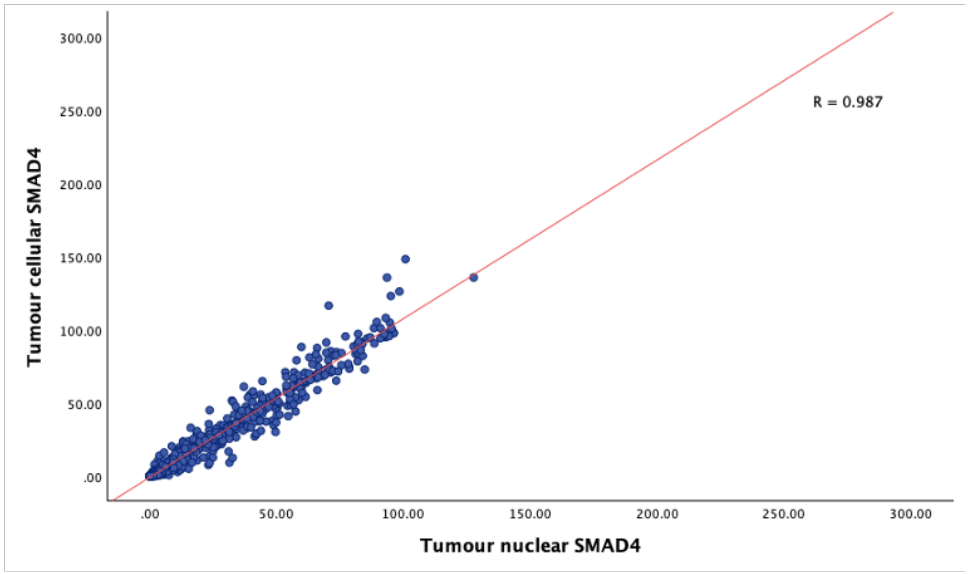
D


Tumour cytoplasmic and nuclear SMAD4 scores
R = 0.946
Tumour nuclear SMAD4
Tumour cytoplasmic SMAD4
300.00
250.00
200.00
150.00
100.00
50.00
.00

E


Tumour cytoplasmic and cellular SMAD4 scores
Tumour cellular SMAD4
Tumour cytoplasmic SMAD4
R = 0.976
300.00
250.00
200.00
150.00
100.00
50.00
.00

F


Tumour nuclear and cellular SMAD4 scores
Tumour cellular SMAD4
Tumour nuclear SMAD4
R = 0.987
300.00
250.00
200.00
150.00
100.00
50.00
.00

G

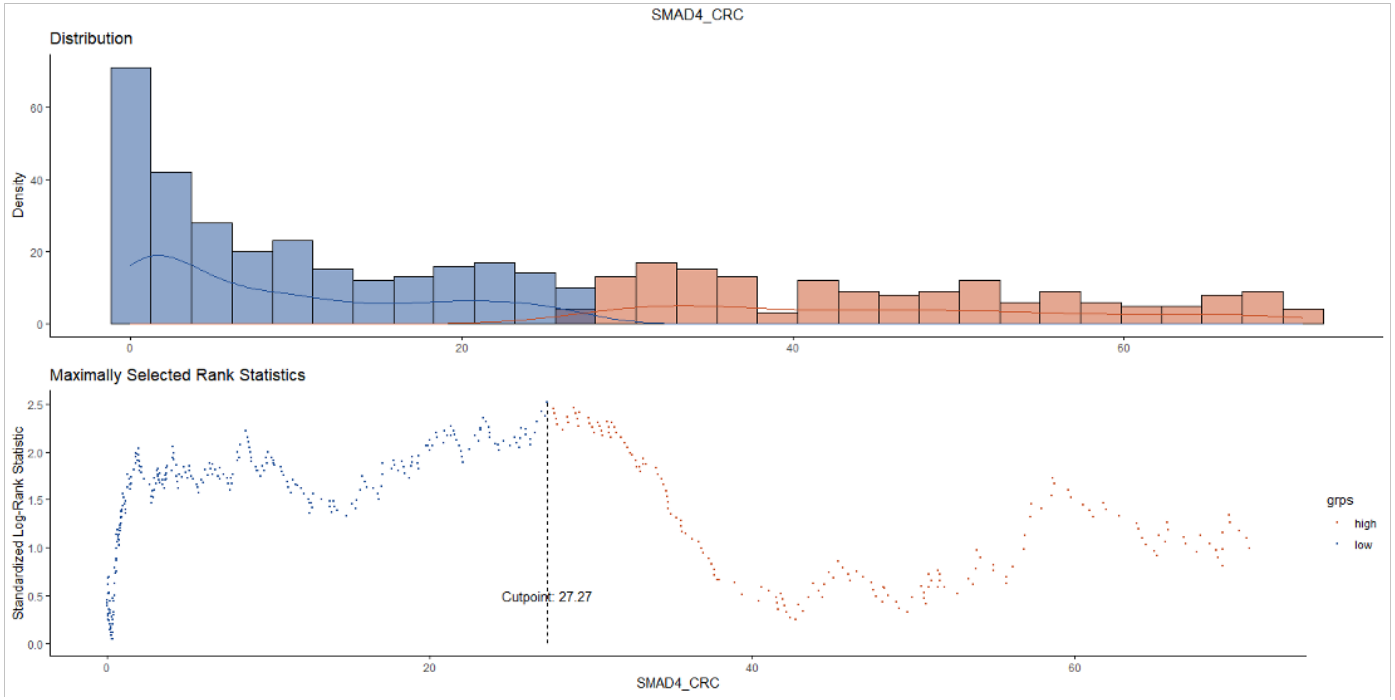

SMAD4_CRC
Distribution
Density
Maximally Selected Rank Statistics
Standardized Log-Rank Statistic
Cutpoint: 27.27
grps
high
low
SMAD4_CRC

Figure 5-4 SMAD4 expression staining pattern in IHC and QuPath classification, with correlation of scores and the cut-off point of expression in CRC stage I-III from the GRI cohort

Representative images of colorectal cancer showing low SMAD4 expression (left) and high SMAD4 expression (right) from IHC staining (top) and classified by QuPath (bottom) (A). A scatter plot of WHS and QuPath scores demonstrated a correlation between these two scores in detecting SMAD4 expression (R = 0.894) (B). A Bland-Altman plot indicated no significant difference between the scores and showed a positive correlation (ICC = 0.931) (C). Scatter plots of QuPath scores revealed a strong correlation between cytoplasmic and nuclear scores (R = 0.946) (D), cytoplasmic and cellular scores (R = 0.976) (E), and nuclear and cellular scores (R = 0.987) (F). The histogram of an optimal cut point was determined by the survminer package using maximally selected rank statistics in R, as shown at 27.27 from the GRI cohort (G). The scale bar = 100 μm in Figures A.

Patients with CRC tumours in the GRI cohort were subsequently stratified into two groups based on SMAD4 expression levels, comprising 435 cases with low SMAD4 expression and 238 cases with high SMAD4 expression. A prognostic analysis was conducted to evaluate the relationship between SMAD4 expression and survival time. The KM curve for the whole cohort showed that low SMAD4 expression in tumour epithelial cells was significantly associated with a reduced cancer-specific survival rate (Log-rank, p = 0.019) (Figure 5-5A). The average survival time for patients with low SMAD4 expression was shorter than for those with high SMAD4 expression, at 192.9 (181.1-204.8) months compared to 209.3 (190.4-209.7) months. The relationship between SMAD4 expression in tumours and clinicopathological features was analysed. Chi-squared results indicated that there was no association between SMAD4 expression and patient characteristics, including sex and age. Regarding tumour features, it was found that low SMAD4 expression was significantly associated with a higher tumour stage (p = 0.004) and histological differentiation grade (p = 0.05), while no significant correlation was observed with tumour location or MMR status. Additionally, disease-free status was found to be associated with SMAD4 expression levels. However, no evidence was shown of a relationship between SMAD4 expression and the Petersen index or tumour budding. Concerning tumour microenvironment parameters, SMAD4 expression levels in tumours were not significantly related to GMS, mGPS, or KMG. Nevertheless, low tumour SMAD4 expression was associated with higher stromal content (p = 0.027) (Figure 5-5B and Table 5-2). Cox regression analysis indicated that age, TNM stage, MMR status, PI, tumour budding, disease-free status, GMS, mGMS, KMG, TSP, and SMAD4 expression were independent predictors in univariate analysis (p < 0.05). High SMAD4 expression was independently linked to better survival outcomes (HR = 0.691, 95% CI = 0.507-0.943, p = 0.002). Multivariate analysis revealed that certain parameters were relatively significant

predictors of outcomes, indicating higher risk in this cohort. These included patients aged 75 years and older (HR = 1.791, 95% CI = 1.212-2.647, p = 0.003), patients with a high Petersen index (HR = 1.719, 95% CI = 1.225-2.412, p = 0.002), and those presenting recurrence (HR = 438.746, 95% CI = 60.794-3166.417, p < 0.001) (Table 5-2).

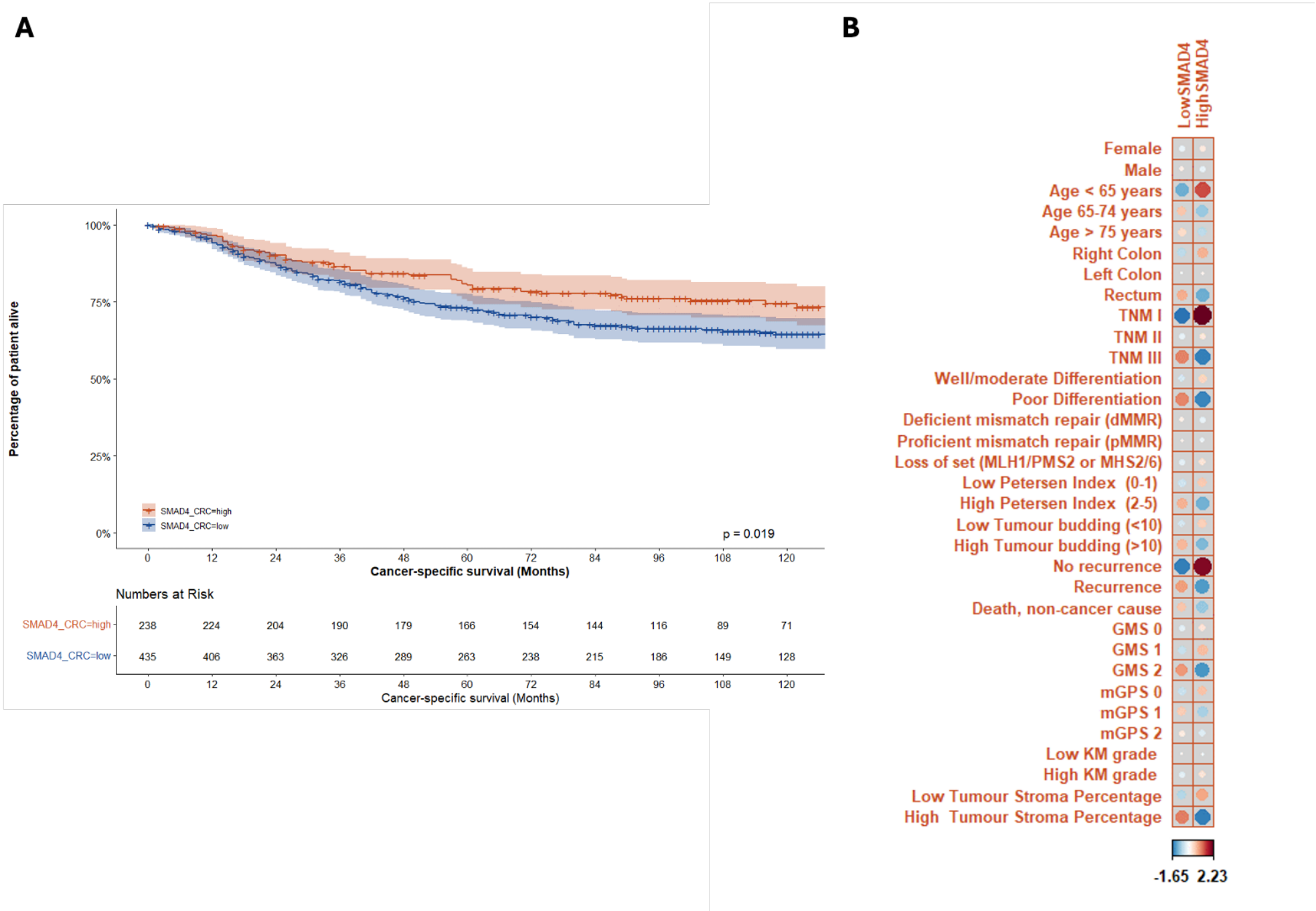

A
100%
75%
50%
25%
0%
Percentage of patient alive
SMAD4_CRC=high
SMAD4_CRC=low
p = 0.019
0
12
24
36
48
60
72
84
96
108
120
Cancer-specific survival (Months)
Numbers at Risk
SMAD4_CRC=high 238 224 204 190 179 166 154 144 116 89 71
SMAD4_CRC=low 435 406 363 326 289 263 238 215 186 149 128
Cancer-specific survival (Months)
B
LowSMAD4
HighSMAD4
Female
Male
Age < 65 years
Age 65-74 years
Age > 75 years
Right Colon
Left Colon
Rectum
TNM I
TNM II
TNM III
Well/moderate Differentiation
Poor Differentiation
Deficient mismatch repair (dMMR)
Proficient mismatch repair (pMMR)
Loss of set (MLH1/PMS2 or MHS2/6)
Low Petersen Index (0-1)
High Petersen Index (2-5)
Low Tumour budding (<10)
High Tumour budding (>10)
No recurrence
Recurrence
Death, non-cancer cause
GMS 0
GMS 1
GMS 2
mGPS 0
mGPS 1
mGPS 2
Low KM grade
High KM grade
Low Tumour Stroma Percentage
High Tumour Stroma Percentage
-1.65 2.23

Figure 5-5 SMAD4 expression in CRC predicting survival outcomes and its correlation with clinical characteristics

The Kaplan-Meier curve for SMAD4 expression in colorectal cancer showed that low levels of SMAD4 are associated with a reduced rate of cancer-specific survival times (Log-rank, $p = 0.019$) (A). The CORR plot illustrated the relationship between SMAD4 expression levels and the clinicopathological characteristics of patients with CRC (B).

Table 5-2 The relationship of SMAD4 expression in CRC stage I-III and clinicopathological characteristics from the GRI cohort

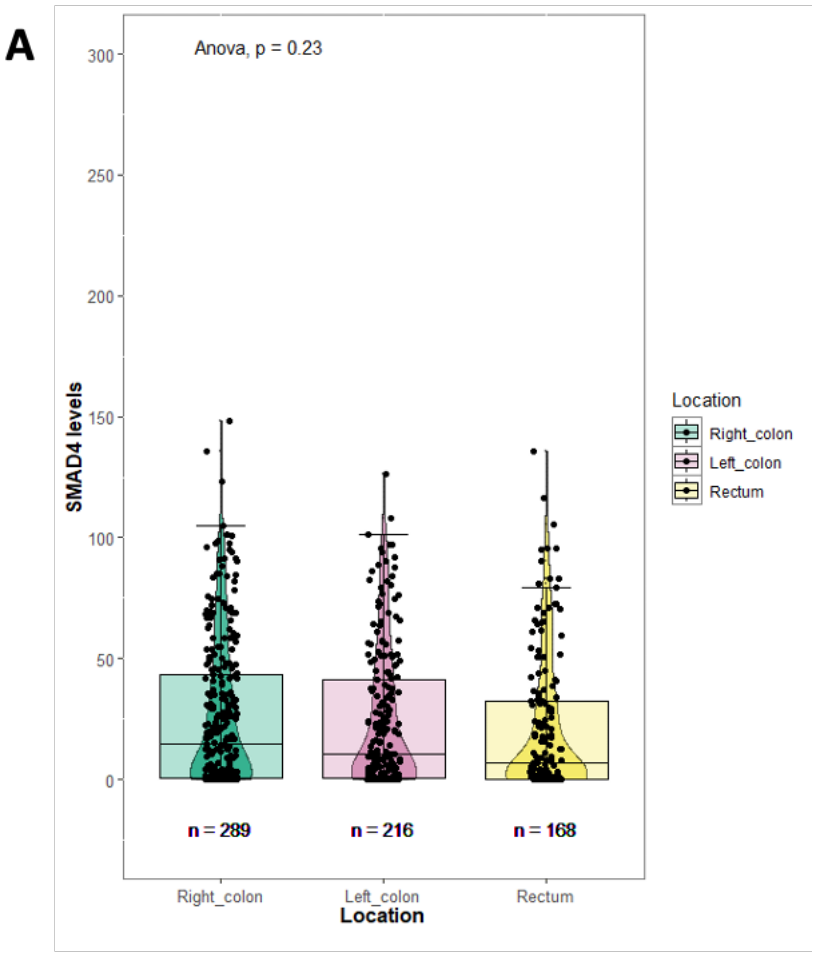
	SMAD4 expression in tumours from GRI cohort			Univariate analysis		Multivariate analysis	
	Low n (%)	High n (%)	*p*	HR (95% CI)	*p*	HR (95% CI)	*p*
Patient characteristics							
Sex							
Female	192 (44.1%)	109 (45.8%)	0.679	1		-	-
Male	243 (55.9%)	129 (54.2%)		1.194 (0.896-1.591)	0.227	-	-
Age							
<65 years	125 (28.7%)	88 (37.0%)	0.088	1		1	
65-74 years	147 (33.8%)	69 (29.0%)		1.185 (0.822-1.709)	0.364	1.285 (0.856-1.931)	0.226
≥ 75 years	163 (37.5%)	81 (34.0%)		1.661 (1.174-2.349)	0.004	1.791 (1.212-2.647)	0.003
Tumour characteristics							
Location							
Right colon	179 (41.1%)	110 (46.2%)	0.246	1		-	-
Left colon	139 (32.0%)	77 (32.4%)		1.102 (0.796-1.525)	0.559	-	-
Rectum	117 (26.9%)	51 (21.4%)		1.034 (0.722-1.482)	0.854	-	-
TNM stage							
TNM I	41 (9.4%)	41 (17.2%)	0.004	1		1	
TNM II	204 (46.9%)	116 (48.7%)		2.831 (1.301-6.158)	0.009	1.254 (0.526-2.991)	0.610
TNM III	190 (43.7%)	81 (34%)		6.413 (2.99-13.753)	< 0.001	2.009 (0.859-4.695)	0.107
Differentiation							
Well to moderately	377 (87.1%)	217 (91.9%)	0.05	1		-	-
Poor	56 (12.9%)	19 (8.1%)		1.462 (0.96-2.225)	0.077	-	-
Mismatch repair (MMR) status							
Deficient mismatch repair (dMMR)	79 (18.5%)	41 (17.7%)	0.887	1		1	
Proficient mismatch repair (pMMR)	283 (66.1%)	152 (65.5%)		0.781 (0.551-1.107)	0.165	0.939 (0.639-1.378)	0.747
Loss of set (MLH1/PMS2 or MHS2/6)	66 (15.4%)	39 (16.8%)		0.522 (0.309-0.882)	0.015	0.710 (0.389-1.296)	0.264

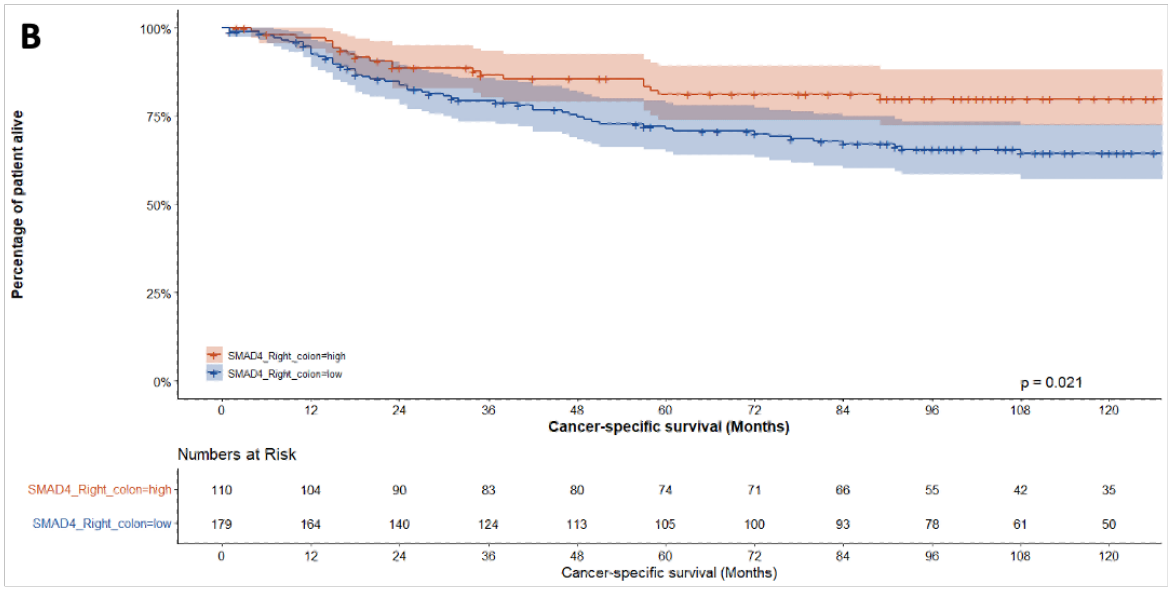
Table 5-2 The relationship of SMAD4 expression in CRC stage I-III and clinicopathological characteristics from the GRI cohort (continue)

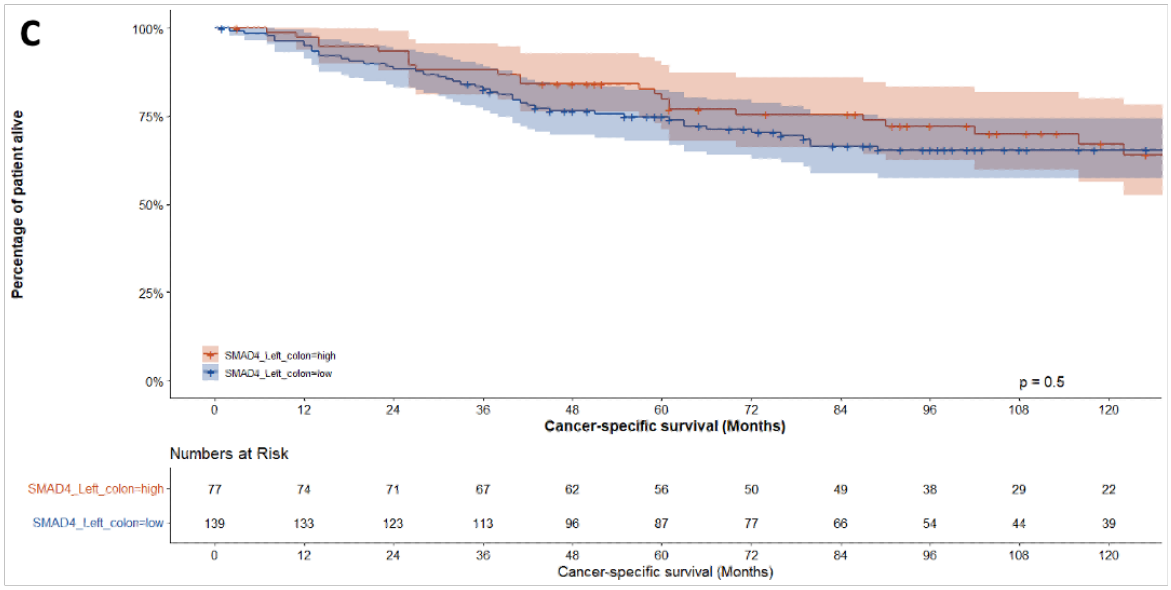
	SMAD4 expression in tumours from GRI cohort			Univariate analysis		Multivariate analysis	
	Low n (%)	High n (%)	*p*	HR (95% CI)	*p*	HR (95% CI)	*p*
Recurrence factors							
Petersen Index (PI)							
Low (0-1)	331 (76.1%)	193 (81.1%)	0.132	1		1	
High (2-5)	104 (23.9%)	45 (18.9%)		2.53 (1.887-3.394)	< 0.001	1.719 (1.225-2.412)	0.002
Tumour budding TB)							
low (<10)	285 (67.5%)	166 (72.8%)	0.162	1		1	
high (>10)	137 (32.5%)	62 (27.2%)		2.099 (1.571-2.804)	< 0.001	1.073 (0.771-1.494)	0.675
Disease-free survival (DFS) status							
No recurrence	121 (28.5%)	94 (40.7%)	0.006	1		1	
Recurrence	144 (34.0%)	62 (26.8%)		496.013 (69.246-3552.964)	< 0.001	438.746 (60.794-3166.417)	< 0.001
Death, non-cancer cause	159 (37.5%)	75 (32.5%)		0 (0-1.466E+43)	0.888	0.003 (0-2.294E+14)	0.771
Tumour microenvironment							
Glasgow Microenvironment Score (GMS)							
GMS0	66 (15.8%)	39 (17.0%)	0.166	1		1	
GMS1	257 (61.5%)	153 (66.5%)		1.817 (1.105-2.988)	0.019	1.811 (0.974-3.367)	0.060
GMS2	95 (22.7%)	38 (16.5%)		3.314 (1.947-5.642)	< 0.001	1310.254 (0-8.700E+41)	0.875
Modified Glasgow Prognostic Score (mGPS)							
mGPS0	253 (58.2%)	150 (63.0%)	0.446	1		1	
mGPS1	109 (25.1%)	51 (21.4%)		1.495 (1.07-2.087)	0.018	1.277 (0.876-1.861)	0.204
mGPS2	73 (16.8%)	37 (15.5%)		1.947 (1.356-2.796)	< 0.001	1.488 (0.967-2.289)	0.071
Klintrup-Mäkinen grade (KMG)							
Low KMG	352 (84.2%)	191 (83.0%)	0.7	1		-	-
High KMG	66 (15.8%)	39 (17.0%)		0.469 (0.288-0.762)	0.002	-	-
Tumour stromal percentages (TSP)							
Low TSP	315 (75.2%)	190 (82.6%)	0.027	1		1	
High TSP	104 (24.8%)	40 (17.4%)		1.822 (1.334-2.488)	< 0.001	0.002 (0-1.099E+36)	0.888
SMAD4 expression in tumours							
Low SMAD4	-	-	-	1		1	
High SMAD4	-	-	-	0.691 (0.507-0.943)	0.02	1.070 (0.743-1.540)	0.716

5.4 SMAD4 expression in CRC and sites of tumour development predicts survival outcomes

Although the previous result showed no association between SMAD4 expression levels and tumour location, tumours arising at different sites may be driven by various biological mechanisms that specifically influence tumour size and survival outcomes in these patients. To explore this, we examined the levels of SMAD4 expression in tumours from the right colon, the left colon, and the rectum. The findings indicated no difference in SMAD4 expression levels across the three locations (ANOVA, p = 0.23) (Figure 5-6A). When patients were grouped based on tumour location, Kaplan-Meier survival analysis demonstrated that low SMAD4 expression was significantly associated with higher cancer mortality in the right-sided colon compared to patients with high SMAD4 (p = 0.021) (Figure 5-6B), but there was no difference in the left-sided colon (p = 0.5) or rectum (p = 0.37) (Figure 5-6C and D).




B
Percentage of patient alive
100%
75%
50%
25%
0%
SMAD4_Right_colon=high
SMAD4_Right_colon=low
p = 0.021
0
12
24
36
48
60
72
84
96
108
120
Cancer-specific survival (Months)
Numbers at Risk
SMAD4_Right_colon=high 110 104 90 83 80 74 71 66 55 42 35
SMAD4_Right_colon=low 179 164 140 124 113 105 100 93 78 61 50
Cancer-specific survival (Months)


C
Percentage of patient alive
100%
75%
50%
25%
0%
SMAD4_Left_colon=high
SMAD4_Left_colon=low
p = 0.5
0
12
24
36
48
60
72
84
96
108
120
Cancer-specific survival (Months)
Numbers at Risk
SMAD4_Left_colon=high 77 74 71 67 62 56 50 49 38 29 22
SMAD4_Left_colon=low 139 133 123 113 96 87 77 66 54 44 39
Cancer-specific survival (Months)

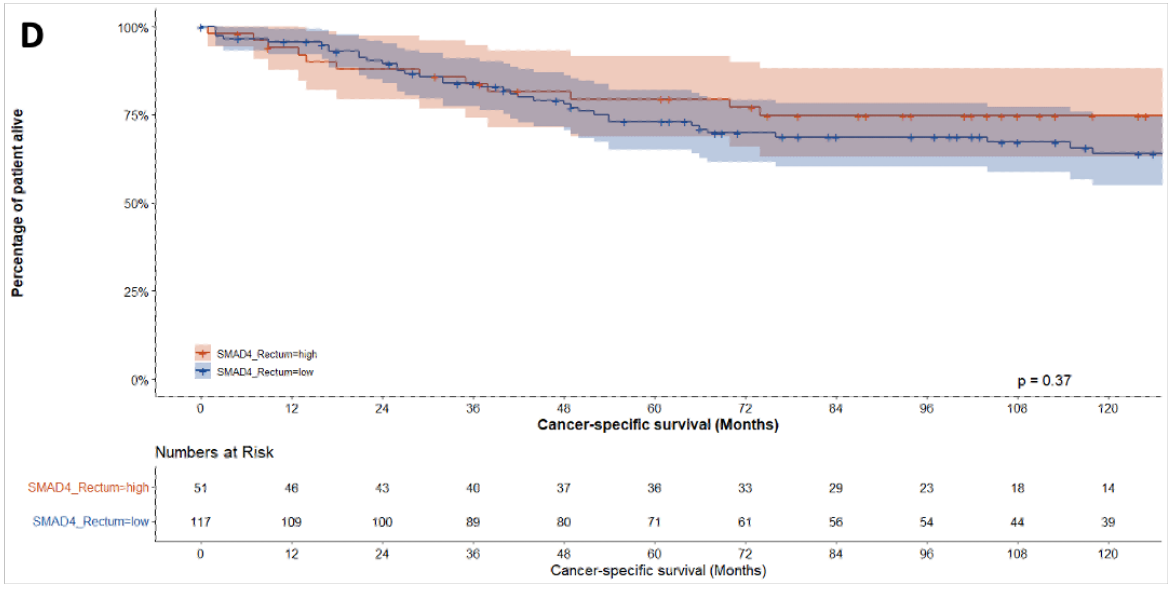

D
Percentage of patient alive
100%
75%
50%
25%
0%
SMAD4_Rectum=high
SMAD4_Rectum=low
p = 0.37
0
12
24
36
48
60
72
84
96
108
120
Cancer-specific survival (Months)
Numbers at Risk
SMAD4_Rectum=high 51 46 43 40 37 36 33 29 23 18 14
SMAD4_Rectum=low 117 109 100 89 80 71 61 56 54 44 39
Cancer-specific survival (Months)

Figure 5-6 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal cancer, stratified by tumour location

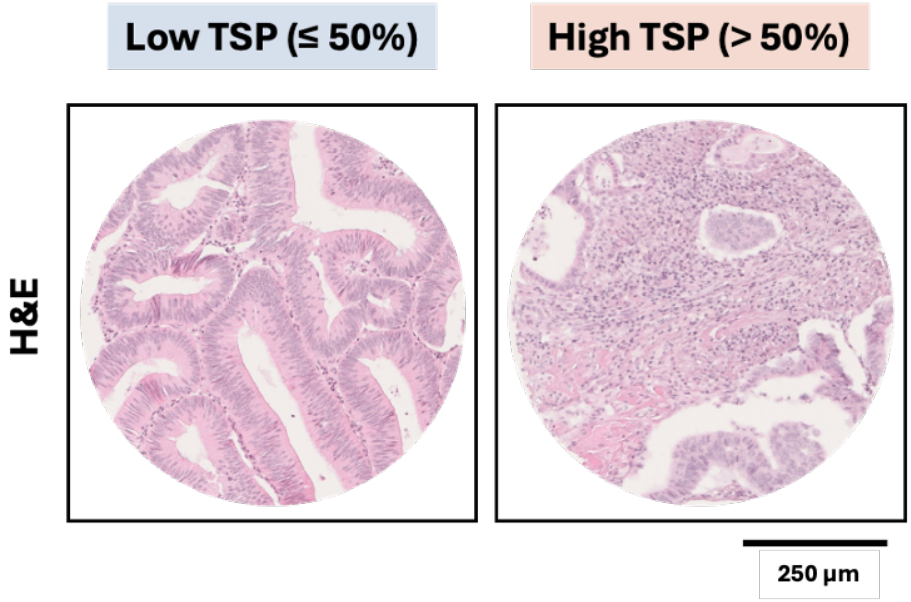
A boxplot showed no significant difference in SMAD4 expression levels among right-colon, left-colon, and rectum (ANOVA, p = 0.23) (A). Kaplan-Meier curve of SMAD4 expression in right colonal tumours showed that low SMAD4 expression was associated with a decresing cancer-specific survival time (Log-rank, p = 0.021) (B). Whereas, Kaplan-Meier curve of SMAD4 expression in tumours arising at left colon (C) and rectum (D) showed that low SMAD4 expression was not associated with differece survival outcomes compared to the other group (Log-rank, p = 0.05 and 0.37, respectively).

5.5 SMAD4 expression in CRC tumours with high tumour stroma predicts survival outcomes and contributes to transcriptomic analysis

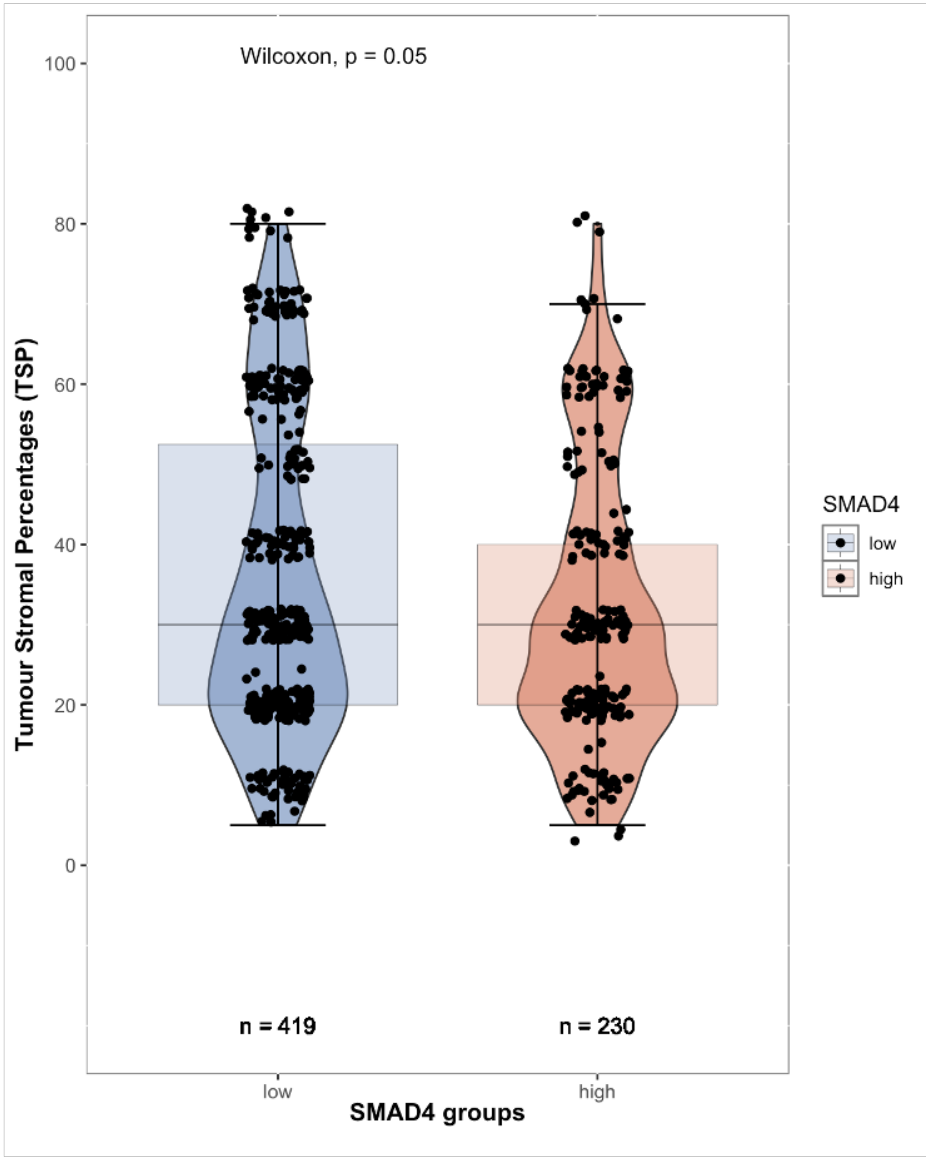
5.5.1 Combining tumour SMAD4 expression levels with tumour stroma percentages in CRC and their association with survival prognosis

Since tumour stroma was suggested to influence the TME during carcinogenesis and survival outcomes, and our findings showed a correlation between tumour SMAD4 expression and stroma, this study combined TSP and tumour SMAD4 expression to assess their prognostic significance. Previously, the TSP for each case was evaluated using H&E staining and classified into low and high TSP based on whether stromal content was less than or more than 50% of the tissue in the full section (Figure 5-7A). When the percentage of stroma was compared between the low and high SMAD4 groups of CRC patients (n = 419 and 230, respectively), the results showed that a higher TSP was observed in the low SMAD4 group (Wilcoxon test, p = 0.05) (Figure 5-7B). Subsequently, the CRC patients were divided into four groups: low TSP with low SMAD4 in tumours (n = 315), low TSP with high SMAD4 in tumours (n = 190), high TSP with low SMAD4 in tumours (n = 104), and high TSP with high SMAD4 in tumours (n = 40). Survival analysis indicated that patients with high stroma and low SMAD4 expression in the tumour were predicted to have the worst survival outcomes compared to the other groups (Log-rank, p < 0.001; Figure 5-7C).

A


Low TSP (≤ 50%)
High TSP (> 50%)
H&E
250 µm

B


Wilcoxon, p = 0.05
Tumour Stromal Percentages (TSP)
100
80
60
40
20
0
n = 419
n = 230
low
high
SMAD4 groups
SMAD4
low
high

C

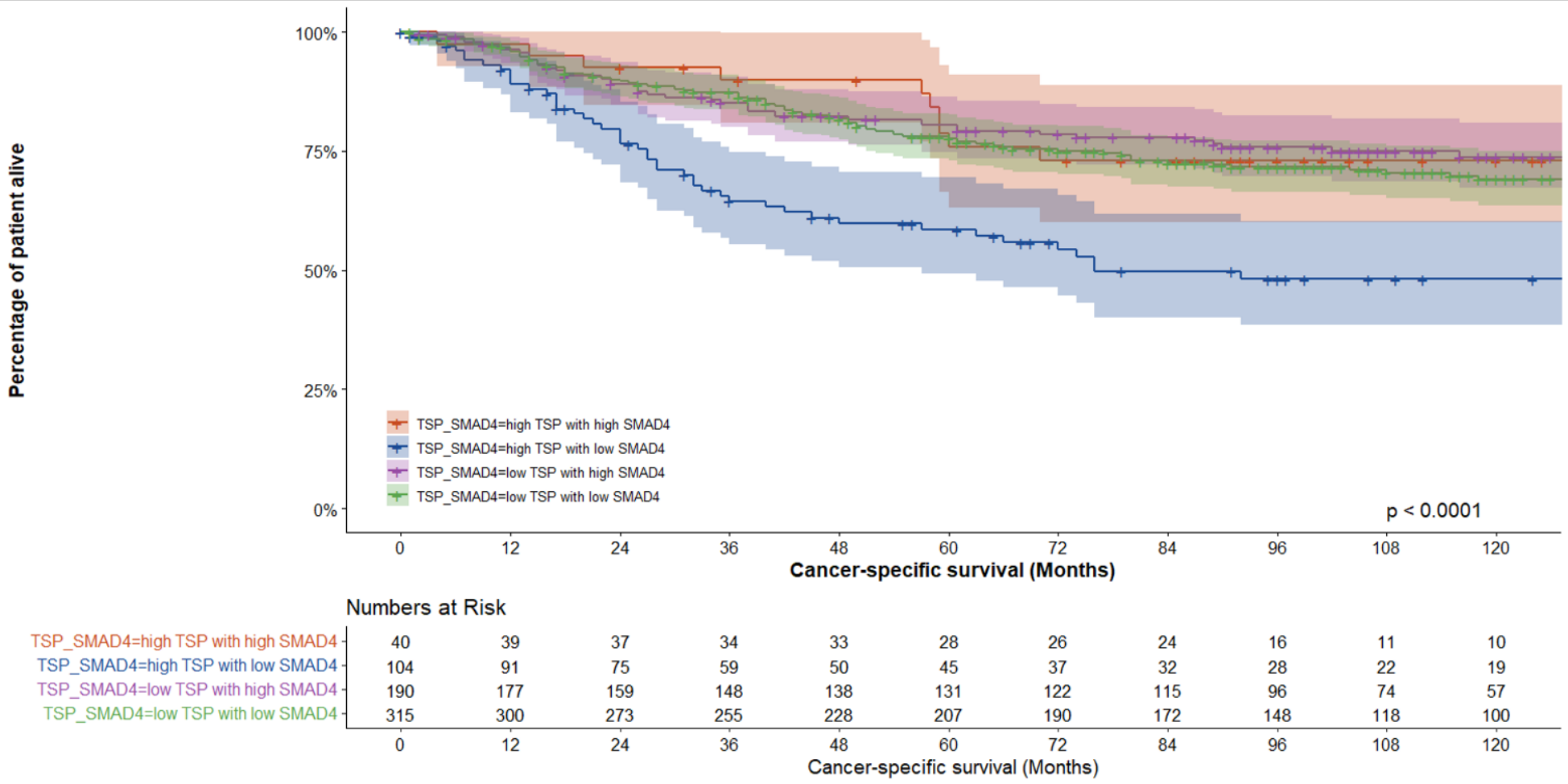

Percentage of patient alive
100%
75%
50%
25%
0%
TSP_SMAD4=high TSP with high SMAD4
TSP_SMAD4=high TSP with low SMAD4
TSP_SMAD4=low TSP with high SMAD4
TSP_SMAD4=low TSP with low SMAD4
p < 0.0001
0
12
24
36
48
60
72
84
96
108
120
Cancer-specific survival (Months)
Numbers at Risk
TSP_SMAD4=high TSP with high SMAD4 40 39 37 34 33 28 26 24 16 11 10
TSP_SMAD4=high TSP with low SMAD4 104 91 75 59 50 45 37 32 28 22 19
TSP_SMAD4=low TSP with high SMAD4 190 177 159 148 138 131 122 115 96 74 57
TSP_SMAD4=low TSP with low SMAD4 315 300 273 255 228 207 190 172 148 118 100
0 12 24 36 48 60 72 84 96 108 120
Cancer-specific survival (Months)

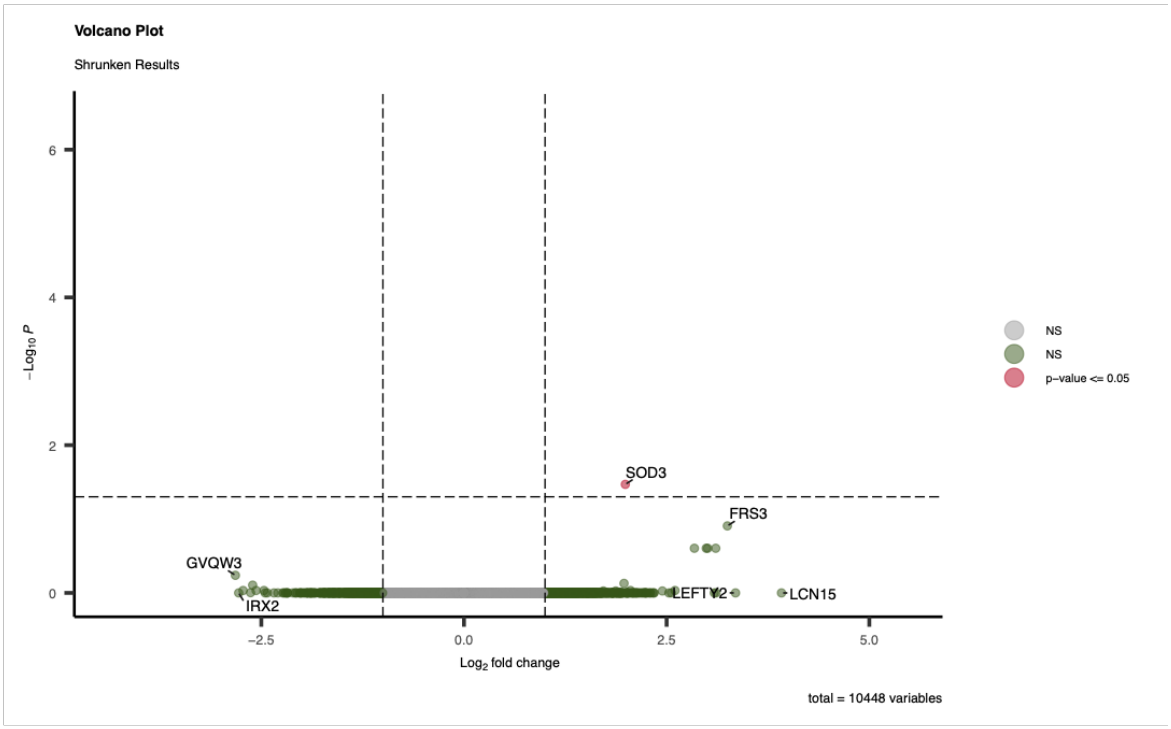
Figure 5-7 Relationship between SMAD4 expression levels and clinical outcomes in patients with colorectal cancer, stratified by tumour stroma percentage (TSP)

Representative images of colorectal cancer tissues showing low TSP (left) and high TSP (right) from H&E staining (A). A boxplot revealed a difference in TSP levels between tumours with low and high SMAD4 expression (Wilcoxon test, $p = 0.05$) (B). Kaplan-Meier curve of combining TSP and tumour SMAD4 expression showed that patients with high stroma and low tumour SMAD4 expression were associated with a decreasing cancer-specific survival time (Log-rank, $p = 0.021$) (C). The scale bar = 250 μm in Figures A.

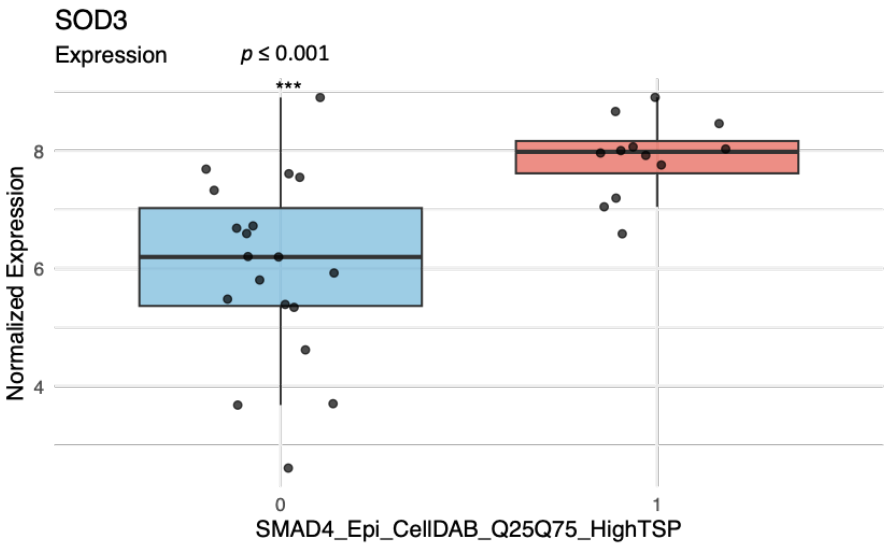
5.5.2 Transcriptomic analysis of tumours with low and high SMAD4 expression involving high stroma in CRC

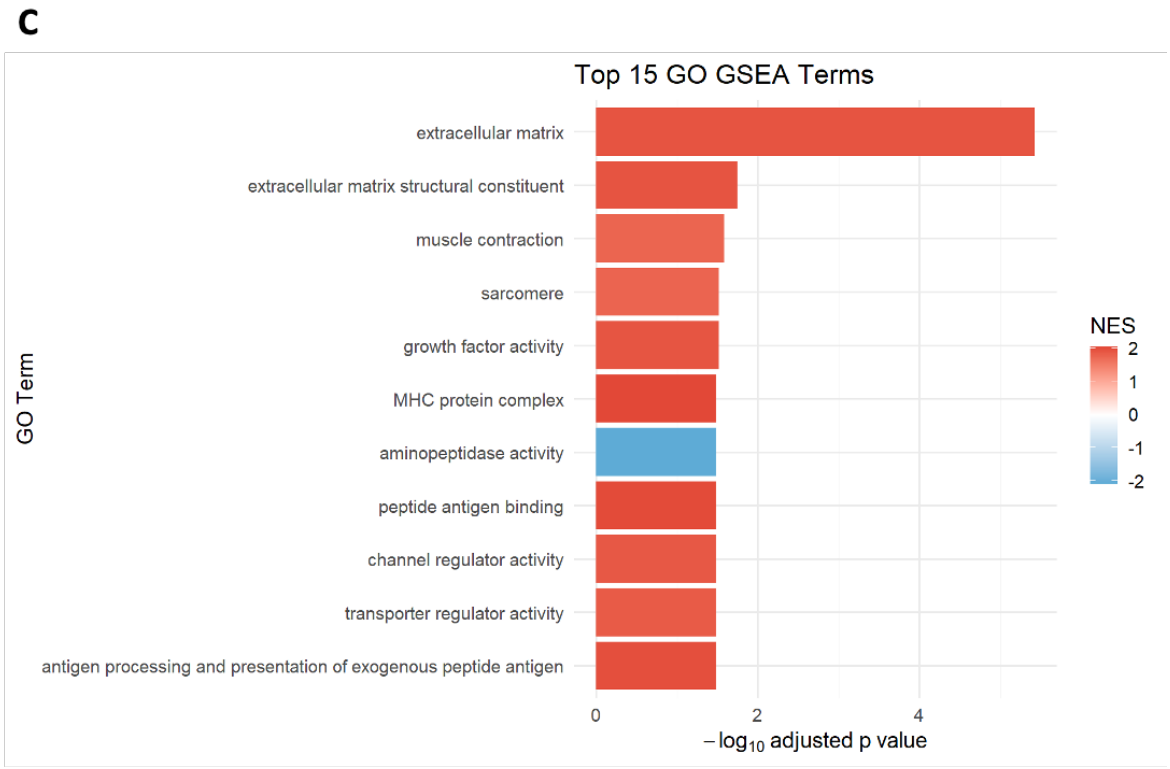
To enhance understanding of dysregulation at the gene expression level involving changes in SMAD4 protein levels in tumours with high stroma, we compared transcriptomic analyses of CRC patients with high TSP, focusing on tumours with low and high SMAD4 expression from the first and fourth quartiles. These represent samples with very low (n = 37) and high SMAD4 levels (n = 30), respectively, as described in Section 2.2.2.6. After differential expression analysis, the volcano plot showed that *SOD3* was upregulated in the high SMAD4 group (adjusted p-value ≤ 0.05) (Figure 5-8A). The boxplot illustrated that the mRNA level of *SOD3* was significantly decreased in the low SMAD4 group (Wilcoxon test, $p \leq 0.001$) (Figure 5-8B). Subsequently, GSEA analysis revealed that 11 gene sets in Gene Ontology (GO) were enriched (adjusted p-value ≤ 0.05). The gene sets for extracellular matrix, extracellular matrix structural constituent, muscle contraction, sarcomere, growth factor activity, MHC protein complex, peptide antigen binding, channel regulator activity, transporter regulator activity, and antigen processing and presentation of exogenous peptide antigen were enriched in a high SMAD4 group with NES > 1.5. In contrast, aminopeptidase activity was enriched in a low SMAD4 group with NES < -1.5 (Figure 5-8C). For GSEA with the KEGG pathways, several pathways were enriched in a high SMAD4 group (NES > 1.5, adjusted p-value ≤ 0.05), such as cytoskeleton in muscle cells, ECM-receptor interaction, PI3K-Akt signalling pathway (Figure 5-8D).

A


Volcano Plot
Shrunken Results
−Log10 P
6
4
2
0
SOD3
FRS3
GVQW3
IRX2
LEFTY2
LCN15
−2.5
0.0
2.5
5.0
Log2 fold change
NS
NS
p−value <= 0.05
total = 10448 variables

B


SOD3
Expression
p ≤ 0.001
***
Normalized Expression
8
6
4
0
1
SMAD4_Epi_CellDAB_Q25Q75_HighTSP



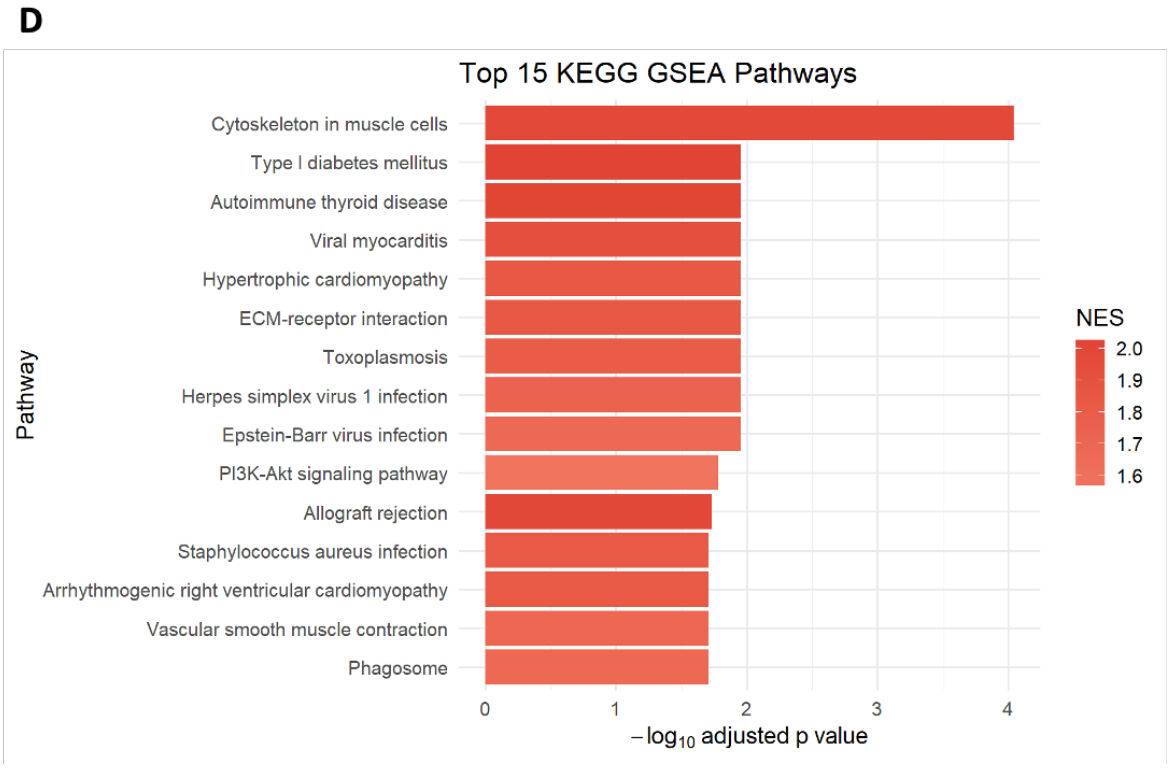


Figure 5-8 Transcriptomic analysis comparing low and high SMAD4 expression in CRC tumours with high stroma

A volcano plot displayed differential gene expression between the low SMAD4 group (n = 37) and the high SMAD4 group (n = 30), with upregulated *SOD3* in the high SMAD4 tumour with high stroma (adjusted p-value ≤ 0.05) (A). A boxplot comparing *SOD3* gene expression levels between the low- and high-SMAD4 protein groups showed a significant increase in RNA levels in the high group (Wilcoxon test, *** $p \leq 0.001$) (B). GSEA analysis plots showed that 11 gene sets were enriched

based on Gene Ontology (GO) (C), and the top 15 gene sets were enriched based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (D). Normalised Enrichment Score (NES) > 1.5 is represented in red and < -1.5 is represented in blue. Significance at adjusted p-value ≤ 0.05.

5.6 Discussion

TGF-β/BMP signalling pathways are involved in maintaining tissue homeostasis. Dysregulation of these components leads to altered functions and ultimately to disease development (162). SMAD4 dysregulation has been identified in CRC carcinogenesis and serves as a valuable tool for predicting metachronous lesions, especially in patients with tubulovillous polyps, based on our previous findings in the colorectal polyp study. Additionally, other studies have demonstrated that the *SMAD4* gene mutation in CRC patients was associated with colon cancer rather than rectal cancer, more frequent in female patients than males, related to higher TNM stage, lymph node metastasis, mucinous differentiation, *RAS* mutation status, and shorter overall survival times (243, 246).

Here, we explored the prognostic significance of SMAD4 protein alterations and their relationship with clinical characteristics in CRC patients. The Chi- squared results from two CRC cohorts, including the early CRC cohort (n = 129) and the GRI cohort of stages I-III (n = 673), showed that lower SMAD4 expression in tumour epithelium was significantly linked to higher tumour stages (p = 0.002 in the early CRC cohort and p = 0.004 in the GRI cohort). Although it was suggested that SMAD4 expression was associated with tumour locations in the early CRC cohort (p = 0.05), this was not statistically significant in the GRI cohort (p = 0.246). Furthermore, SMAD4 levels correlated with histological differentiation grade (p = 0.05), and notably, decreased SMAD4 was associated with higher tumour stroma (p = 0.027). These findings indicated that the status of SMAD4 expression in CRC tumours, as assessed by immunohistochemistry, was linked to more aggressive features of CRC. Similarly, loss of SMAD4 previously analysed in a cohort of 364 CRC patients has been suggested to be significantly associated with higher tumour and nodal stage, in addition to being correlated with other parameters, such as adjuvant therapy use, fewer tumour-infiltrating lymphocytes (TIL), and lower peritumoural lymphocyte aggregate (PLA) scores (all, p < 0.04). Contrastingly, their results were presented in relation to MMR (p = 0.024), but not with regard to tumour location (p = 0.22) (237). Another study (n = 1,281) also confirmed that the loss of nuclear SMAD4 was associated with a

higher TNM stage (p < 0.01) and suggested more frequent distant metastasis (p = 0.001), as well as presenting lymphovascular and perineural invasion, and fewer tumour-infiltrating lymphocytes (p < 0.001). However, they found a contrasting result with the relationship between SMAD4 expression and tumour budding (p < 0.001), along with no correlation with tumour location (p = 0.222) and differentiation grade (p = 0.923) (247).

Considering the prognostic significance of SMAD4 expression in CRC tumours and survival outcomes, a low SMAD4 level predicted decreased cancer-specific survival in CRC patients from the GRI cohort (p = 0.019). Similar findings have been reported in other studies, demonstrating the prognostic value of SMAD4 status via IHC in predicting poorer survival outcomes (165, 237, 243, 247, 248). Although it has been suggested that the right-sided tumours were linked to higher tumour stage, poor tumour differentiation, elevated tumour lymphocytic infiltrate and systemic inflammatory response (249), along with predicted worse DFS (250), there was limited understanding of the prognostic significance of SMAD4 levels in right-sided tumours. Therefore, when stratifying CRC patients based on tumour location, the results indicated that tumours arising in the right colon with reduced SMAD4 expression were associated with poorer survival compared to tumours with higher SMAD4 expression (p = 0.021). In cases of left-sided tumours, no significant difference in survival trends was observed between low and high levels of SMAD4 expression. Furthermore, the combination of tumour SMAD4 expression levels and stromal percentage in this study demonstrated that the worst survival outcome occurred in patients with tumours showing low SMAD4 expression and high stroma, compared to cases with either low SMAD4 expression and low stroma or high SMAD4 expression tumours with both low and high stroma (p < 0.001). Similarly, a study of 135 patients with stage I–II colon cancer also found that worse survival was linked to high stromal content and decreased IHC SMAD4 (251). This suggests that tumours with high stroma and reduced SMAD4 expression could be classified as a high-risk group in terms of survival outcomes.

We conducted a transcriptomic analysis of CRC tumours, focusing on high stromal content, by comparing the low and high SMAD4 groups. The results showed that *SOD3* was upregulated in tumours with higher SMAD4 levels and identified several gene sets enriched in this high SMAD4 group, such as those related to the extracellular matrix, growth factor activity, and PI3K-Akt signalling. Conversely,

aminopeptidase activity was observed in the low SMAD4 group. In other studies, *SOD3*, also known as extracellular superoxide dismutase, was found to be downregulated in CRC (n = 80, $p < 0.01$) and associated with poorer survival outcomes ($p < 0.05$). Additionally, overexpression of *SOD3* in mouse models has been suggested to reduce Ki-67 expression, as well as prevent cell proliferation, tumour growth, EMT, and liver metastasis (252). Aminopeptidase N (APN) is a membrane dimeric metallopeptidase that plays a role in differentiation, development, and proliferation across various tissues. Higher Aminopeptidase N activity in CRC tissue was observed to predict better overall survival (n = 81, $p < 0.05$) (253). Nonetheless, a gene expression-based SMAD4-modulated profile in 250 colorectal cancer patients revealed that SMAD4 regulated fifteen genes within BMP/Wnt target genes, such as *JAG1*, *TCF7*, and *MYC*, which were consistently associated with SMAD4 expression in *in vitro* analysis using colorectal cancer tumouroids and predicted disease-free survival in CRC stages II and III (254). Moreover, integrated multi-omics data of SMAD4 mutant colorectal cancer revealed that missense mutations are more commonly observed in SMAD4 mutants. These patients were predicted to have poorer survival outcomes, and tumour tissues with the mutation exhibited a reduction in immune cells, such as CD4+ memory T cells and memory B cells. Additionally, many differentially expressed immune-related genes were identified compared to the non-mutated SMAD4 group (255). These could serve as potential target biomarkers in further studies to enhance our understanding of the biological processes and pathological relevance underlying the dysregulation of SMAD4 tumours with high stroma, facilitating clinical applications and the development of more advanced therapies.

In conclusion, this chapter demonstrated that SMAD4 expression is an independent prognostic factor in CRC tumours. Low SMAD4 expression was significantly associated with more aggressive features, including higher cancer stage and increased tumour stromal content. Overall, patients with low SMAD4 expression in tumours had poorer survival outcomes, especially those with tumours in the right colon. Furthermore, the worst survival outcomes were observed in patients with tumoural low SMAD4 expression and high stroma. Downregulation of *SOD3* and increased aminopeptidase activity were seen in the low SMAD4 expression group, as revealed by transcriptomic profiling of tumours with high stroma. These findings could provide potential new biomarkers for identifying and stratifying

patients for suitable therapeutic management. Moreover, combining the analysis of SMAD4 with other SMAD family members could enhance the prognostic value of TGF-β/BMP signalling pathways in CRC.

Chapter 6 Signalling mechanism alterations related to SMAD4 levels in CRC cells, synergistic effects of TGF-β and BMP signalling inhibition in combination with standard chemotherapy in a metastatic CRC cell, and developing a model of metastasis

6.1 Introduction

Metastasis of CRC indicates an advanced stage of the disease, which is associated with unfavourable survival outcomes due to the more limited treatment options available. It was highlighted that more than 20% of CRC patients develop metastasis, while fewer than 20% survive for five years after diagnosis (7). Several dysregulations of signalling pathways are observed in CRC tumours during carcinogenesis, which significantly influence the tumour's ability to spread to other sites (2, 118). Moreover, alterations in the tumour microenvironment at both the primary and the target sites facilitate metastatic CRC. However, a greater understanding of the complex mechanisms underlying CRC metastasis is necessary in order to develop more targeted treatments for patients (7, 131, 132).

Activation of the canonical TGF-β/BMP signalling pathways regulates cellular processes in tissue homeostasis through the phosphorylation of receptor-regulated SMADs, which form a complex with SMAD4 for signal transmission (162, 163). Dysregulation of these pathways, especially SMAD4, has been linked to CRC development and metastasis, and is associated with poor survival outcomes and resistance to therapy (17, 124, 164-166). Although a standard chemotherapy regimen based on 5-FU has been widely used in CRC treatment, some patients with loss of SMAD4 have been observed to respond less favourably to the drug, resulting in worse disease-free survival (165). Additionally, among the high heterogeneity of CRC tumours regarding CMS subtypes, patients in the CMS4 group, characterised by prominent TGF–β activation and stromal infiltration, were associated with the worst relapse-free survival outcomes (57). Therefore, effective therapies against dysregulated TGF-β/BMP signalling pathways are required to overcome these limitations. Inhibition of TGF-β/BMP signalling, such as type I BMP receptor inhibitors using small-molecule compounds (e.g.,

Dorsomorphin, LDN-212854, and K02288), could provide an alternative approach for treating CRC and preventing metastasis (163, 172, 175-178).

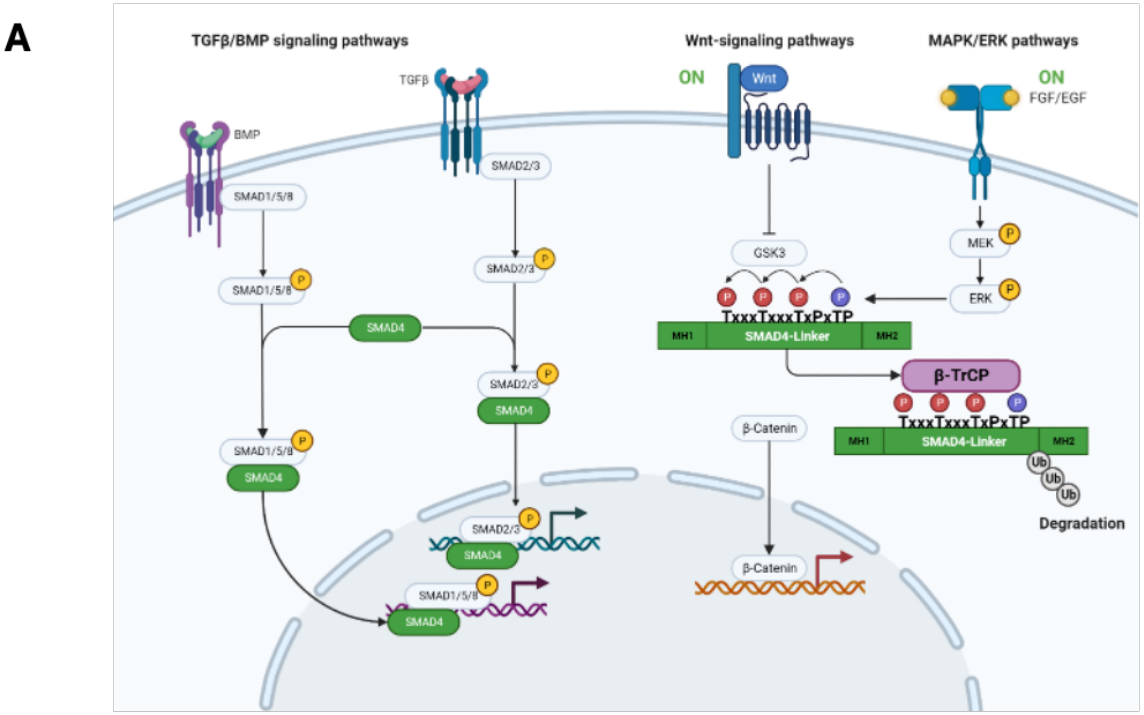
Models of human CRC are essential for understanding the altered mechanisms of CRC cells and their microenvironment, supporting preclinical research to develop new treatments. Advances in technology have made *in vitro* 3D culture more advantageous for creating models with a 3D structure and improved physiological interactions among different cells, compared to traditional 2D cultures. However, existing CRC models remain limited due to the complex biology involved, especially for predicting drug efficacy in metastasis (179, 180). Developing a model of metastatic CRC that includes other cell types, such as those in metastatic sites like the liver, lung, peritoneum and bone, remains challenging. Thus, this chapter aims to examine alterations in the signalling mechanisms involved in regulating SMAD4 activity of CRC cell lines, as well as investigate the inhibitory effects of novel inhibitors of the TGF-β and BMP signalling pathways, both alone and in combination with chemotherapeutic agents in CRC. The final part of this chapter develops a 3D model of CRC metastasis to investigate drug efficacy in a more physiological model.

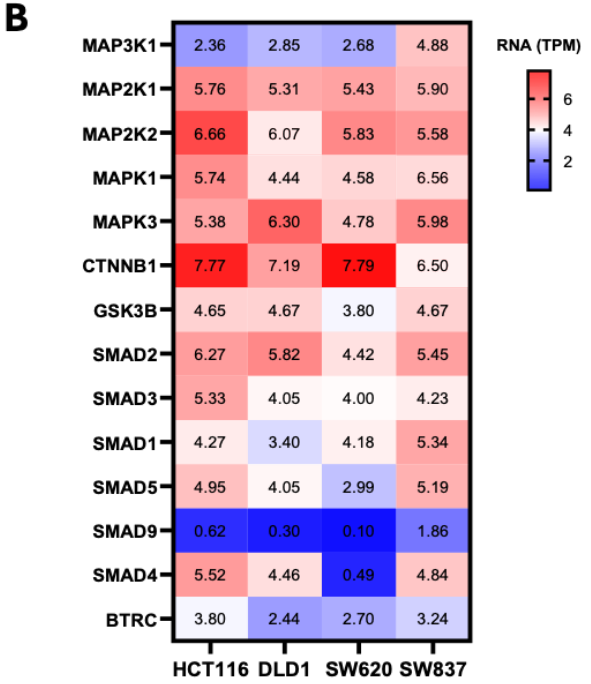
6.2 Alterations in signalling mechanisms involved in regulating SMAD4 activity.

6.2.1 Differential expression patterns of pathways which crosstalk with SMAD4 activity regulation were observed across CRC cell lines.

The functional activity of SMAD4 can be regulated through the interaction of glycogen synthase kinase-3 (GSK3) and mitogen-activated protein kinase (MAPK) phosphorylation, which co-activates Wnt and MAPK signalling pathways. Sequential GSK3 phosphorylations on the SMAD4 linker region prime the protein, inhibiting transcriptional activation and leading to polyubiquitination and degradation by the E3 ligase β-TrCP (Figure 6-1A) (256). To investigate the crosstalk pathway activities in CRC, we analysed expression levels of MAPK/ERK signalling, Wnt/β-catenin signalling, TGF-β/BMP signalling, and BTRC encoding β-TrCP in CRC cell lines, including HCT116, DLD1, SW620, and SW837, using expression data from the DepMap portal. Variations in gene expression were observed across the cell lines. For example, SW620 showed downregulated SMAD4

but upregulated MAPK and Wnt signalling pathways, whereas DLD1 showed upregulated SMAD4 and the crosstalk pathways. BTRC expression remained similar across all cell lines (Figure 6-1B). Protein expression and phosphorylation levels were also evaluated in the five CRC cell lines, including the COLO205 cell line. Different patterns of protein and phosphorylation activity were observed among the cells. Loss of SMAD4 protein was found in SW620 and COLO205; however, high activity of TGF-β/BMP signalling was present in SW620. The activation statuses of the MAPK/ERK and Wnt/β-catenin signalling pathways highlighted differences between colon and rectal cancer cells (SW837). Variations in β-TrCP activity were also identified across CRC cell lines. Additionally, we investigated the PI3K/AKT pathway, a key regulator of GSK3 and non-canonical SMAD signalling pathways, by assessing phospho-Akt substrates. Western blotting indicated activation of various substrates in each cell (Figure 6-1C).


A
TGFβ/BMP signaling pathways
Wnt-signaling pathways
MAPK/ERK pathways
TGFβ
BMP
Wnt
ON
FGF/EGF
SMAD1/5/8
SMAD2/3
SMAD4
GSK3
MEK
ERK
TxxxTxxxTxPxTP
MH1
SMAD4-Linker
MH2
β-TrCP
β-Catenin
Ub
Degradation


B
RNA (TPM)
6
4
2
MAP3K1 2.36 2.85 2.68 4.88
MAP2K1 5.76 5.31 5.43 5.90
MAP2K2 6.66 6.07 5.83 5.58
MAPK1 5.74 4.44 4.58 6.56
MAPK3 5.38 6.30 4.78 5.98
CTNNB1 7.77 7.19 7.79 6.50
GSK3B 4.65 4.67 3.80 4.67
SMAD2 6.27 5.82 4.42 5.45
SMAD3 5.33 4.05 4.00 4.23
SMAD1 4.27 3.40 4.18 5.34
SMAD5 4.95 4.05 2.99 5.19
SMAD9 0.62 0.30 0.10 1.86
SMAD4 5.52 4.46 0.49 4.84
BTRC 3.80 2.44 2.70 3.24
HCT116 DLD1 SW620 SW837

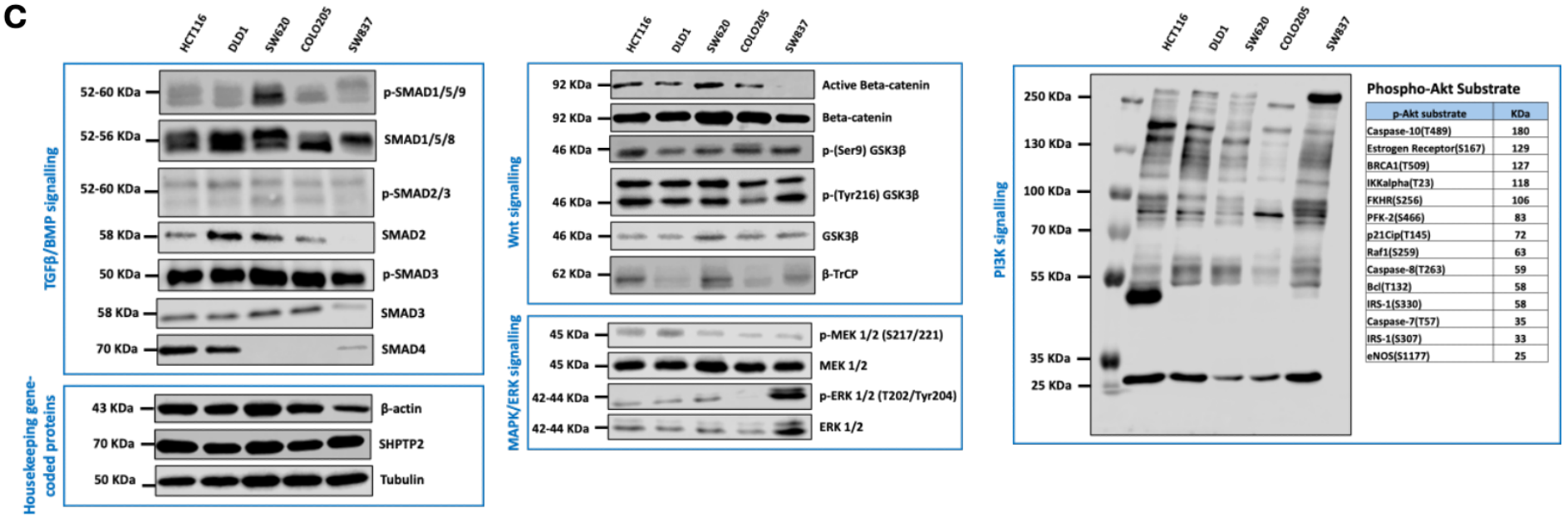

C
HCT116 DLD1 SW620 COLO205 SW837
TGFβ/BMP signalling
52-60 KDa p-SMAD1/5/9
52-56 KDa SMAD1/5/8
52-60 KDa p-SMAD2/3
58 KDa SMAD2
50 KDa p-SMAD3
58 KDa SMAD3
70 KDa SMAD4
Housekeeping gene-coded proteins
43 KDa β-actin
70 KDa SHPTP2
50 KDa Tubulin
Wnt signalling
92 KDa Active Beta-catenin
92 KDa Beta-catenin
46 KDa p-(Ser9) GSK3β
46 KDa p-(Tyr216) GSK3β
46 KDa GSK3β
62 KDa β-TrCP
MAPK/ERK signalling
45 KDa p-MEK 1/2 (S217/221)
45 KDa MEK 1/2
42-44 KDa p-ERK 1/2 (T202/Tyr204)
42-44 KDa ERK 1/2
PI3K signalling
250 KDa
130 KDa
100 KDa
70 KDa
55 KDa
35 KDa
25 KDa
Phospho-Akt Substrate
p-Akt substrate KDa
Caspase-10(T489) 180
Estrogen Receptor(S167) 129
BRCA1(T509) 127
IKKalpha(T23) 118
FKHR(S256) 106
PFK-2(S466) 83
p21Cip(T145) 72
Raf1(S259) 63
Caspase-8(T263) 59
Bcl(T132) 58
IRS-1(S330) 58
Caspase-7(T57) 35
IRS-1(S307) 33
eNOS(S1177) 25

Figure 6-1 Crosstalk pathways of TGF-β/BMP signalling, Wnt signalling, and MAPK/ERK signalling involving SMAD4 activity regulation in CRC cell lines

Schematic diagram of crosstalk pathways in SMAD4 activity regulation. This figure was created using BioRender.com and adapted from Demagny H., Araki T., and De Robertis E.M., 2014 (A). The heatmap displayed differential mRNA expression patterns of the crosstalk pathways across CRC cell lines (B). Western blotting showed varying protein levels of crosstalk pathway members across CRC cell lines, including TGF-β/BMP signalling, Wnt signalling, MAPK/ERK signalling, and PI3K/AKT signalling pathways (C).

6.2.2 Investigation of crosstalk pathways in regulating SMAD4 expression through stimulation with growth factors such as BMP4 and EGF, as well as treatment with Bortezomib.

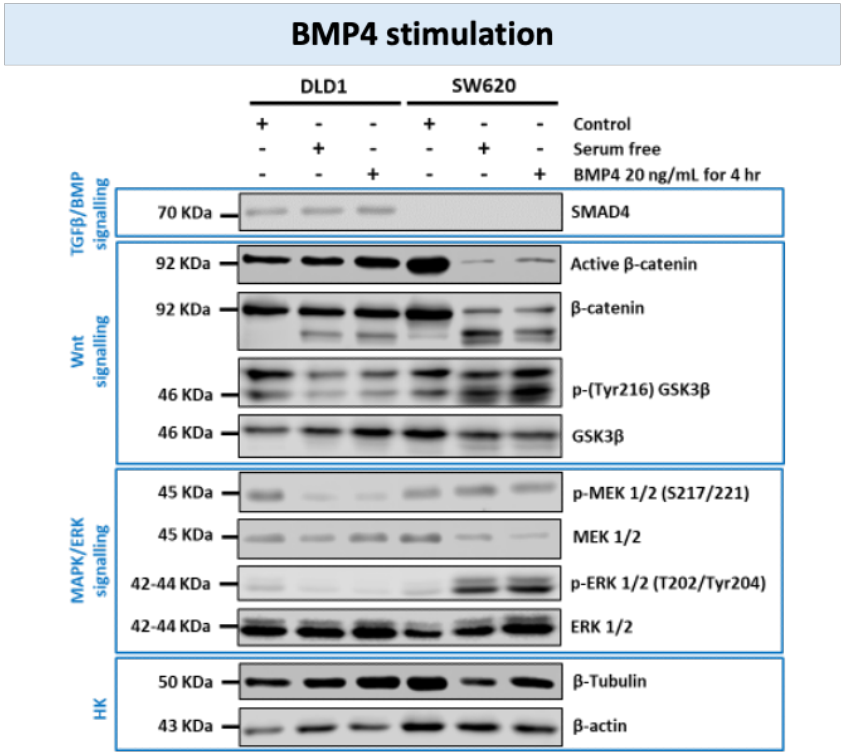
BMP4 and EGF are ligands involved in activating the BMP and EGF signalling pathways respectively, both of which are crucial mechanisms in the intestinal crypts necessary for tissue homeostasis. These pathways can activate intracellular MAPK signalling (22-30). Additionally, dysregulation of BMP and EGF signalling pathways are highlighted as contributing to the development and progression of CRC (3, 31, 32). Furthermore, our previous investigation of gene expression for TGF-β and BMP signalling using DepMap data in Chapter 3 showed that BMP4 was upregulated across CRC cell lines. To examine activation of MAPK/ERK signalling, Wnt/β-catenin signalling, TGF-β/BMP signalling in regulating SMAD4 activity, we performed a growth stimulation experiment. We used the growth factors BMP4 or EGF in two CRC cell lines: DLD1, which is a primary CRC cell line with high SMAD4 expression, and SW620, a metastatic CRC cell line with loss of SMAD4 expression, as described in Section 2.2.3.7. We also treated the cells with Bortezomib to inhibit the 20S proteasome, which is involved in the ERK signalling pathways.

After treating both cell lines with 20 ng/mL BMP4 for 4 hours, the Western blot results showed no change in SMAD4 protein levels. DLD1 cells active β-catenin levels stayed relatively constant under all conditions, with p-MEK and p-ERK levels decreasing when cells were cultured in serum starvation alone or serum starved then stimulated with either BMP-4 or EGF, compared to cell culture in serum (Figure 6-2A and B). Interestingly, when the SW620 cells were subjected to serum starvation the level of active β-catenin decreased and the level of p-ERK increased compared to cells cultured in serum. Serum starvation followed by stimulated with BMP4, resulted in a decrease in active β-catenin and an increase in p-ERK in SW620 cells similar to the levels seen when cells were grown in serum starvation (Figure

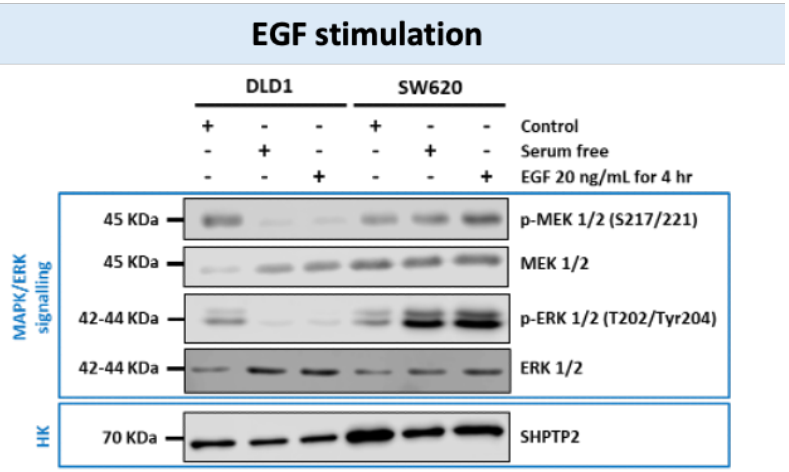
6-2A). Similarly, increased p-ERK levels were detected in SW620 cells after 4 hours of stimulation with 20 ng/mL EGF (Figure 6-2B). Bortezomib treatments at 10 and 30 μM for 24 hours on both cell lines resulted in increased β-TrCP levels; however, SMAD4 levels remained unchanged (Figure 6-2C). Overall, these results suggest that CRC have different dependencies on the signalling pathways with SW620 activating the Wnt/β-catenin pathway under stress, whereas the MAPK pathway was more active in the presence of serum and growth factor stimulation.

Due to time limitations, further investigation into the entire signalling cascades and multiple replicates would enhance understanding of these complex crosstalk mechanisms.

A


BMP4 stimulation
DLD1
SW620
Control
Serum free
BMP4 20 ng/mL for 4 hr
TGFβ/BMP signalling
70 KDa
SMAD4
Wnt signalling
92 KDa
Active β-catenin
92 KDa
β-catenin
46 KDa
p-(Tyr216) GSK3β
46 KDa
GSK3β
MAPK/ERK signalling
45 KDa
p-MEK 1/2 (S217/221)
45 KDa
MEK 1/2
42-44 KDa
p-ERK 1/2 (T202/Tyr204)
42-44 KDa
ERK 1/2
HK
50 KDa
β-Tubulin
43 KDa
β-actin

B


EGF stimulation
DLD1
SW620
Control
Serum free
EGF 20 ng/mL for 4 hr
MAPK/ERK signalling
45 KDa
p-MEK 1/2 (S217/221)
45 KDa
MEK 1/2
42-44 KDa
p-ERK 1/2 (T202/Tyr204)
42-44 KDa
ERK 1/2
HK
70 KDa
SHPTP2

C

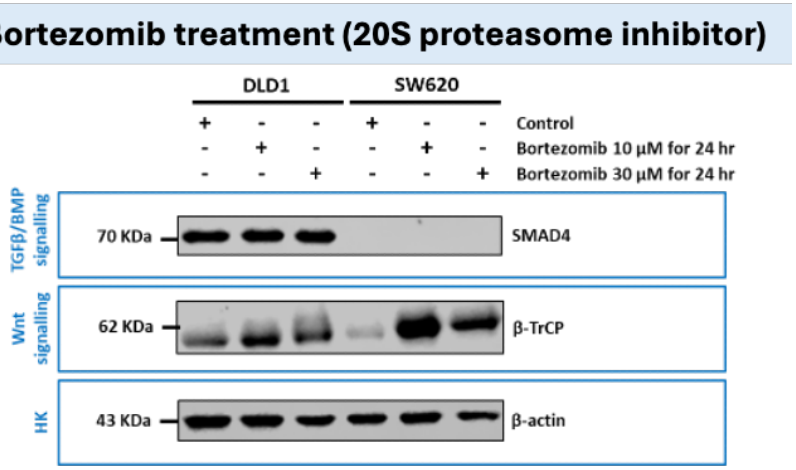

Bortezomib treatment (20S proteasome inhibitor)
DLD1
SW620
Control
Bortezomib 10 μM for 24 hr
Bortezomib 30 μM for 24 hr
TGFβ/BMP signalling
70 KDa
SMAD4
Wnt signalling
62 KDa
β-TrCP
HK
43 KDa
β-actin

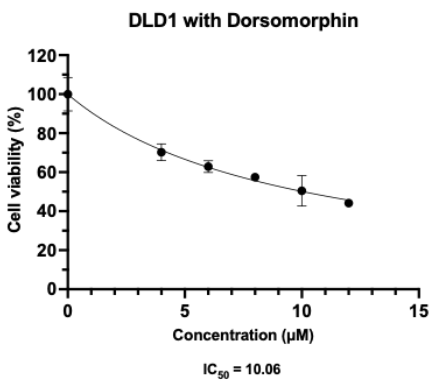
Figure 6-2 Effects of BMP4 and EGF stimulation and Bortezomib treatment on protein levels in crosstalk pathways regulating SMAD4 activity in DLD1 and SW620 cell lines

Western blotting showed that the levels of specific proteins involved in TGF-β/BMP signalling, Wnt signalling, and MAPK/ERK signalling were detected in DLD1 and SW620 cells after stimulation with 20 ng/mL of BMP4 (A) and EGF (B) for 4 hours, as well as treatment with 10 and 30 μM Bortezomib for 24 hours (C). HK is referred to as a housekeeping gene.

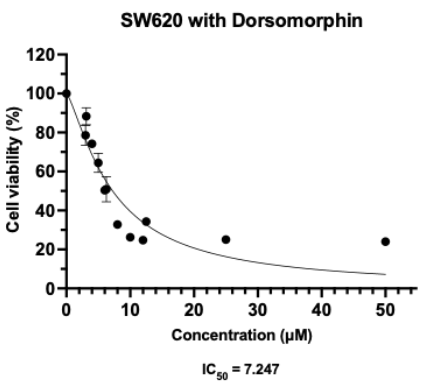
6.3 Inhibitory effects of BMPR1 inhibitors (Dor, K02288, and LDN-212854) on the TGF-β/BMP signalling pathways and a standard chemotherapeutic agent (5-FU) in CRC cell lines.

In addition to the signalling ligands, BMP receptors such as BMPR1A were also found to be upregulated across CRC cells, as shown by DepMap data analysis in Chapter 3. Some studies suggested that dysregulation of BMPR1A is related to CRC carcinogenesis (22, 23). Furthermore, the TGF-β/BMP signalling pathways interact with other key signalling pathways, such as the Wnt signalling pathway, which is involved in the complex process of contributing to CRC development and progression (2, 17, 18, 124, 125). Blocking signal activation of the TGF-β/BMP signalling in CRC cells may alter their cellular processes, especially those that promote cancer aggressiveness. Dorsomorphin (Dor), and two new derivatives of Dor; K02288, and LDN-212854 are small molecules designed to inhibit the kinase activity of BMP type I receptors (ALK1-3 and 6) (175-178). To evaluate their inhibitory effects in CRC, DLD1 and SW620 cell lines were treated with these compounds for 72 hours, followed by assessment of cell viability using the WST-1 Proliferation assay (as described in Section 2.2.3.4). IC50 values for each drug were determined in DLD1 cells: 10.06 μM for Dor, 28.07 μM for K02288, and 3.04 μM for LDN-212854 (Figure 6-3A, C, and E, respectively). Likewise, in SW620, the IC50 values were approximately 7.25 μM for Dor, 11.64 μM for K02288, and 1.85 μM for LDN-212854 (Figure 6-3B, D, and F, respectively). Furthermore, the inhibitory effects of 5-FU on cell viability were confirmed in both cell lines to determine the IC50 for subsequent experiments. Results showed less response to 5-FU after 48 hours of exposure, while 72-hour treatment decreased cell viability: IC50 = 18.31 μM for DLD1 and 18.95 μM for SW620 (Figure 6-3G and H, respectively).

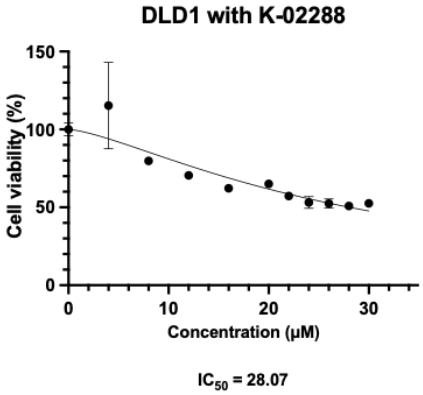
A


DLD1 with Dorsomorphin
Cell viability (%)
Concentration (μM)
IC50 = 10.06

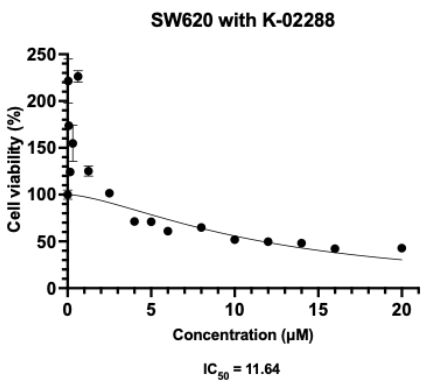
B


SW620 with Dorsomorphin
Cell viability (%)
Concentration (μM)
IC50 = 7.247

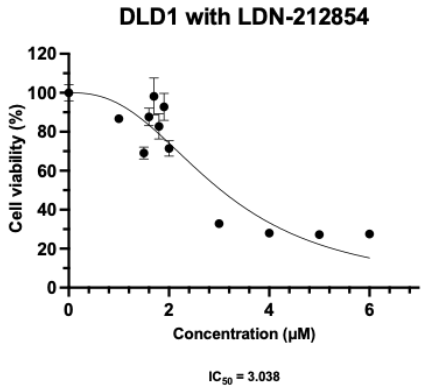
C


DLD1 with K-02288
Cell viability (%)
Concentration (μM)
IC50 = 28.07

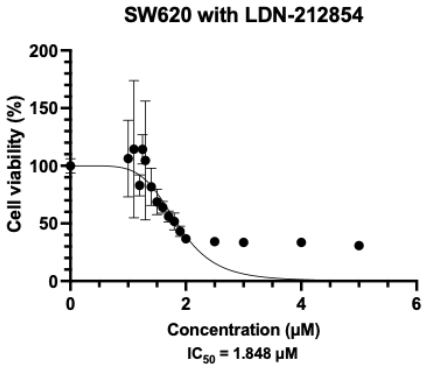
D


SW620 with K-02288
Cell viability (%)
Concentration (μM)
IC50 = 11.64

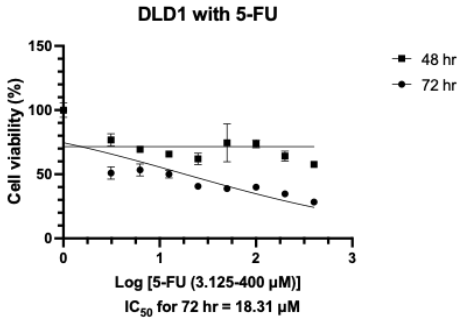
E


DLD1 with LDN-212854
Cell viability (%)
Concentration (μM)
IC50 = 3.038

F


SW620 with LDN-212854
Cell viability (%)
Concentration (μM)
IC50 = 1.848 μM

G


DLD1 with 5-FU
Cell viability (%)
48 hr
72 hr
Log [5-FU (3.125-400 μM)]
IC50 for 72 hr = 18.31 μM

H

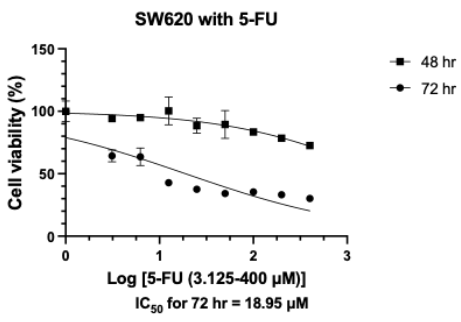

SW620 with 5-FU
Cell viability (%)
48 hr
72 hr
Log [5-FU (3.125-400 μM)]
IC50 for 72 hr = 18.95 μM

Figure 6-3 Inhibitory effects on cell viability of BMPR1 inhibitors (Dor, K02288, and LDN-212854) and a standard chemotherapeutic agent (5-FU) in DLD1 and SW620 cell lines

The graphs showed decreasing cell viability of DLD1 and SW620 cells after treatment with Dor (A and B), K02288 (C and D), and LDN-212854 (E and F) at 72 hours, analysed using the WST-1 assay, and IC50 values for each drug were calculated. Additionally, the inhibitory effects of 5-FU on cell viability were assessed at 48 and 72 hours, with a stronger response observed at the longer time point in DLD1 and SW620 cells (G and H).

6.4 Targeting TGF-β/BMP signalling pathways both alone and in combination with chemotherapy.

6.4.1 Synergistic effects of the drug combination between LDN-212854 and 5-FU on cell viability.

Overall, among the three BMP type I receptor inhibitors, LDN-212854 appeared to have a more potent effect on CRC cell viability. Additionally, when treating the two cell lines with 5-FU for 72 hours, a better drug response was observed compared to 48 hours. We then decided to use these two drugs to repeat the experiment, aiming to obtain more precise results with drug concentrations at 0.5X, 1X, and 2X of the IC50 at 48 and 72 hours. Furthermore, we evaluated the effects of the drug combination using the Chou-Talalay method, as described in Section 2.2.3.5. Representative images of DLD1 cells showed cell morphology from the vehicle as a reference (Figure 6-4A), which changed after treatment with 20 μM 5-FU, 3 μM LDN-212854, and the combination of 20 μM 5-FU and 3 μM LDN-212854 for 72 hours (Figure 6-4B, C, and D, respectively). The inhibitory effects of 5-FU were not detected at 48 hours, and an IC50 of 101.4 μM was identified in DLD1 cells treated with 5-FU (Figure 6-4E). During LDN-212845 treatment in DLD1 cells, results showed IC50 values of 4.05 μM and 2.324 μM at 48 and 72 hours, respectively (Figure 6-4F).

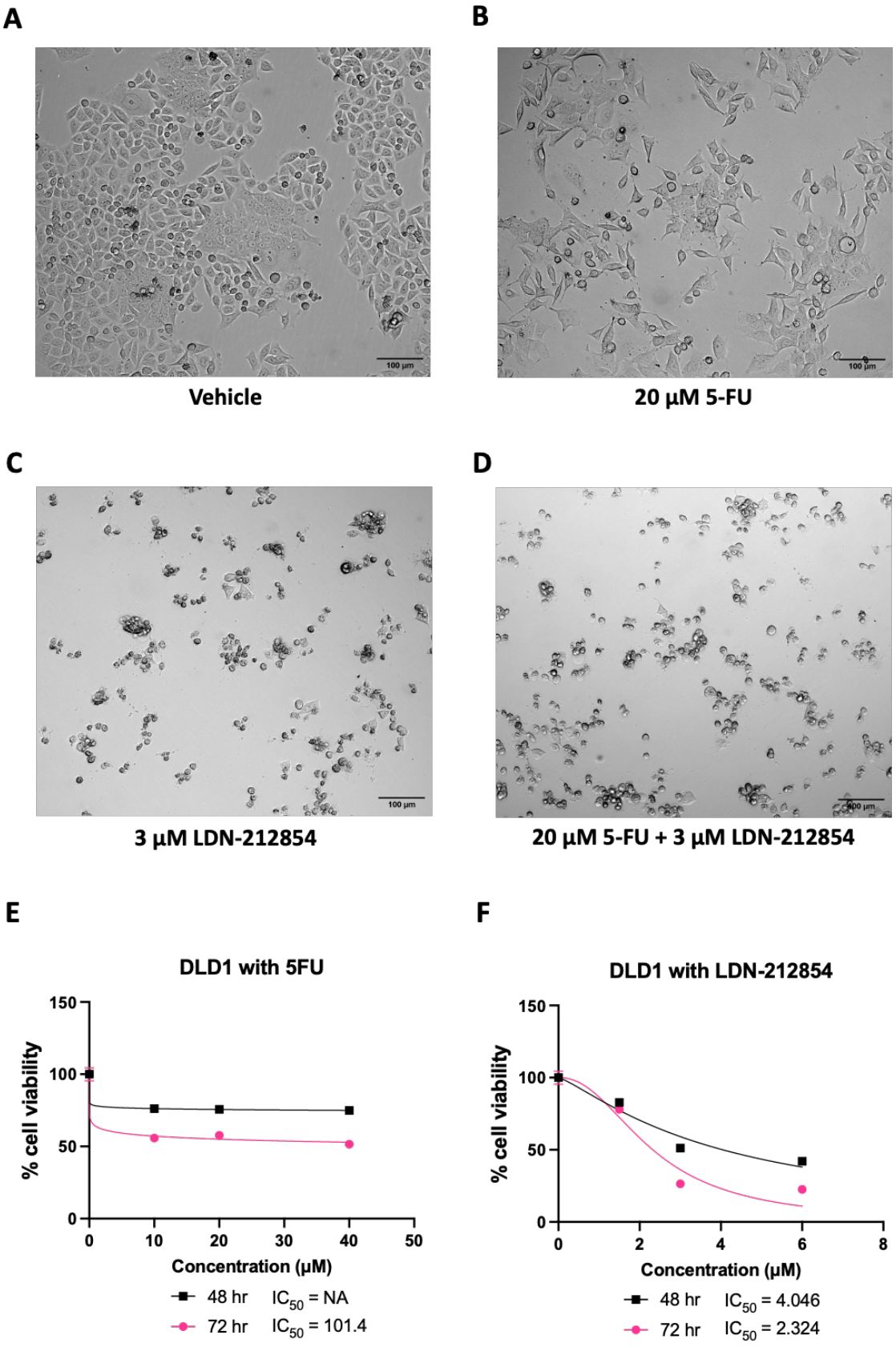


Figure 6-4 Inhibitory effects on cell viability of LDN-212854 and 5-FU in DLD1 cell line

Images demonstrated the morphology of DLD1 cells treated with vehicle (A), 20 µM 5-FU (B), 3 µM LDN-212854 (C), and a combination of 20 µM 5-FU and 3 µM LDN-212854 (D) after 72 hours. The graphs showed that the inhibitory effects of 5-FU (E) and LDN-212854 (F) on cell viability were assessed at 48 and 72 hours using the WST-1 assay, with a greater reduction in cell viability

observed at the longer time point and lower IC50 values calculated. The scale bar = 100 μm in Figures A-D.

For the treatment in SW620 cells, representative images showing the cell morphology of SW620 cells in the vehicle condition (Figure 6-5A) and changes under treatment conditions with 20 μM 5-FU, 2 μM LDN-212854, and the combination of 20 μM 5-FU and 2 μM LDN-212854 for 72 hours (Figure 6-5B, C, and D, respectively). The cell viability results showed a less responsive effect to 5-FU and LDN-212854 at 48-hour treatment, with IC50 values of 110.9 μM and 2.73 μM, respectively. Better response was observed after 72 hours of treatment, with IC50 values of 19.74 μM for 5-FU and 1.69 μM for LDN-212854 (Figure 6-5E and F).

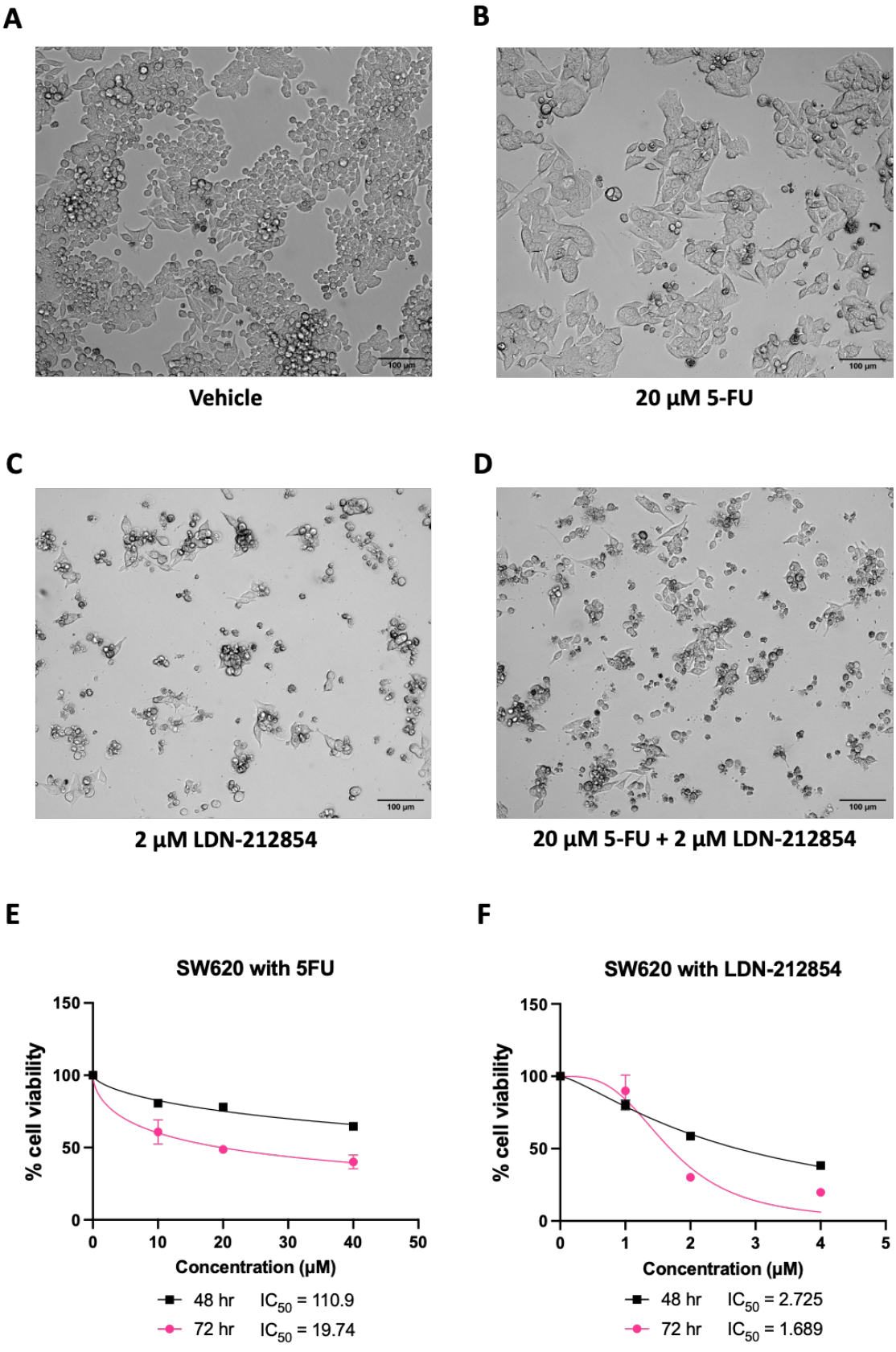


Figure 6-5 Inhibitory effects on cell viability of LDN-212854 and 5-FU in SW620 cell line

Images demonstrated the morphology of SW620 cells treated with vehicle (A), 20 µM 5-FU (B), 2 µM LDN-212854 (C), and a combination of 20 µM 5-FU and 2 µM LDN-212854 (D) after 72 hours. The graphs showed that the inhibitory effects of 5-FU (E) and LDN-212854 (F) on cell viability were assessed at 48 and 72 hours using the WST-1 assay, with a greater reduction in cell viability

observed at the longer time point and lower IC50 values calculated. The scale bar = 100 μm in Figures A-D.

Subsequently, the pharmacodynamic drug interaction results showed that all CI values for treatment with 3 μM LDN-212854 and 10 to 40 μM 5-FU on DLD1 cells over 48 and 72 hours were below 1, indicating synergism between these two drugs in combination. Treating the cells with 20 μM 5-FU and 1.5 to 3 μM LDN-212854 demonstrated synergistic effects; however, when the LDN-212854 concentration increased to 6 μM, both additivity (CI = 1) and antagonism (CI > 1) were observed over time. The Fa value at 0.5 was calculated, indicating 50% inhibition. The results revealed that approximately 2.5 μM LDN-212854 combined with 103.23 μM 5-FU was necessary for a 72-hour treatment. Although 4.13 μM LDN-212854 was determined for 48-hour treatment, the specific dose of 5-FU was not identified (Figure 6-6A). Similarly, the combination treatment on SW620 cells with 2 μM LDN-212854 and 10 to 40 μM 5-FU showed synergistic effects over 48 and 72 hours (all CI < 1). Synergism was also observed with the combined treatment of 20 μM 5-FU and 1 to 2 μM LDN-212854, while additivity and antagonism appeared in 20 μM 5-FU combined with 4 μM LDN-212854 at the two time points, respectively. At Fa of 0.5, 2.73 μM LDN-212854 and 127.62 μM 5-FU were required for half-inhibition effects at 48 hours of treatment. The doses of 1.99 μM LDN-212854 and 19.79 μM 5-FU were determined for 72-hour treatment (Figure 6-6B).

After obtaining preliminary results on the effects of the drug combination on cell viability using the WST-1 assay, we confirmed the optimal drug concentrations for subsequent experiments by measuring cell viability between 48 and 96 hours with the Trypan blue exclusion assay as described in Section 2.2.3.2. The DLD1 cells were treated with 20 μM 5-FU and 3 μM LDN-212854, either alone or in combination. The results showed a significant decrease in the percentage of viable DLD1 cells across all treatment conditions and time points compared to the vehicle (Figure 6-6C). Similarly, the number of viable SW620 cells declined in all treatment conditions and time points (Figure 6-6D). Overall, nonviable DLD1 and SW620 cells did not show a significant increase in cell numbers (Figure 6-6E and F, respectively).

A

Pharmacodynamic drug interaction (DLD1)

Treatment		48 hr		72 hr	
Dose LDN	Dose 5FU	Effect	CI	Effect	CI
3.0	10.0	0.51119	0.70372	0.73454	0.68268
3.0	20.0	0.52645	0.67287	0.74321	0.66594
3.0	40.0	0.51780	0.69019	0.74344	0.66552
Dose LDN	Dose 5FU	Effect	CI	Effect	CI
1.5	20.0	0.34842	0.5755	0.57826	0.51819
3.0	20.0	0.51628	0.69329	0.74602	0.66051
6.0	20.0	0.59613	1.09309	0.79147	1.14639

Combination index (CI) theorem	
Synergism	CI <1
Additivity	CI =1
Antagonism	CI >1

CI values at:ComboED50
Fraction affected (Fa) = 0.5

Treatment	Fa	Dose LDN	Dose 5FU
48 hr	0.5	4.12535	3.99E+11
72 hr	0.5	2.50319	103.228

B

Pharmacodynamic drug interaction (SW620)

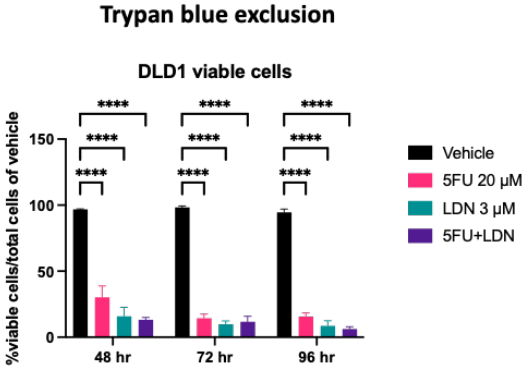
Treatment		48 hr		72 hr	
Dose LDN	Dose 5FU	Effect	CI	Effect	CI
2.0	10.0	0.51454	0.77165	0.72063	0.80346
2.0	20.0	0.51894	0.82949	0.72933	0.88288
2.0	40.0	0.53348	0.91254	0.77582	0.88326
Dose LDN	Dose 5FU	Effect	CI	Effect	CI
1.0	20.0	0.43877	0.67561	0.70048	0.61109
2.0	20.0	0.54626	0.75242	0.74582	0.8348
4.0	20.0	0.63348	1.03902	0.80467	1.26175

Combination index (CI) theorem	
Synergism	CI <1
Additivity	CI =1
Antagonism	CI >1

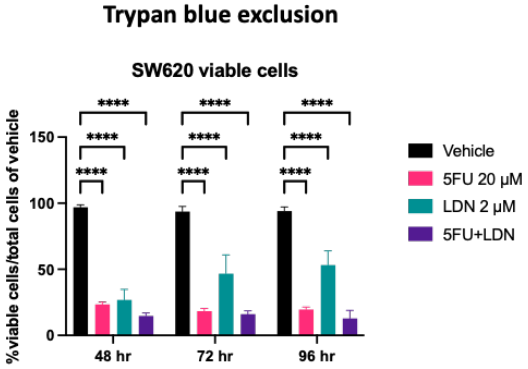
CI values at:ComboED50
Fraction affected (Fa) = 0.5

Treatment	Fa	Dose LDN	Dose 5FU
48 hr	0.5	2.73473	127.621
72 hr	0.5	1.98639	19.791

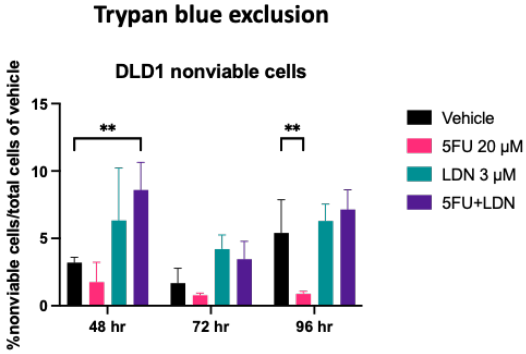
C



D



E



F

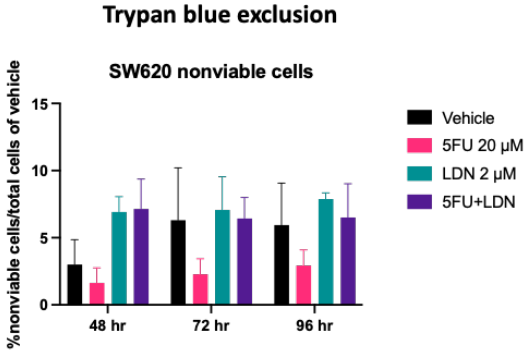


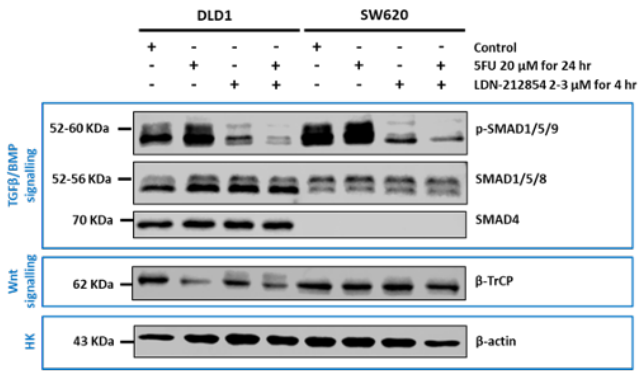
Figure 6-6 Synergistic effects of the drug combination between LDN-212854 and 5-FU on cell viability of DLD1 and SW620 cell lines

The table displayed the pharmacodynamic drug interaction of 5-FU combined with LDN-212854 on the DLD1 (A) and SW620 (B) cell lines after 48 and 72 hours of treatment. The combination index (CI) indicated: CI < 1 (Synergism), CI = 1 (Additivity), and CI > 1 (Antagonism). Histograms comparing the percentage of viable and non-viable cells, as determined by trypan blue exclusion, showed a decrease in viable cells after treatment across all conditions between 48 and 96 hours in DLD1 cells (C) and SW620 cells (D). Meanwhile, the number of non-viable cells was not related to the differences observed in all treatment conditions in DLD1 cells (E) and SW620 cells (F). Dunnett's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, mean ± SD).

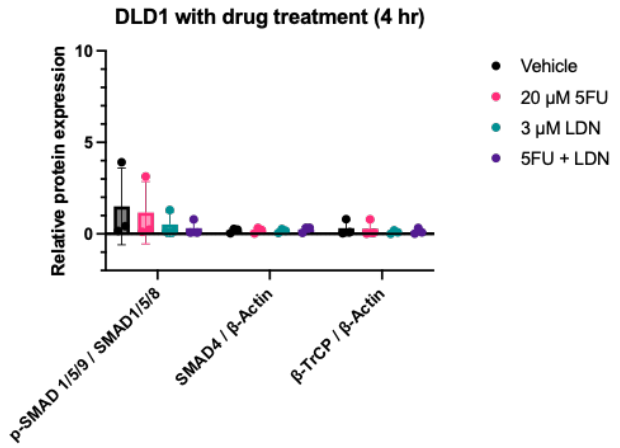
6.4.2 Inhibition effects of the drug combination between LDN-212854 and 5-FU on protein levels of signalling mechanisms

Regarding the inhibition of TGF-ß/BMP signalling pathways by targeting the type I BMP receptor with LDN-212854 and the cytotoxic effects of 5-FU, next we examined the on-target effects of these drug treatments. DLD1 and SW620 cell lines were treated, and protein expression changes were evaluated using Western blotting as described in section 2.2.6. DLD1 cells were exposed to 20 µM 5-FU, 3 µM LDN-212854, and the combination of 20 µM 5-FU and 3 µM LDN-212854, while SW620 cells were treated with 20 µM 5-FU, 2 µM LDN-212854, and the combination of 20 µM 5-FU and 2 µM LDN-212854. After 4 hours of treatment, the protein levels of p-SMAD1/5/9, SMAD1/5/8, SMAD4, and ß-TrCP were assessed. A decrease in p-SMAD1/5/9 levels was observed in both cell types with LDN-212854 and the drug combination, although the results were not statistically significant due to high standard deviation. Levels of SMAD4 and ß-TrCP remained unchanged (Figure 6-7A to C). Nonetheless, a notable reduction in the phosphorylation levels of SMAD1/5/9 was seen after 24 hours of LDN-212854 treatment, evidenced by a significant difference in DLD1 with the drug combination, as well as in SW620 with either LDN-212854 alone or in combination, compared to the vehicle. These did not affect SMAD4 expression levels (Figure 6-7D to F). Moreover, the effects of 5-FU on inducing cell death were investigated by examining increased phosphorylated H2AX in DLD1 cells treated with either 5-FU alone or in combination, compared to the vehicle at 48 hours. While a higher level of cleaved Caspase-3 was observed in SW620 following 5-FU treatment, although a statistically significant difference was not observed (Figure 6-7G to I). A longer treatment time point may offer more precise results.

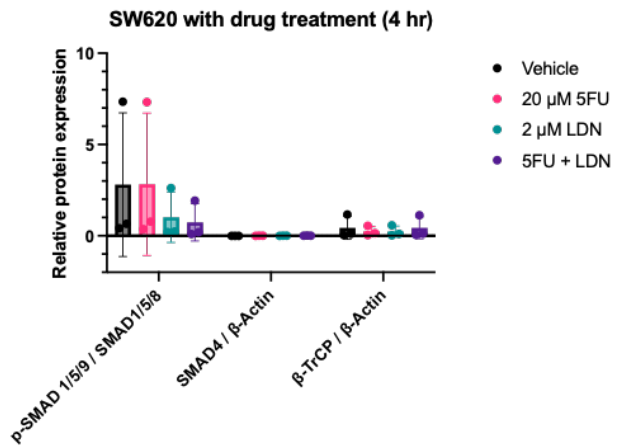
A


DLD1
SW620
Control
5FU 20 μM for 24 hr
LDN-212854 2-3 μM for 4 hr
52-60 KDa
p-SMAD1/5/9
52-56 KDa
SMAD1/5/8
70 KDa
SMAD4
62 KDa
β-TrCP
43 KDa
β-actin
TGFβ/BMP signalling
Wnt signalling
HK

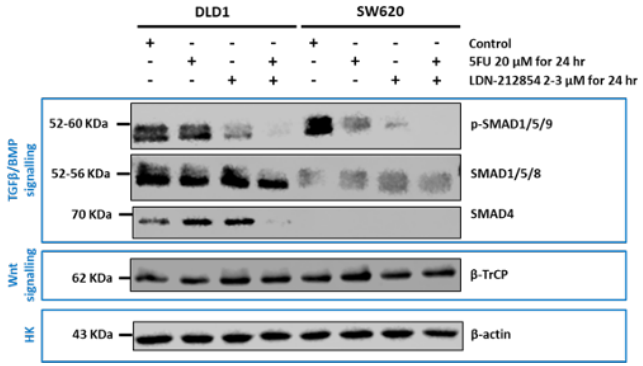
B


DLD1 with drug treatment (4 hr)
Relative protein expression
Vehicle
20 μM 5FU
3 μM LDN
5FU + LDN
p-SMAD 1/5/9 / SMAD1/5/8
SMAD4 / β-Actin
β-TrCP / β-Actin

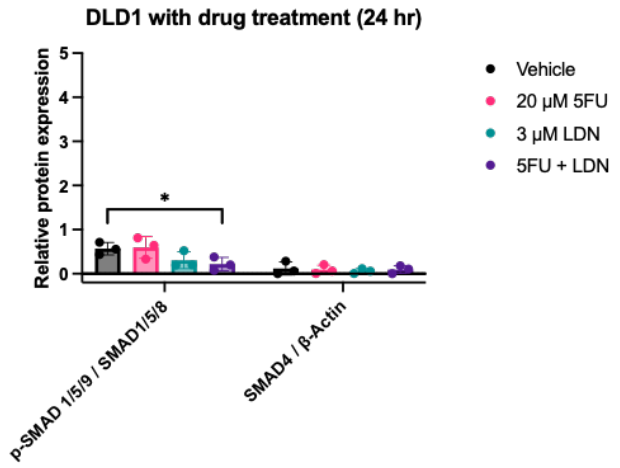
C


SW620 with drug treatment (4 hr)
Relative protein expression
Vehicle
20 μM 5FU
2 μM LDN
5FU + LDN
p-SMAD 1/5/9 / SMAD1/5/8
SMAD4 / β-Actin
β-TrCP / β-Actin

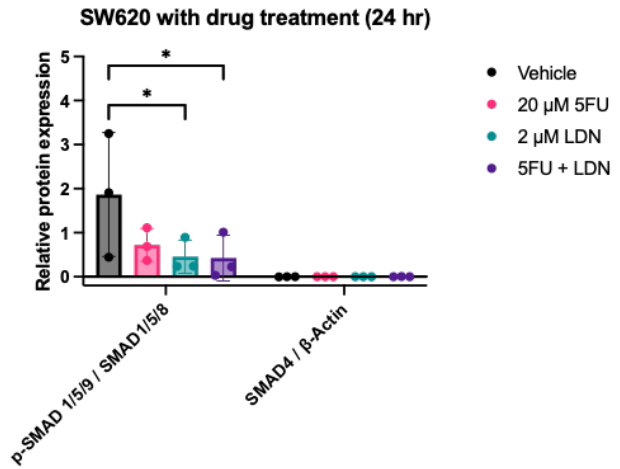
D


DLD1
SW620
Control
5FU 20 μM for 24 hr
LDN-212854 2-3 μM for 24 hr
52-60 KDa
p-SMAD1/5/9
52-56 KDa
SMAD1/5/8
70 KDa
SMAD4
62 KDa
β-TrCP
43 KDa
β-actin
TGFβ/BMP signalling
Wnt signalling
HK

E


DLD1 with drug treatment (24 hr)
Relative protein expression
Vehicle
20 μM 5FU
3 μM LDN
5FU + LDN
*
p-SMAD 1/5/9 / SMAD1/5/8
SMAD4 / β-Actin

F


SW620 with drug treatment (24 hr)
Relative protein expression
Vehicle
20 μM 5FU
2 μM LDN
5FU + LDN
*
*
p-SMAD 1/5/9 / SMAD1/5/8
SMAD4 / β-Actin

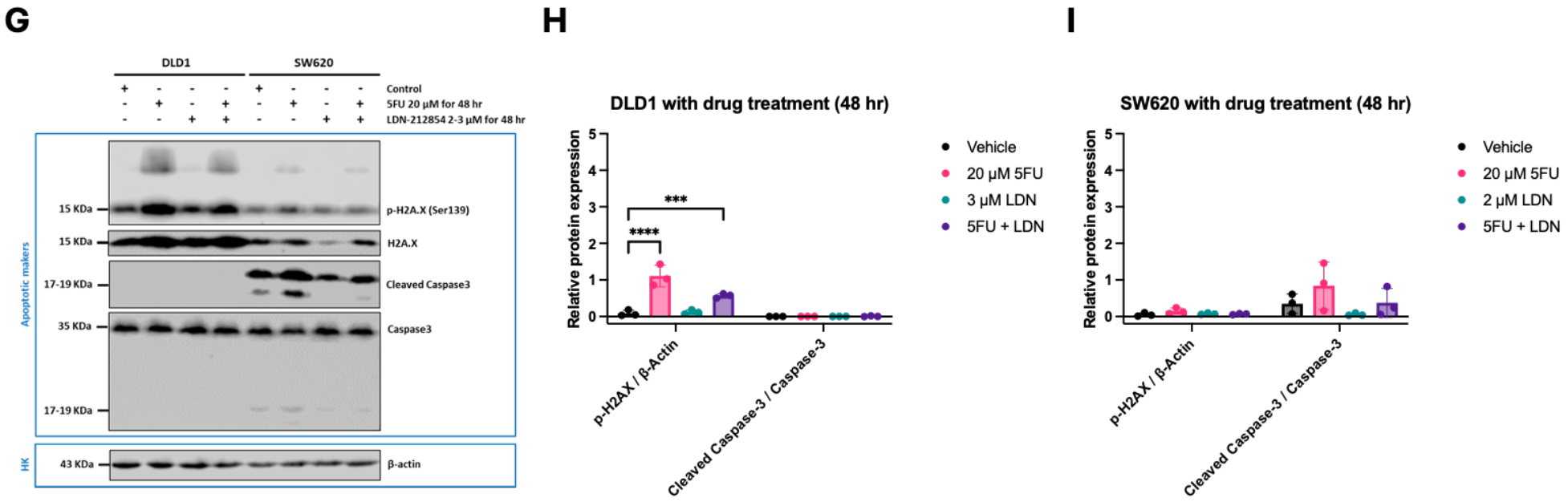
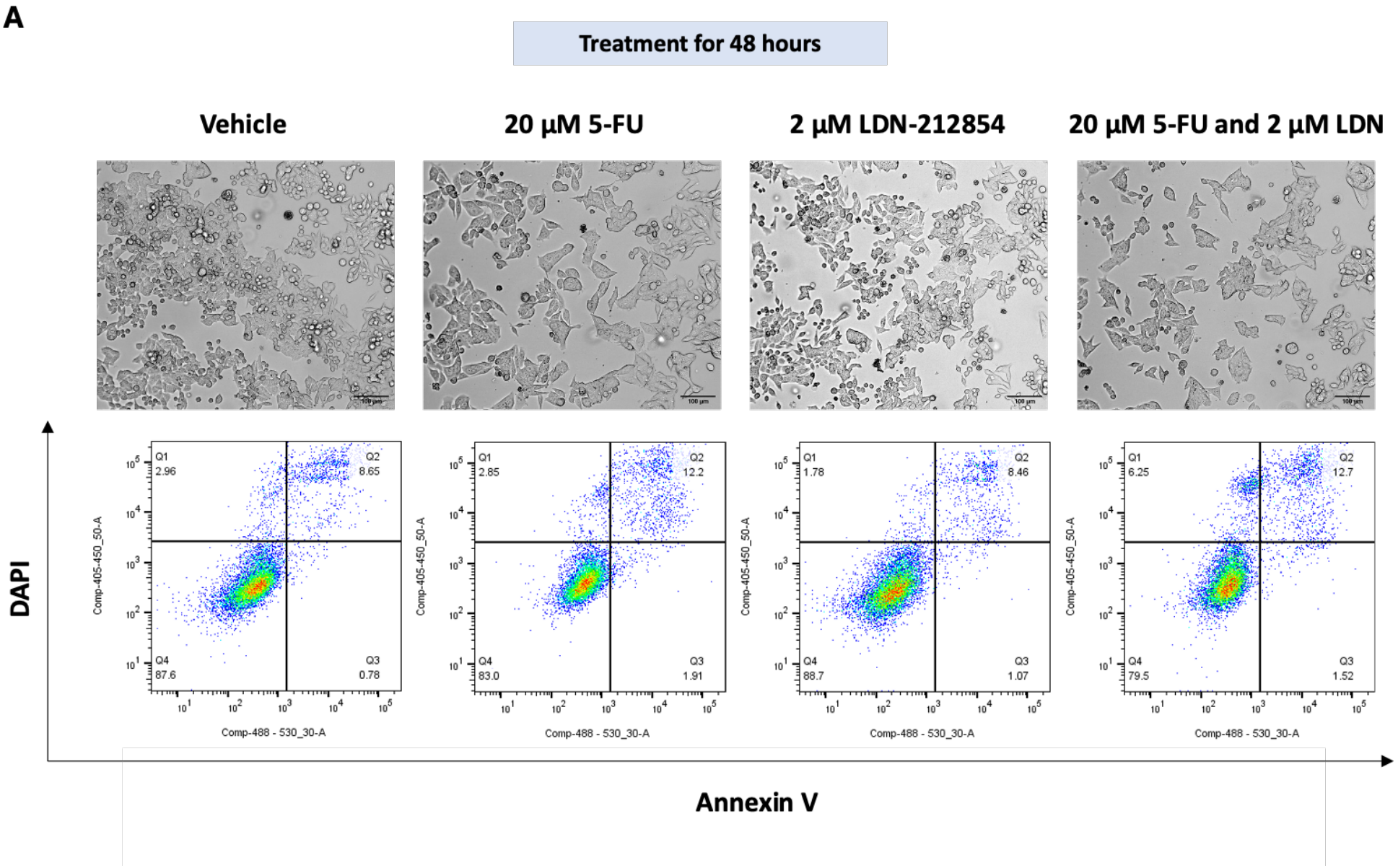


Figure 6-7 Inhibition effects of the drug combination between LDN-212854 and 5-FU on protein levels of signalling mechanisms

Western blotting results showed a decrease in SMAD1/5/9 phosphorylation after treating DLD1 and SW620 cells with LDN-212854 alone and in combination for 4 hours. There was no change in SMAD4 and β-TrCP levels (A). The histograms from densitometry showed the relative protein expression levels after 4-hour treatment in DLD1 cells (B) and SW620 cells (C). Western blotting results showed dephosphorylation of SMAD1/5/9 in DLD1 and SW620 cells after treatment with LDN-212854 alone and in combination for 24 hours. There was no change in SMAD4 levels (D). The histograms from densitometry displayed the relative protein expression levels after 24-hour treatment in DLD1 cells (E) and SW620 cells (F). Western blotting results showed detection of apoptotic markers, p-H2AX and cleaved caspase-3 in DLD1 and SW620 cells after treatment with 5-FU alone and in combination for 48 hours (G). The histograms from densitometry displayed the relative protein expression levels after 48-hour treatment in DLD1 cells (E) and SW620 cells (F).

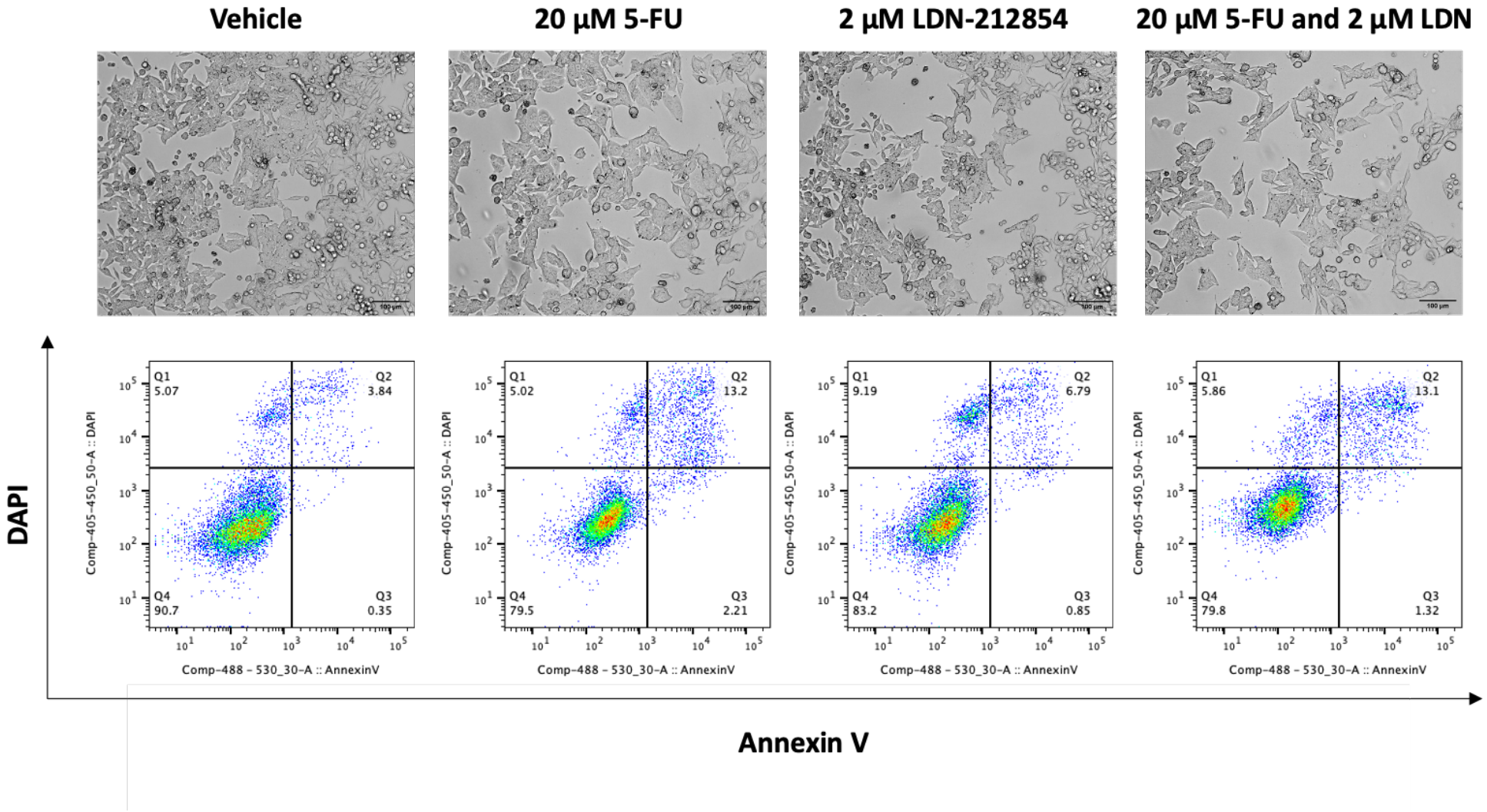
6.4.3 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on apoptosis

Next, to evaluate the effects of the drug treatment alone and in combination, on inducing apoptosis, in a metastatic CRC cell line, SW620 cells were exposed to 20 μM 5-FU, 2 μM LDN-212854, and the drug combination for 48 and 72 hours. The percentage of apoptotic cells (FITC Annexin V+) was then measured using FACS, as described in Section 2.2.4.1. Representative images of cell morphology (above) and corresponding FlowJo™ density plots (below) under different treatment conditions showed changes in cell morphology and cell populations in each apoptotic cell state (FITC+/- and DAPI+/-) over 48 and 72 hours of treatment (Figure 6-8A and B, respectively). Additionally, the trypan blue exclusion assay was performed to assess the number of viable and non-viable cells. These preliminary results revealed a reduction in cell viability in the drug treatment groups compared to the vehicle control (Figure 6-8C). Furthermore, the number of non-viable cells under drug treatment increased by a very slight amount, less than 5% (Figure 6-8D). According to the FACS results, most cells were detected as viable (FITC- and DAPI-), while approximately 15% or fewer were identified as early apoptotic (FITC+ and DAPI-), late apoptotic (FITC+ and DAPI+), or necrotic (FITC- and DAPI+) at both 48 hours (Figure 6-8E) and 72 hours (Figure 6-8F). These findings were obtained from only one replicate due to a low detectable rate of apoptotic cells. A higher drug concentration may improve detection of inhibitory effects on apoptosis.

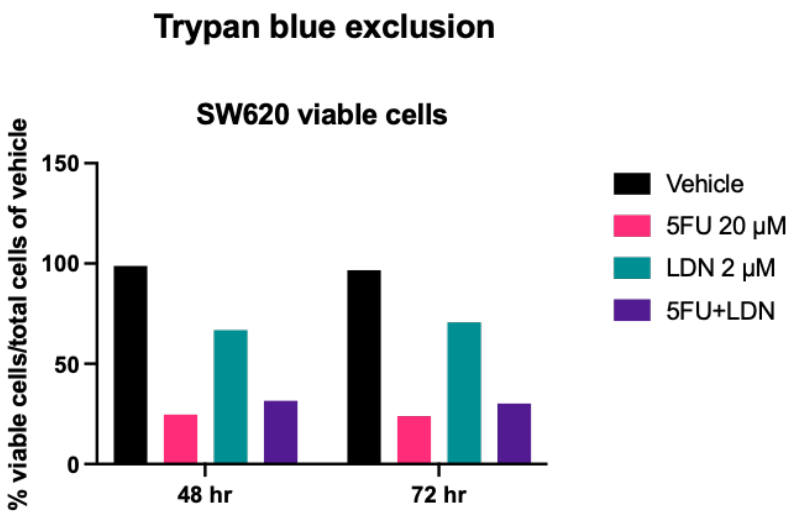

A
Treatment for 48 hours
Vehicle
20 µM 5-FU
2 µM LDN-212854
20 µM 5-FU and 2 µM LDN
DAPI
Annexin V
Comp-405-450_50-A
Comp-488 - 530_30-A
Q1 2.96
Q2 8.65
Q4 87.6
Q3 0.78
Q1 2.85
Q2 12.2
Q4 83.0
Q3 1.91
Q1 1.78
Q2 8.46
Q4 88.7
Q3 1.07
Q1 6.25
Q2 12.7
Q4 79.5
Q3 1.52

B

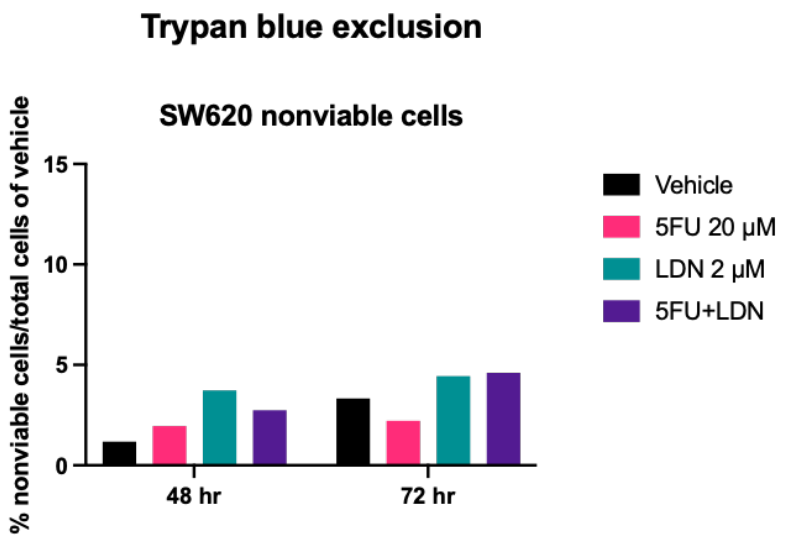
Treatment for 72 hours



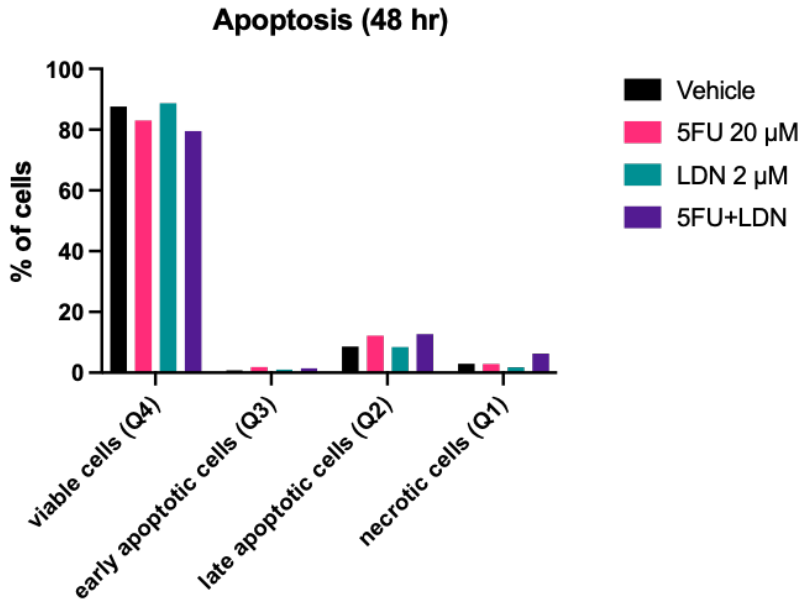
C


Trypan blue exclusion
SW620 viable cells
% viable cells/total cells of vehicle
150
100
50
0
48 hr
72 hr
Vehicle
5FU 20 μM
LDN 2 μM
5FU+LDN

D


Trypan blue exclusion
SW620 nonviable cells
% nonviable cells/total cells of vehicle
15
10
5
0
48 hr
72 hr
Vehicle
5FU 20 μM
LDN 2 μM
5FU+LDN

E


Apoptosis (48 hr)
% of cells
100
80
60
40
20
0
viable cells (Q4)
early apoptotic cells (Q3)
late apoptotic cells (Q2)
necrotic cells (Q1)
Vehicle
5FU 20 μM
LDN 2 μM
5FU+LDN

F

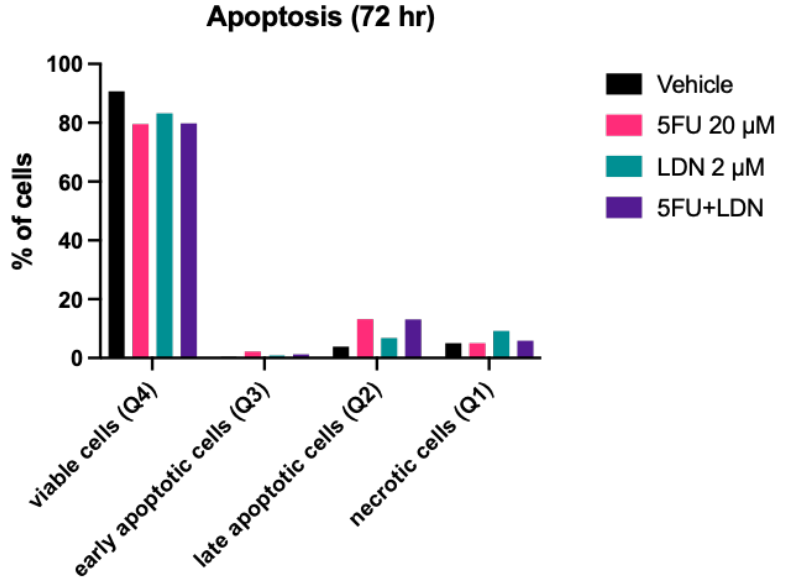
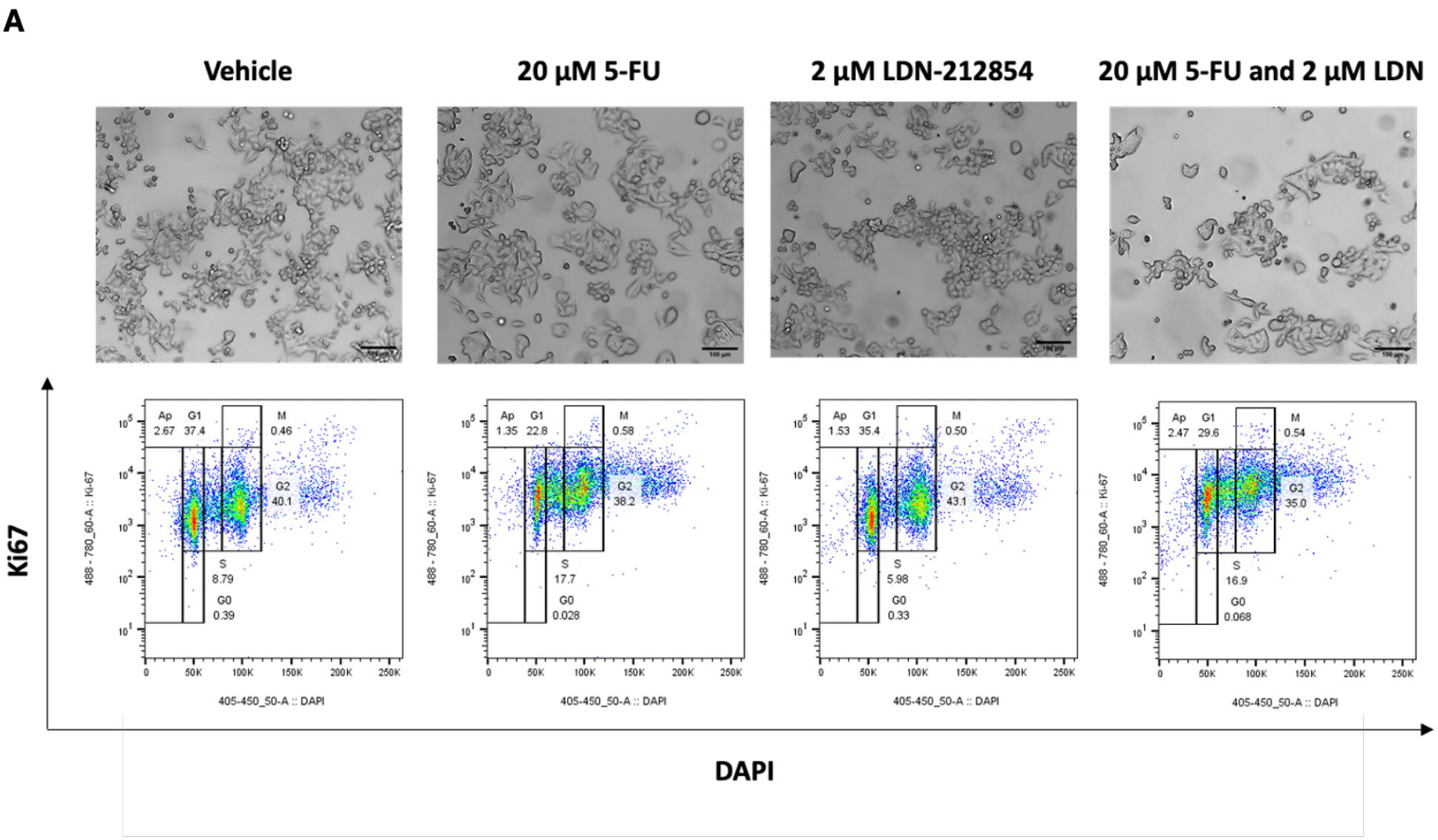

Apoptosis (72 hr)
% of cells
100
80
60
40
20
0
viable cells (Q4)
early apoptotic cells (Q3)
late apoptotic cells (Q2)
necrotic cells (Q1)
Vehicle
5FU 20 μM
LDN 2 μM
5FU+LDN

Figure 6-8 Inhibition effects of the drug combination between LDN-212854 and 5-FU on apoptosis analysis of the SW620 cell line

Images illustrated the morphology (above) and the corresponding FlowJo™ density plots (below) of SW620 cells treated with vehicle, 20 μM 5-FU, 2 μM LDN-212854, and a combination of 20 μM 5-FU and 2 μM LDN-212854 after 48 hours (A) and 72 hours (B). The histogram from the trypan blue exclusion assay showed a decrease in the percentage of viable cells across treatment conditions compared to the vehicle (C). A slight increase in non-viable cells was observed (D). The histogram from apoptosis analysis revealed variation in the percentage of cells in each apoptotic state (FITC+/- and DAPI+/-) under different treatment conditions, including viable cells (FITC- and DAPI-), early apoptotic (FITC+ and DAPI-), late apoptotic (FITC+ and DAPI+), and necrotic (FITC- and DAPI+) at 48 hours (E) and 72 hours (F), (n = 1). The scale bar = 100 μm in Figures A and B.

6.4.4 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell cycle analysis

We examined the inhibitory effects on the cell cycle of 20 μM 5-FU, 2 μM LDN-212854, and the drug combination in SW620 cells over 48 hours using FACS with Ki-67 cell cycle analysis, as described in Section 2.2.4.2. Representative images of cell morphology (above) and corresponding FlowJo™ density plots (below) under treatment conditions showed changes in cell morphology and cell populations throughout the 48-hour period (Figure 6-9A). The number of viable and non-viable cells was counted using the Trypan blue exclusion assay. The results indicated a significant decrease in the percentage of viable cells after treatment with both drugs alone and in combination. However, there was no difference in the percentage of non-viable cells across conditions (Figure 6-9B). Furthermore, the cell cycle analysis revealed that treating the cells with 20 μM 5-FU caused cell cycle arrest at the S phase, while most of the 2 μM LDN-212854-treated cells were in the G1 and G2 phases. Additionally, a similar cell cycle pattern to that seen with 5-FU alone was observed in cells treated with the drug combination, which also induced cell cycle arrest at the S phase (Figure 6-9C).


A
Vehicle
20 µM 5-FU
2 µM LDN-212854
20 µM 5-FU and 2 µM LDN
Ki67
DAPI
488 - 780_60-A :: Ki-67
405-450_50-A :: DAPI
Ap 2.67
G1 37.4
M 0.46
G2 40.1
S 8.79
G0 0.39
Ap 1.35
G1 22.8
M 0.58
G2 38.2
S 17.7
G0 0.028
Ap 1.53
G1 35.4
M 0.50
G2 43.1
S 5.98
G0 0.33
Ap 2.47
G1 29.6
M 0.54
G2 35.0
S 16.9
G0 0.068

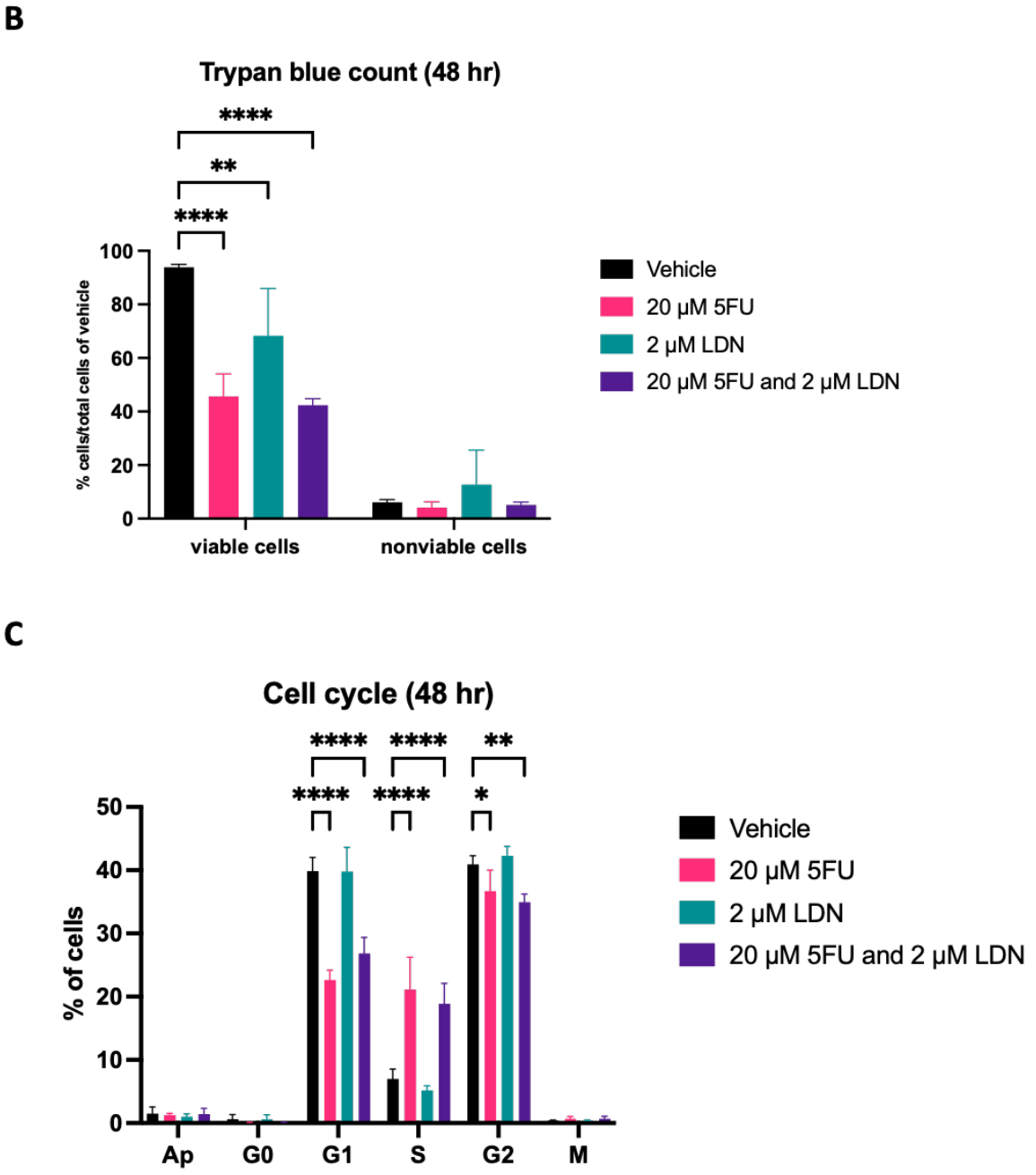


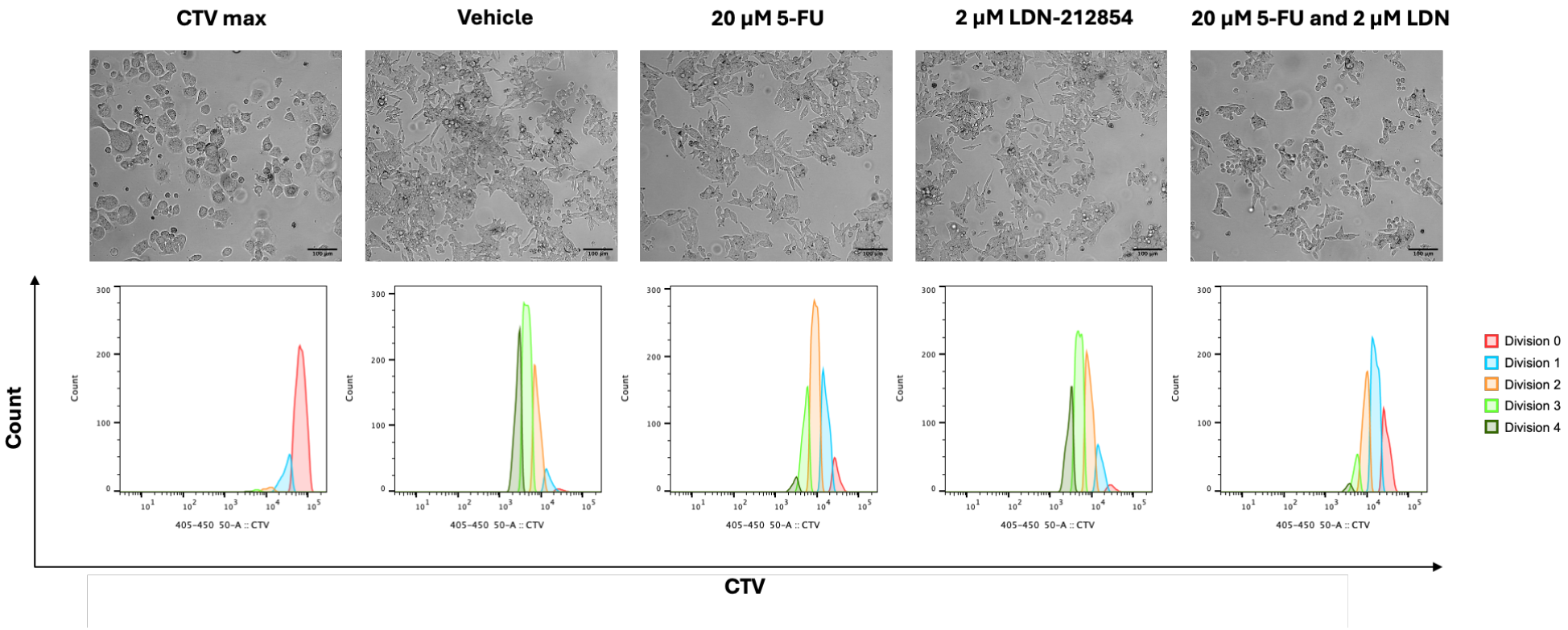
Figure 6-9 Inhibition effects of the drug combination between LDN-212854 and 5-FU on cell cycle analysis of the SW620 cell line

Images illustrated the morphology (above) and the corresponding FlowJo™ density plots (below) of SW620 cells treated with vehicle, 20 μM 5-FU, 2 μM LDN-212854, and a combination of 20 μM 5-FU and 2 μM LDN-212854 after 48 hours (A). The histogram from the trypan blue exclusion assay showed the difference in the percentage of viable cells between treatment conditions and the vehicle. However, no difference in non-viable cells was observed (B). The histogram from cell cycle analysis revealed variation in the percentage of the cell population in each stage of the cell cycle across treatment conditions (C). Dunnett's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, mean ± SD). The scale bar = 100 μm in Figure A.

6.4.5 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell proliferation using CTV analysis.

Subsequently, the inhibitory effects on cell proliferation were assessed by tracking cell divisions using a CTV assay as described in Section 2.2.4.3. Briefly, SW620 cells were stained with 5 μM of CTV before being treated with 20 μM 5-FU, 2 μM LDN-212854, and the drug combination in SW620 cells for 96 hours. Representative images of cell morphology (above) and the corresponding FlowJo™ proliferation analysis histogram plots (below) under treatment conditions demonstrated changes in cell morphology and cell divisions after 96-hour treatment (Figure 6-10A). The density plots of CTV fluorescence intensity from each treatment condition were compared, revealing overall differences in cell distribution (Figure 6-10B). When dividing the graphs into each cell division, the results showed that viable cells from the vehicle condition entered cell divisions 3 and 4 within 96 hours. After being treated with 2 μM LDN-212854, live cells were observed at several division stages, with the majority of cells found in division 3. The viable cells treated with 20 μM 5-FU were mostly in cell division 2, while live cells from the combination treatment remained in cell division 1 and appeared to be delayed in cell proliferation (Figure 6-10C).

A


CTV max
Vehicle
20 µM 5-FU
2 µM LDN-212854
20 µM 5-FU and 2 µM LDN
Count
CTV
405-450 50-A :: CTV
Division 0
Division 1
Division 2
Division 3
Division 4

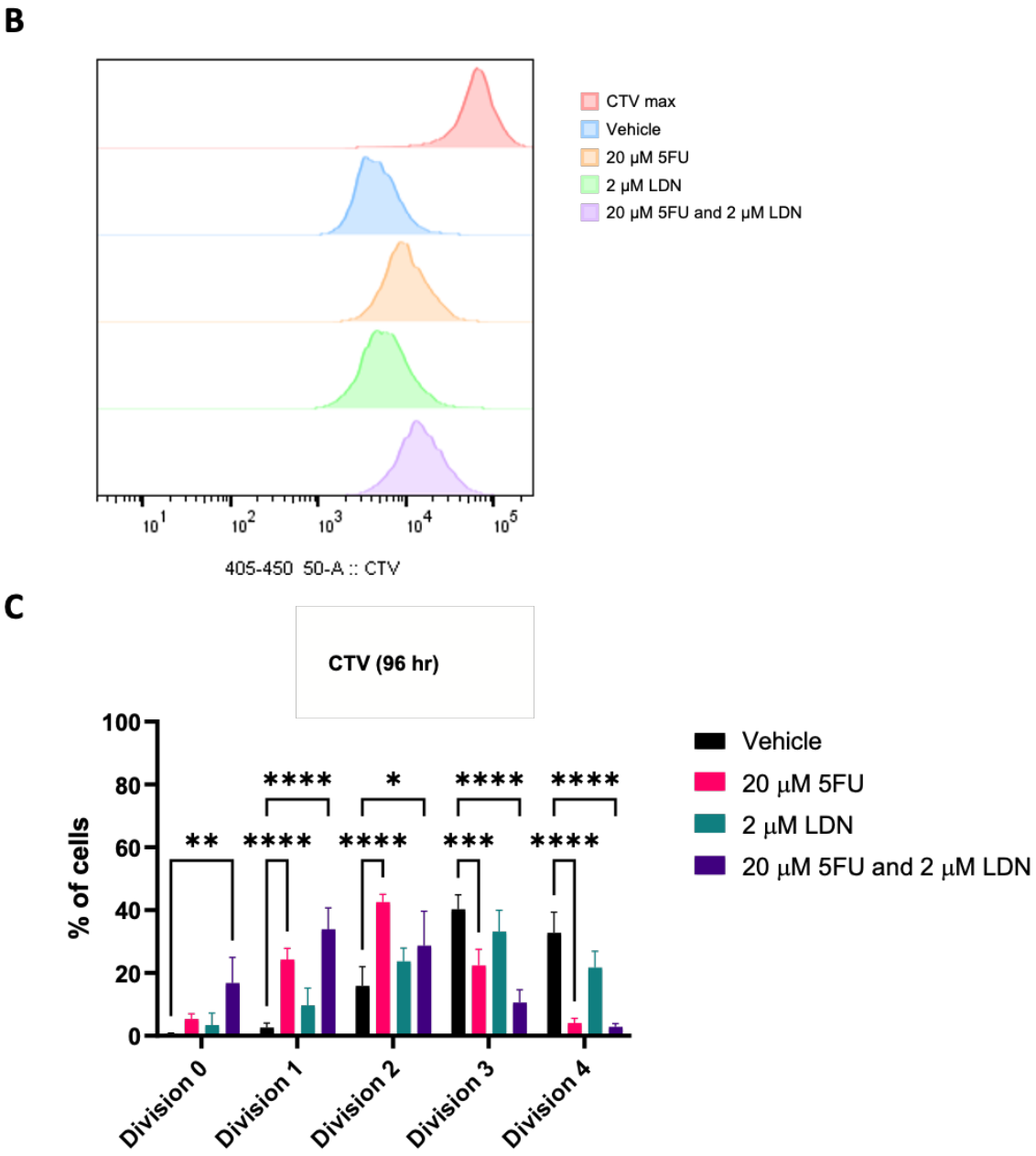


Figure 6-10 Inhibition effects of the drug combination between LDN-212854 and 5-FU on CTV analysis of the SW620 cell line

Images illustrated the morphology (above) and the corresponding FlowJo™ proliferation analysis histogram plots (below) of SW620 cells treated with Colcemid (CTV max), vehicle, 20 µM 5-FU, 2 µM LDN-212854, and a combination of 20 µM 5-FU and 2 µM LDN-212854 after 96 hours (A). The density plots of CTV fluorescence intensity showed cell distribution across treatment conditions (B). The histogram from CTV analysis revealed variation in the percentage of the cell population in each cell division across treatment conditions (C). Dunnett's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, mean ± SD). The scale bar = 100 µm in Figure A.

6.4.6 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell recovery using clonogenic assay

Before conducting the clonogenic assay to evaluate the recovery capacity of SW620 cells after drug treatment, we optimised the seeding density and culture conditions for low-density growth and continued the culture for up to two weeks, as described in Section 2.2.3.6. To determine the optimal conditions, SW620 cells were seeded at densities of 500 up to 1,000 cells per well and compared when cultured with 0.1% DMSO as a vehicle and media only. Four seeding conditions were validated, and representative colony-forming images were obtained, including; standard culture with normal medium (Figure 6-11A), culture with 10% SW620 conditioned medium (Figure 6-11B), pre-coating the culture plate with 1% BME and culturing with normal medium (Figure 6-11C), and pre-coating the culture plate with 1% BME and culturing with SW620 10% conditioned medium (Figure 6-11D). The results from the optimization experiment indicated that higher cell density produced a greater number of colonies. However, there was no difference in colony density between the normal and vehicle control conditions, nor between normal medium and 10% conditioned medium. When comparing standard culture with pre-coated plates, differences in colony morphology were observed. Small, uncountable colonies appeared in standard culture, whereas pre-coating the plate improved colony size, facilitating colony formation and clearer colony distribution at seeding densities between 500 and 750 cells per well.

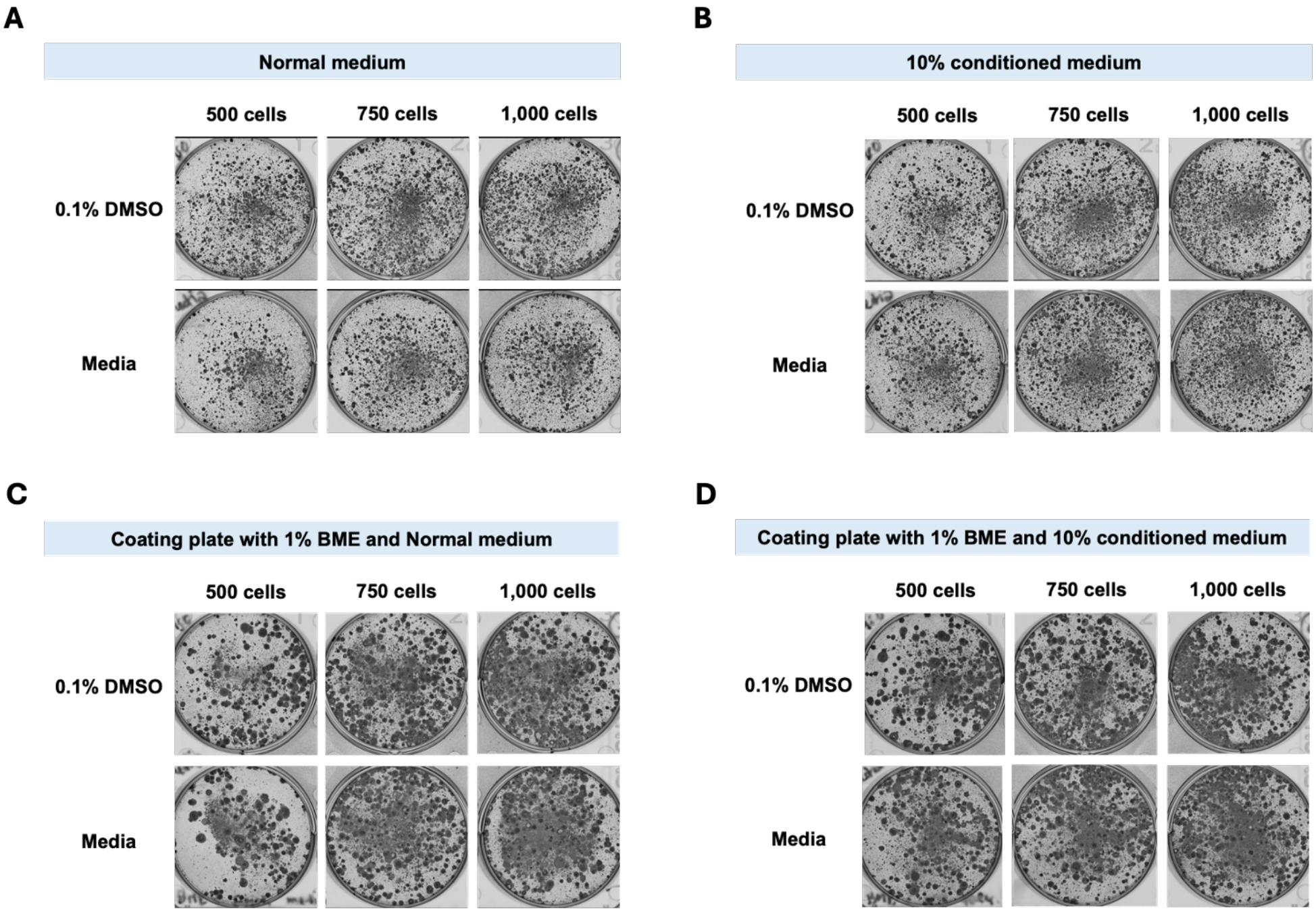

A
Normal medium
500 cells
750 cells
1,000 cells
0.1% DMSO
Media
B
10% conditioned medium
500 cells
750 cells
1,000 cells
0.1% DMSO
Media
C
Coating plate with 1% BME and Normal medium
500 cells
750 cells
1,000 cells
0.1% DMSO
Media
D
Coating plate with 1% BME and 10% conditioned medium
500 cells
750 cells
1,000 cells
0.1% DMSO
Media

Figure 6-11 Optimisation of seeding density and culture conditions for SW620 cell line in Clonogenic assay

Images illustrated colony formation in 0.1% DMSO (treated for 96 hours as a vehicle) and media only after seeding at densities of 500, 750 and 1,000 cells per well, and continued culturing them for 2 weeks. This compared the culture conditions among standard culture with normal medium (A), culture with 10% SW620 conditioned medium (B), pre-coating the culture plate with 1% BME and culturing with normal medium (C), and pre-coating the culture plate with 1% BME and culturing with 10% SW620 conditioned medium (D).

Subsequently, we seeded SW620 cells at 600 cells per well on a culture plate pre-coated with 1% BME and cultured them in normal medium. A low cell density of SW620 was treated both individually and in combination with 5-FU at 5 μM, 10 μM, and 20 μM, as well as LDN-212854 at 0.5 μM, 1 μM, and 2 μM, for 96 hours before continuing culture in normal medium for 2 weeks. Representative images of colony formation from each treatment condition showed fewer colonies at higher drug concentrations (Figure 6-12A). The colony-forming results revealed fewer than 10 colonies were detected in the treatment conditions with 1 to 2 μM LDN-212854 alone and in combination with 5-FU (Figure 6-12B). When comparing 0.5 μM LDN-212854 treatment, 5-20 μM 5-FU, and the drug combination to the vehicle, the percentage of relative colony number indicated that there was a reduction in colony formation with the drug treatments, with notably fewer colonies in the combination treatment than in the single treatment groups (Figure 6-12C).

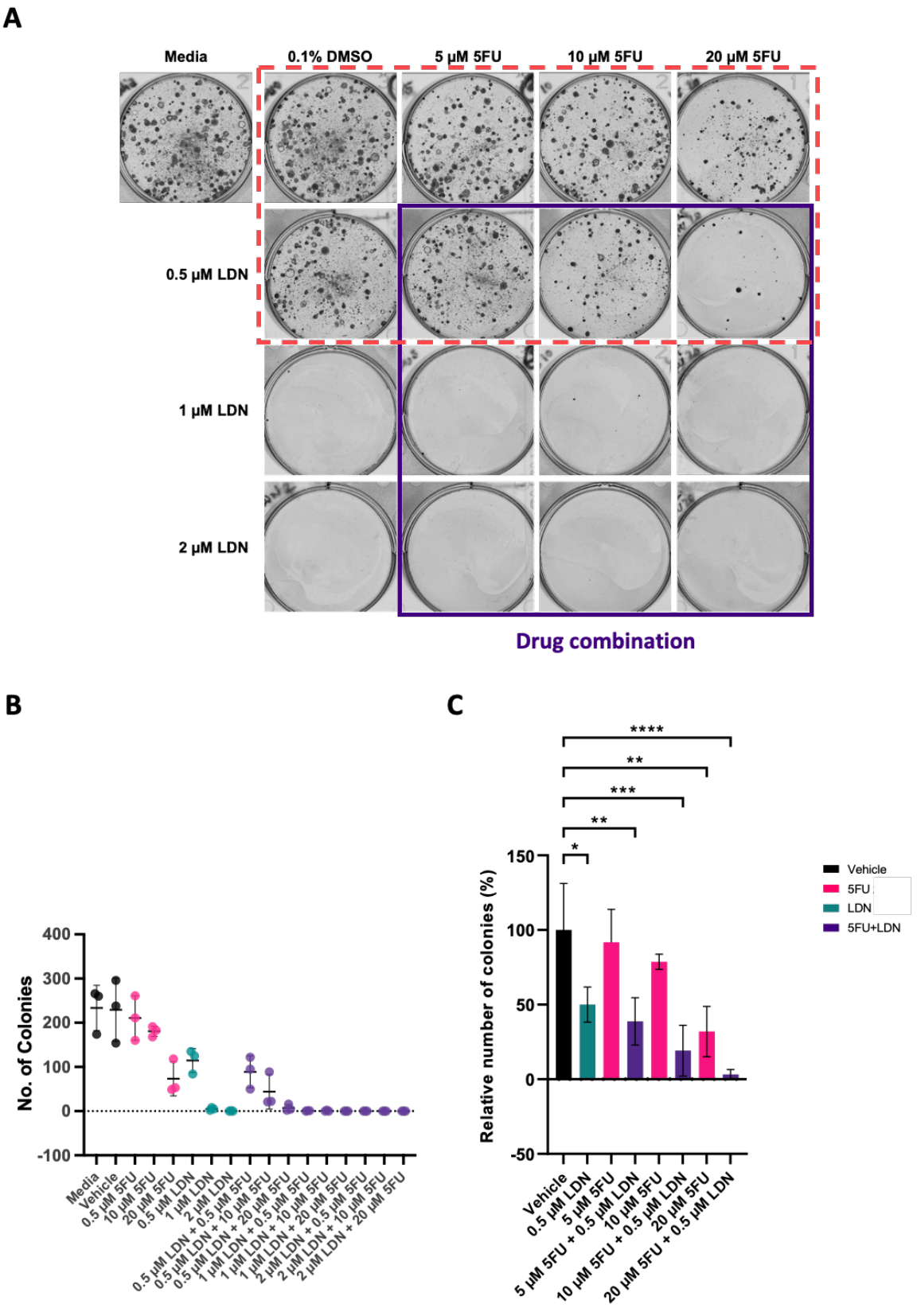


Figure 6-12 Inhibition effects of the drug combination between LDN-212854 and 5-FU on cell recovery of the SW620 cell line using Clonogenic assay

Images illustrated colony formation after treatment with a low density of cells using 5-FU and LDN-212854 for 96 hours. The cells were then recovered and cultured for two weeks (A). The histogram

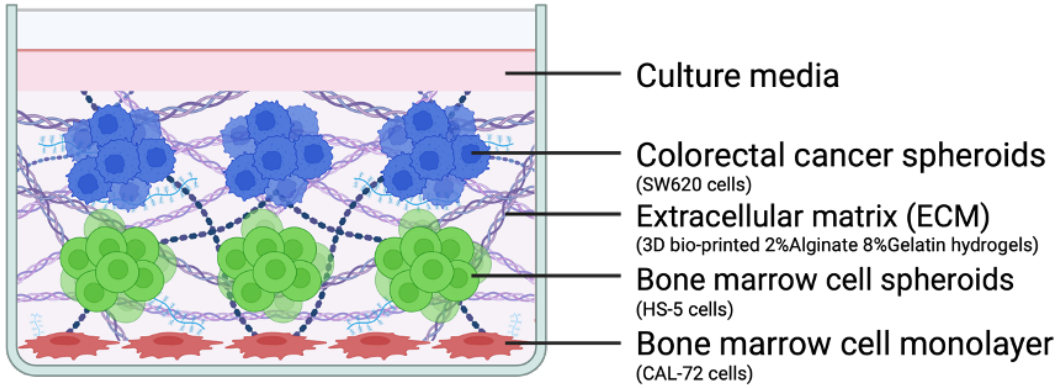
showed variation in the number of colonies under each treatment condition (B). The histogram comparing the relative number of colonies revealed a decrease in the percentage of colonies in drug treatment conditions (C). Dunnett's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, mean ± SD).

6.5 Developing an *in vitro* 3D culture model of CRC metastasis

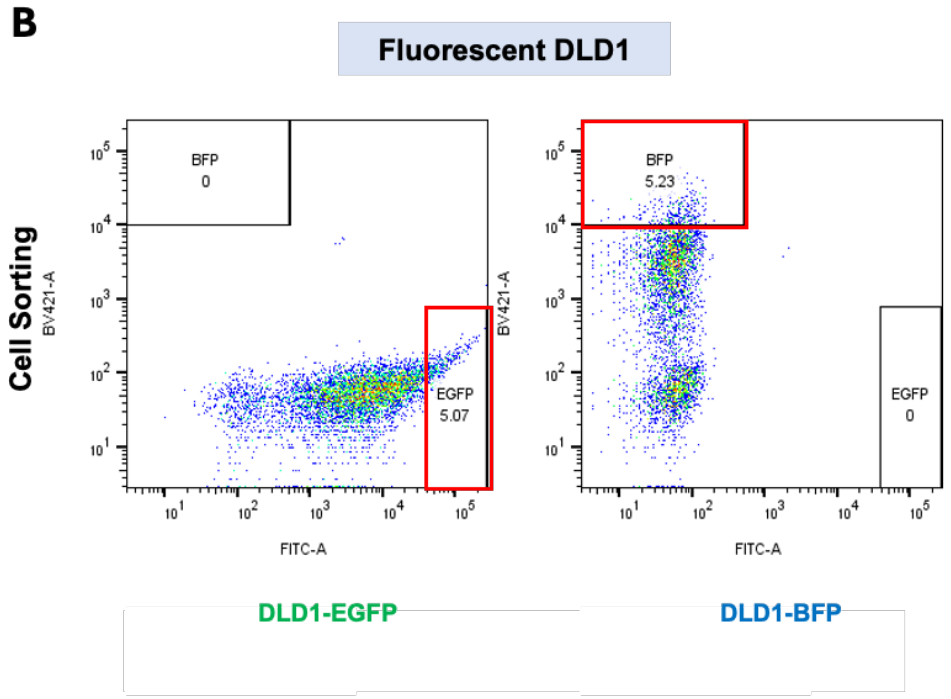
6.5.1 Establishing fluorescent CRC cell lines for 3D culture models

Furthermore, building on our previous findings regarding signalling mechanism alterations in the TGF-ß and BMP pathways and the effects of the drug combination from an *in vitro* 2D model of CRC, we aimed to develop an *in vitro* 3D culture of CRC to establish a model for studying CRC metastasis and testing a novel drug treatment that could prevent and/or treat mCRC. We proposed creating a co-culture model of CRC cells and other cell types, such as a multicell 3D model of metastatic CRC with BM cells. This model will include a monolayer of fluorescent osteocytes, spheroids of stromal cells embedded in an extracellular matrix (ECM), and fluorescent CRC cells (Figure 6-13A). Initially, we generated fluorescent CRC cells from DLD1 and SW620 to enable tracking of the cells in multicellular modelling. The cells were transfected with lentiviruses expressing EGFP and BFP as described in Section 2.2.3.8. Then, the transfected DLD1 and SW620 cells were sorted based on high fluorescence intensity as described in Section 2.2.4.4. Representative images of fluorescent DLD1 and SW620 cells, both expressing EGFP and BFP, were obtained through cell sorting (Figure 6-13B and C, respectively) and observed under a fluorescence microscope (Figure 6-13D and E, respectively). The fluorescent DLD1 and SW620 cells were analysed based on their fluorescent profiles using flow cytometry (Figure 6-13F and G, respectively). Subsequently, the fluorescent cells were expanded and cryopreserved for further experiments.

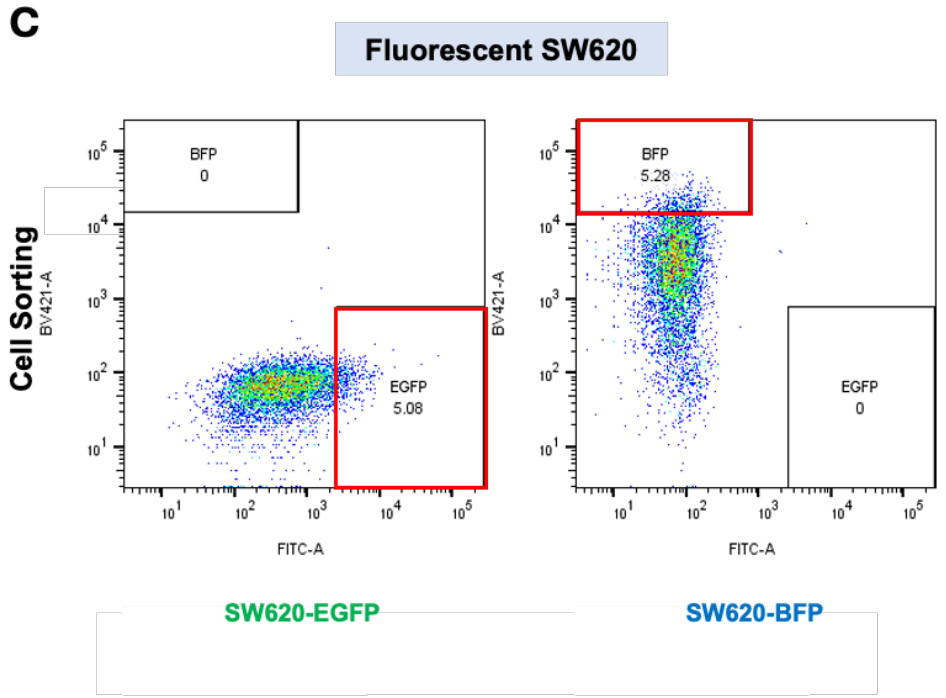
A


Co-culture Model
Culture media
Colorectal cancer spheroids
(SW620 cells)
Extracellular matrix (ECM)
(3D bio-printed 2%Alginate 8%Gelatin hydrogels)
Bone marrow cell spheroids
(HS-5 cells)
Bone marrow cell monolayer
(CAL-72 cells)

B


Fluorescent DLD1
Cell Sorting
BV421-A
FITC-A
BFP
0
EGFP
5.07
BFP
5.23
EGFP
0
DLD1-EGFP
DLD1-BFP

C


Fluorescent SW620
Cell Sorting
BV421-A
FITC-A
BFP
0
EGFP
5.08
BFP
5.28
EGFP
0
SW620-EGFP
SW620-BFP

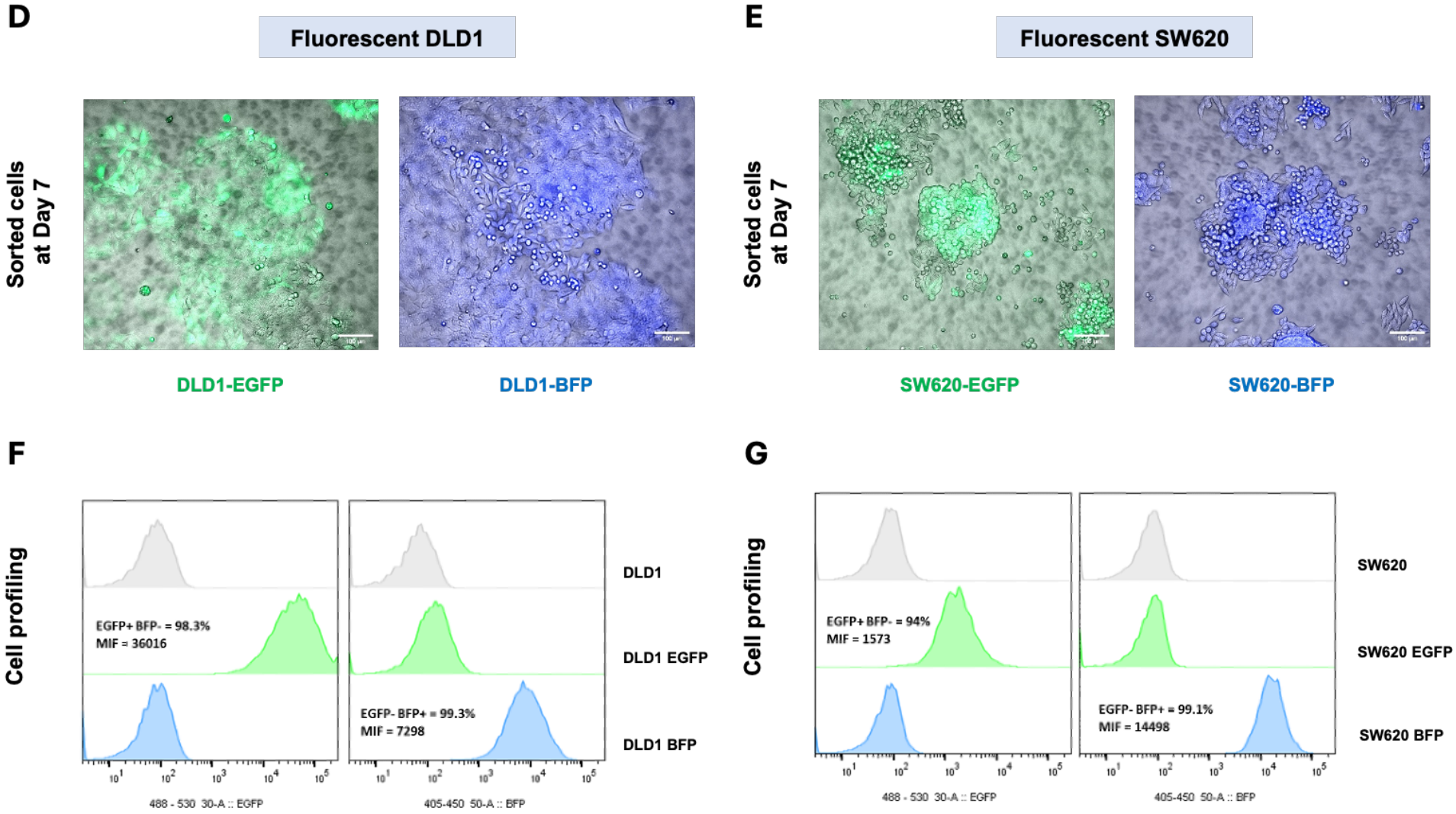

D
Fluorescent DLD1
Sorted cells at Day 7
DLD1-EGFP
DLD1-BFP
E
Fluorescent SW620
Sorted cells at Day 7
SW620-EGFP
SW620-BFP
F
Cell profiling
EGFP+ BFP- = 98.3%
MIF = 36016
EGFP- BFP+ = 99.3%
MIF = 7298
DLD1
DLD1 EGFP
DLD1 BFP
488 - 530 30-A :: EGFP
405-450 50-A :: BFP
G
Cell profiling
EGFP+ BFP- = 94%
MIF = 1573
EGFP- BFP+ = 99.1%
MIF = 14498
SW620
SW620 EGFP
SW620 BFP
488 - 530 30-A :: EGFP
405-450 50-A :: BFP

Figure 6-13 Establishing fluorescent CRC cell lines (SW620 and DLD1) for multicell 3D modelling

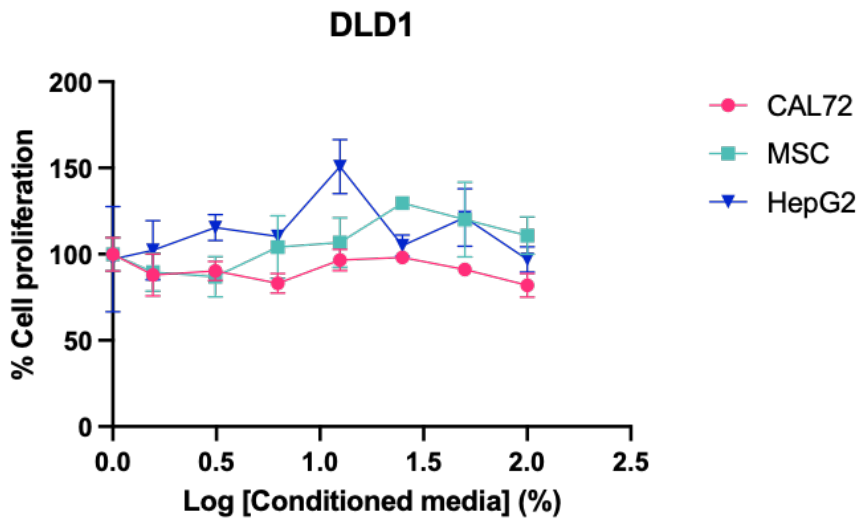
Schematic diagram of multicell 3D modelling for co-culture CRC with BM cells (A). After lentiviral transduction, fluorescent DLD1 cells (B) and SW620 cells (C) were sorted for high expression of EGFP and BFP. Images showed the fluorescent DLD1 cells (D) and SW620 cells (E) expressing EGFP and BFP after sorting and expansion for 7 days. The density plots of EGFP and BFP fluorescence intensity displayed the fluorescence cell profiles of DLD1 cells (F) and SW620 cells (G). The scale bar = 100 μm in Figures D and E.

6.5.2 Cell viability of CRC cell lines in the conditioned media of BM and HCC cells

Before combining different cell types into a co-culture model, we cultured CRC cells in conditioned media from other cells to evaluate their effects on proliferation, as described in Section 2.2.3.9. The results from culturing DLD1 and SW620 cells over 72 hours in conditioned media from CAL72, MSC, and HepG2— with concentrations ranging from 0% to 100%— indicated a decrease in cell growth at approximately 25% conditioned media in most cases. This may be due to a reduction in nutrients in the culture media (Figure 6-14Aand B, respectively). Therefore, we decided to culture both CRC cell lines in 0%, 10%, 20%, and 30% conditioned media from the CAL72, MSC and HepG2 cells, including condition media from DLD1 and SW620 as a control, to assess the effects more precisely. The results showed that 20% conditioned media could alter the growth patterns. A slight increase in the proliferation of DLD1 cells was observed when cultured in conditioned media from MSC and HepG2, whereas a decrease was seen with media from CAL72 and DLD1 (Figure 6-14C). An increase in proliferation was observed in SW620 cells cultured in conditioned media from MSC and SW620, whereas a decrease was noted with media from CAL72 and HepG2 (Figure 6-14D). Representative images of DLD1 and SW620 cells cultured in 20% of each conditioned medium displayed their cell morphology, which was similar to that in normal media (Figure 6-14E and F, respectively).

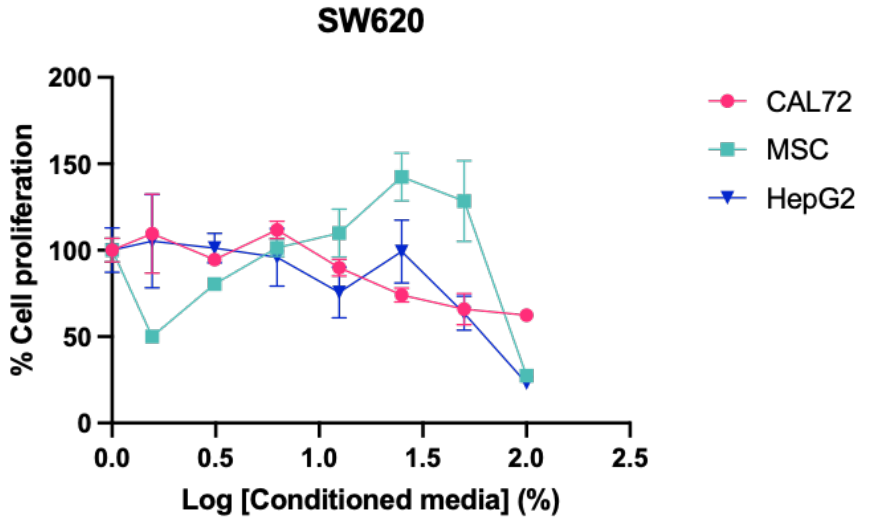
A

DLD1 with Conditioned media at 72 hr

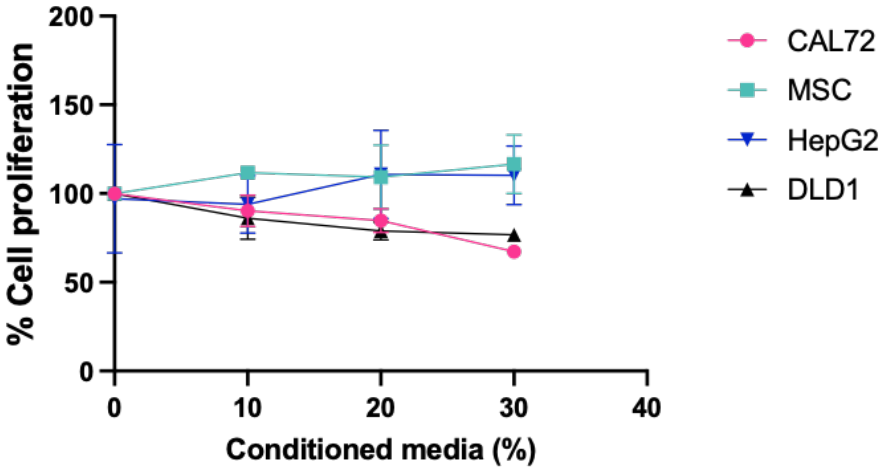


B

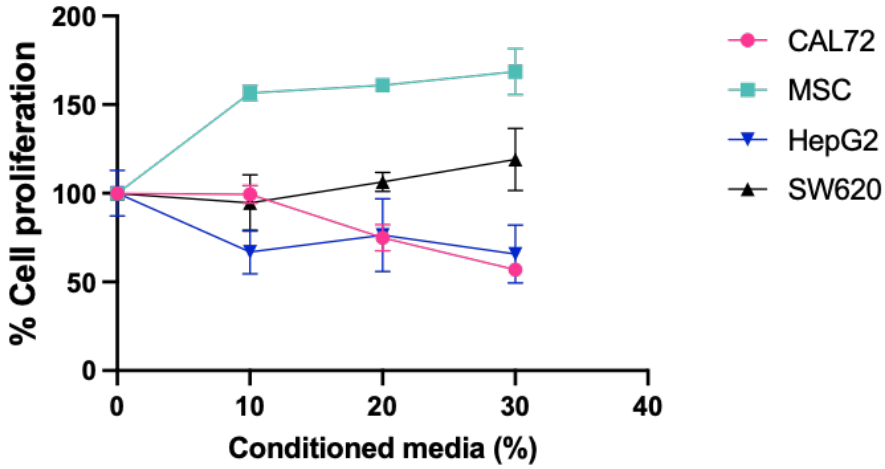
SW620 with Conditioned media at 72 hr



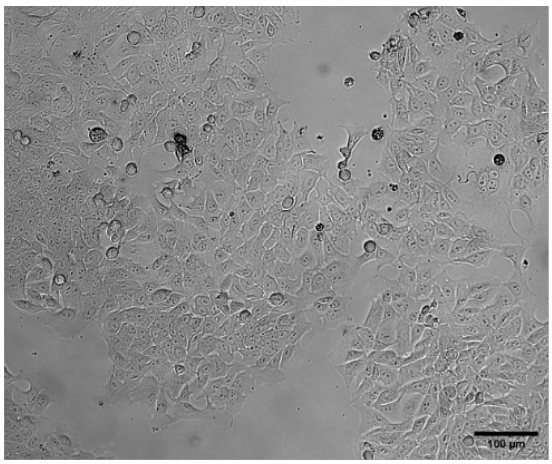
C



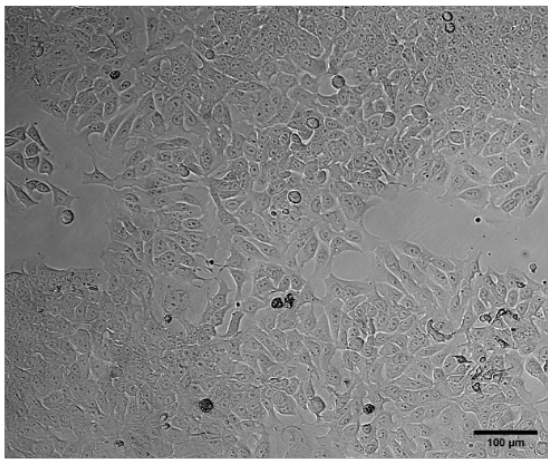
D



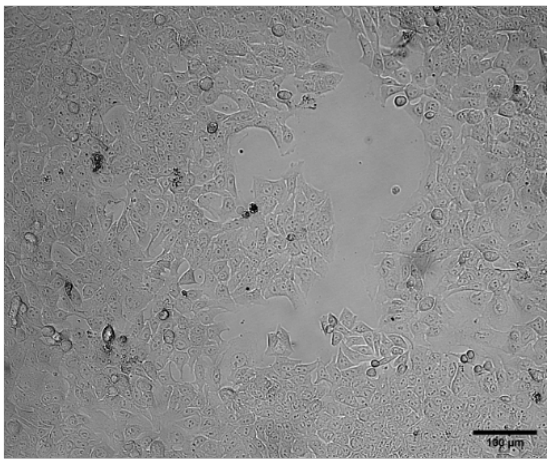
E



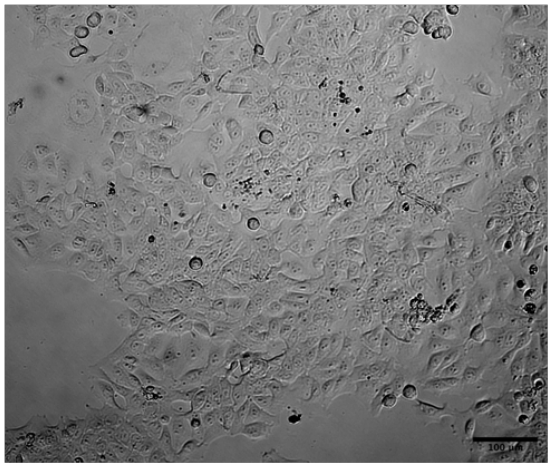
Normal medium of DLD1



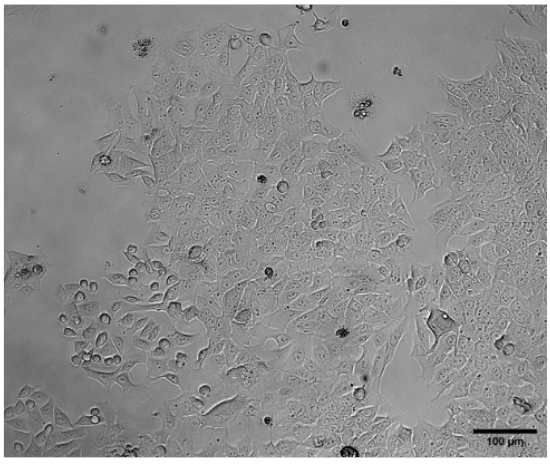
20% Conditioned medium of CAL72



20% Conditioned medium of MSC

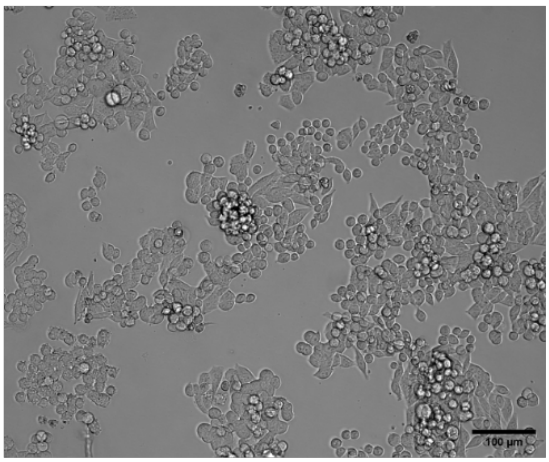


20% Conditioned medium of HepG2

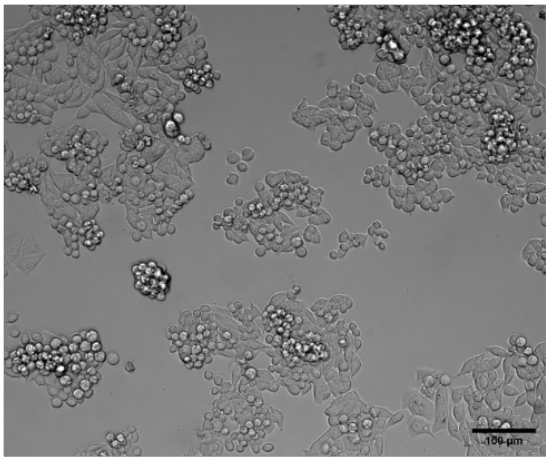


20% Conditioned medium of DLD1

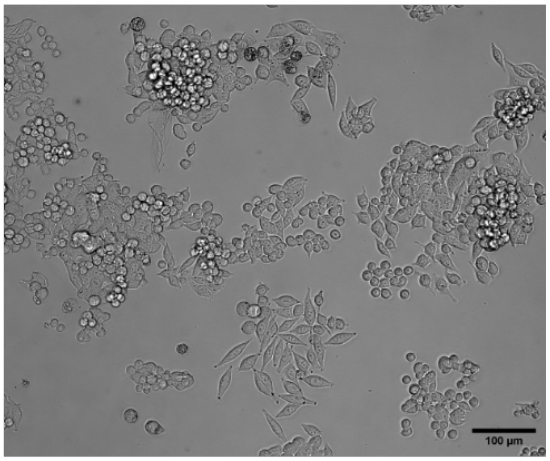
F



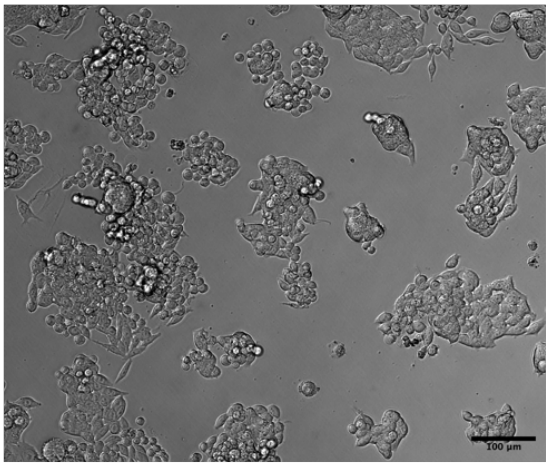
Normal medium of SW620



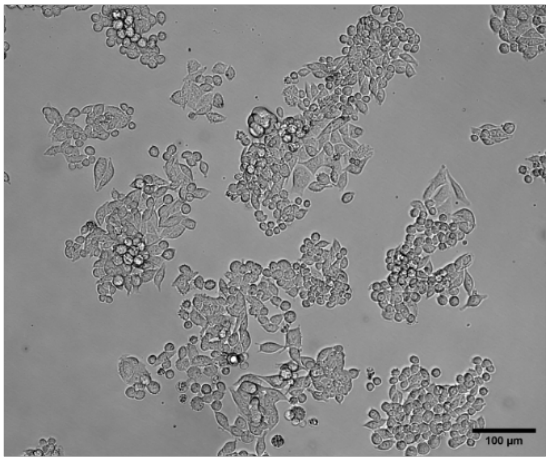
20% Conditioned medium of CAL72



20% Conditioned medium of MSC



20% Conditioned medium of HepG2



20% Conditioned medium of SW620

Figure 6-14 Effects on cell proliferation of culturing CRC cell lines in the conditioned media from BM and HCC cells

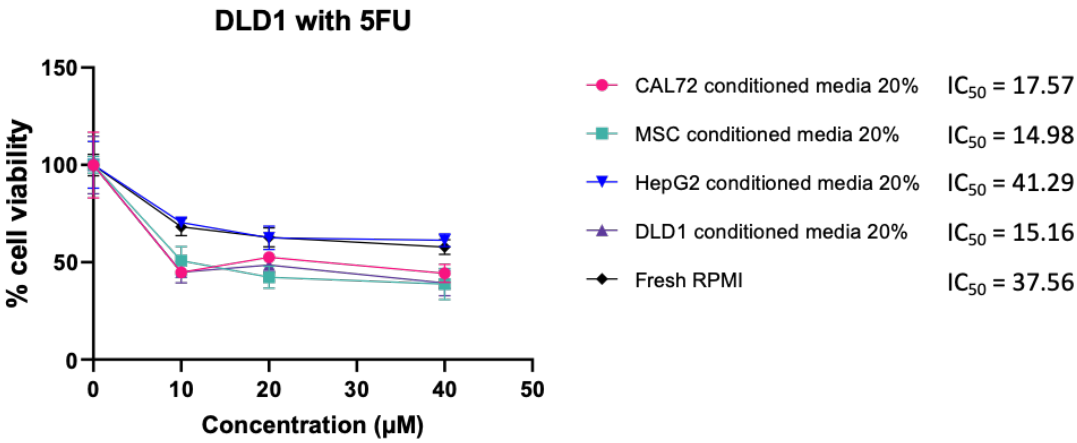
The graphs displayed the percentages of cell proliferation after culturing DLD1 cells (A) and SW620 cells (B) in a range from 0% to 100% conditioned media from CAL72, MSC, and HepG2 for 72 hours, assessed using the WST-1 assay. Additionally, the graphs showed the percentages of cell proliferation after culturing DLD1 cells (C) and SW620 cells (D) in 0%, 10%, 20%, and 30% conditioned media from CAL72, MSC, HepG2, DLD1, and SW620 over 72 hours. Images illustrated the cell morphology of DLD1 cells (E) and SW620 cells (F) cultured in standard medium and in 20% conditioned media conditions. The scale bar = 100 μm in Figures E and F.

6.5.3 Inhibition effects of drug treatment of LDN212854 and 5-FU on the proliferation of CRC cells cultured in conditioned media

Subsequently, we cultured DLD1 and SW620 cells in 20% conditioned media from CAL72, MSC, HepG2, DLD1, and SW620, and treated them for 72 hours with 10, 20, and 40 μM of 5-FU for both cell lines. Additionally, we used 1.5, 3, and 6 μM LDN-212854 for DLD1 cells, and 1, 2, and 4 μM LDN-212854 for SW620 cells, as described in Section 2.2.3.9. The results from the WST-1 assay showed that DLD1 cells responded better to 5-FU when cultured in the conditioned media of CAL72, MSC, and DLD1, with all IC50 values below 20. In contrast, culture with conditioned media from HepG2 and fresh medium was associated with higher cell proliferation rates and higher IC50 values (IC50 = 41.29 and 37.56, respectively) (Figure 6-15A). Similarly, an improved response to LDN-212854 was observed in DLD1 cells cultured in the conditioned media of CAL72, MSC, and DLD1, with all IC50 values below 3, compared to fresh medium (IC50 = 3.56). Conversely, a poorer response was seen with conditioned media from HepG2 (IC50 = 5.35) (Figure 6-15B). Representative images of DLD1 cells under various conditioned media were presented, showing their morphological alterations when treated with 20 μM of 5-FU, as well as 3 μM of LDN-212854, for 72 hours (Figure 6-15C).

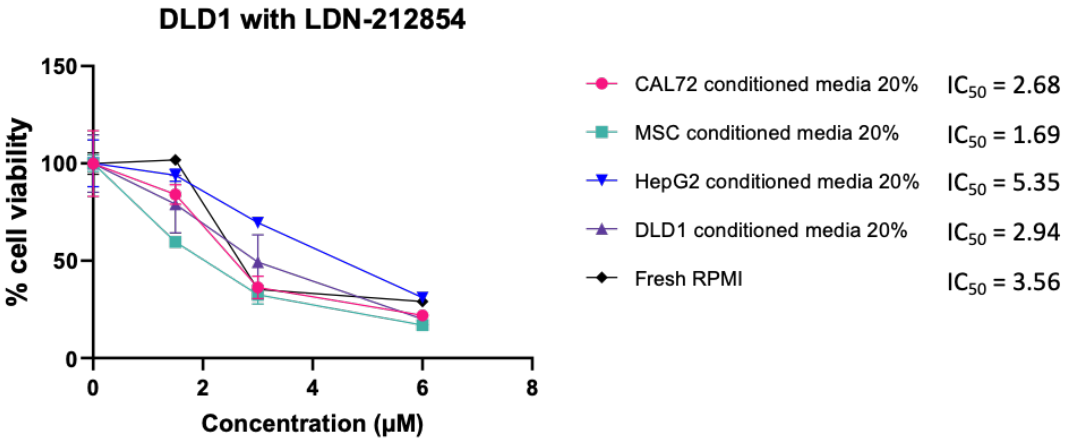
A

DLD1 with Conditioned media with 5-FU treatment at 72 hr



B

DLD1 with Conditioned media with LDN-212854 treatment at 72 hr



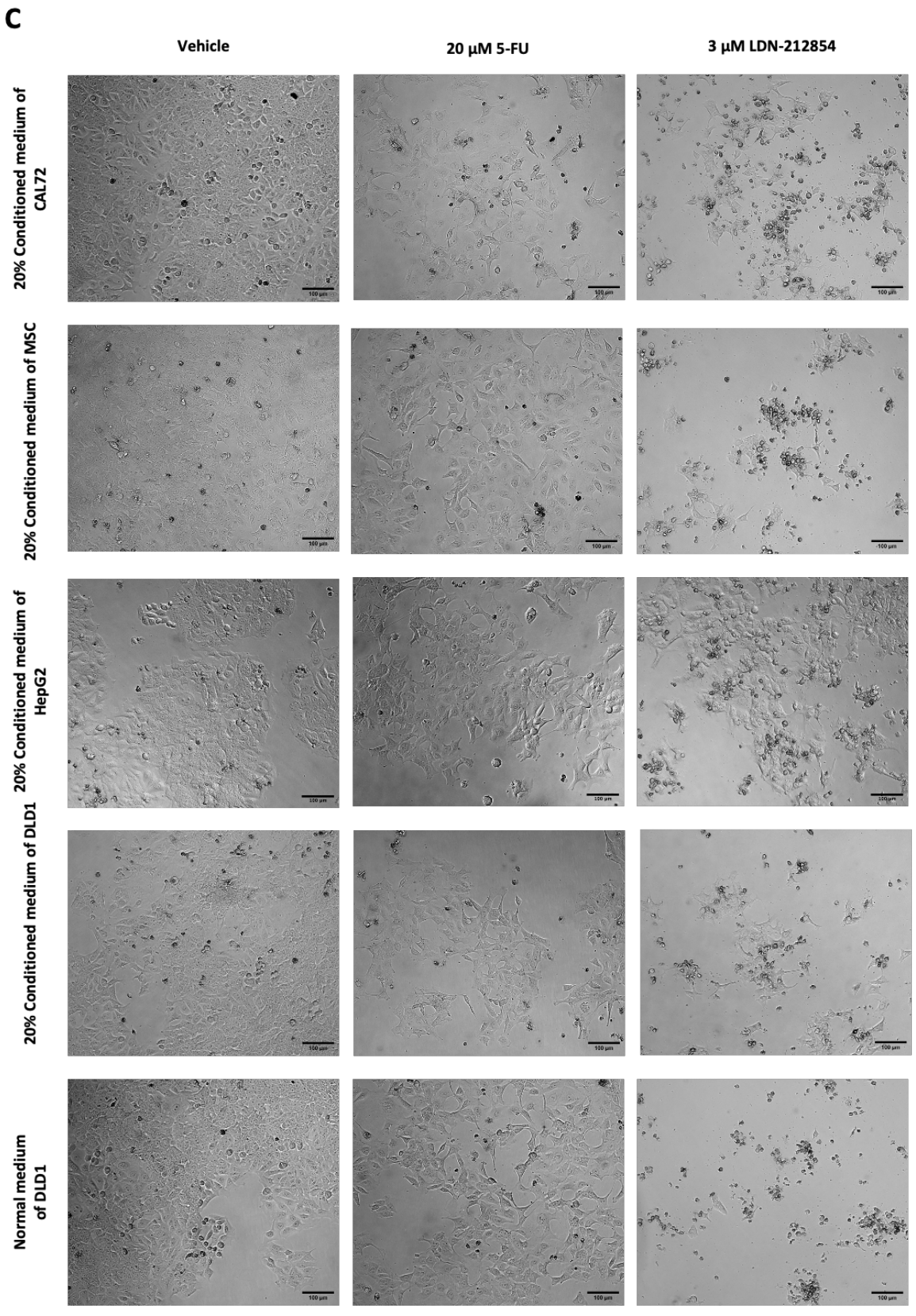

C
Vehicle
20 μM 5-FU
3 μM LDN-212854
20% Conditioned medium of CAL72
20% Conditioned medium of MSC
20% Conditioned medium of HepG2
20% Conditioned medium of DLD1
Normal medium of DLD1

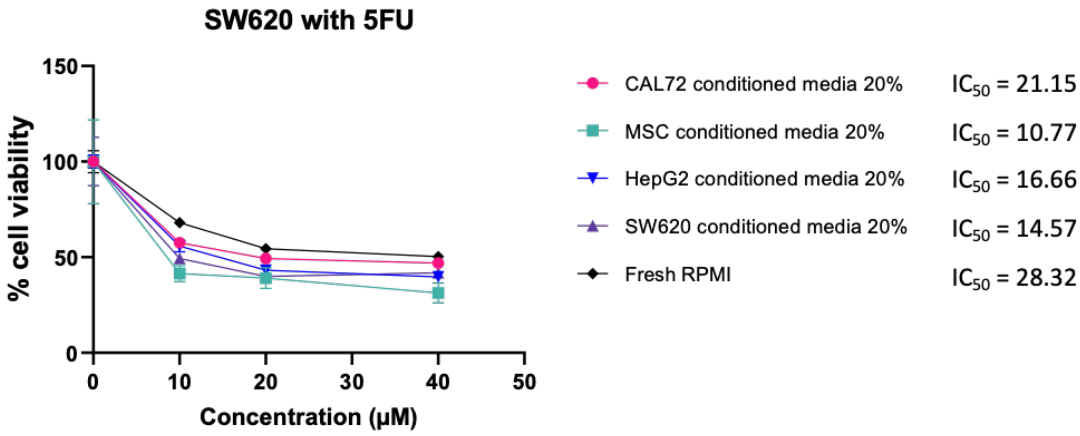
Figure 6-15 Inhibitory effects on cell viability of LDN-212854 and 5-FU treatment in DLD1 cells, which were cultured in 20% conditioned media from BM and HCC cells

The graphs showed that the inhibitory effects of 5-FU (A) and LDN-212854 (B) on cell viability were evaluated at 72 hours using the WST-1 assay, with several IC50 values determined. Images illustrated the morphology of DLD1 cells treated with vehicle, 20 μM 5-FU, and 3 μM LDN-212854 after 72 hours in 20% conditioned media from CAL72, MSC, HepG2, and DLD1, as well as in standard medium (C). The scale bar = 100 μm in Figure C.

When treating SW620 cells with 5-FU, the results showed that all conditioned media had greater inhibitory effects on cell proliferation compared to normal media (IC50 of CAL72 conditioned media = 21.15, IC50 of MSC conditioned media = 10.77, IC50 of HepG2 conditioned media = 16.66, IC50 of SW620 conditioned media = 14.57, and IC50 of normal media = 28.32) (Figure 6-16A). Furthermore, a stronger response to LDN-212854 was observed in cultures with MSC conditioned media (IC50 = 0.77), while IC50 values of other conditioned media were slightly higher than those of fresh media (IC50 of CAL72 conditioned media = 2.00, IC50 of HepG2 conditioned media = 2.34, IC50 of SW620 conditioned media = 2.16, and IC50 of normal media = 1.84) (Figure 6-16B). Representative images of SW620 cells under various conditioned media displayed morphological changes when treated with 20 μM of 5-FU and 2 μM of LDN-212854 for 72 hours (Figure 6-16C). However, these results were obtained from a single experiment; multiple replicate analyses would provide more substantial evidence.

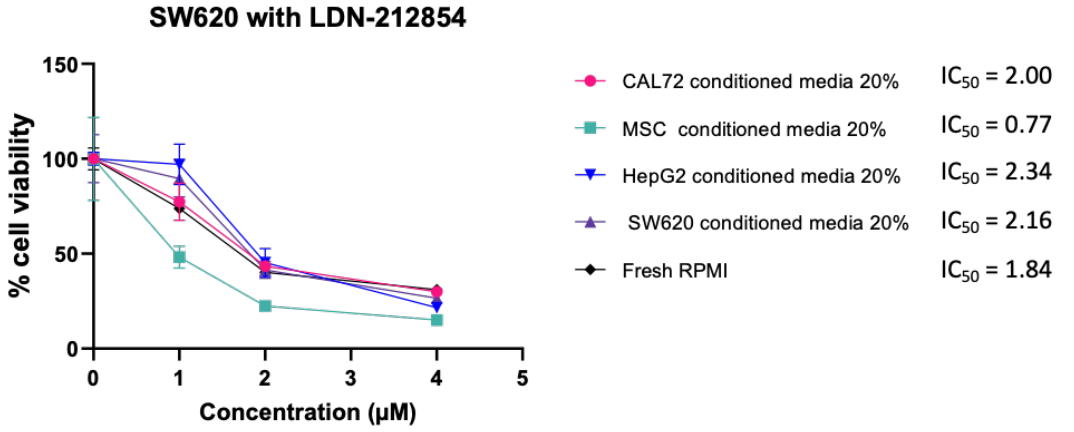
A

SW620 with Conditioned media with 5-FU treatment at 72 hr



B

SW620 with Conditioned media with LDN-212854 treatment at 72 hr



C

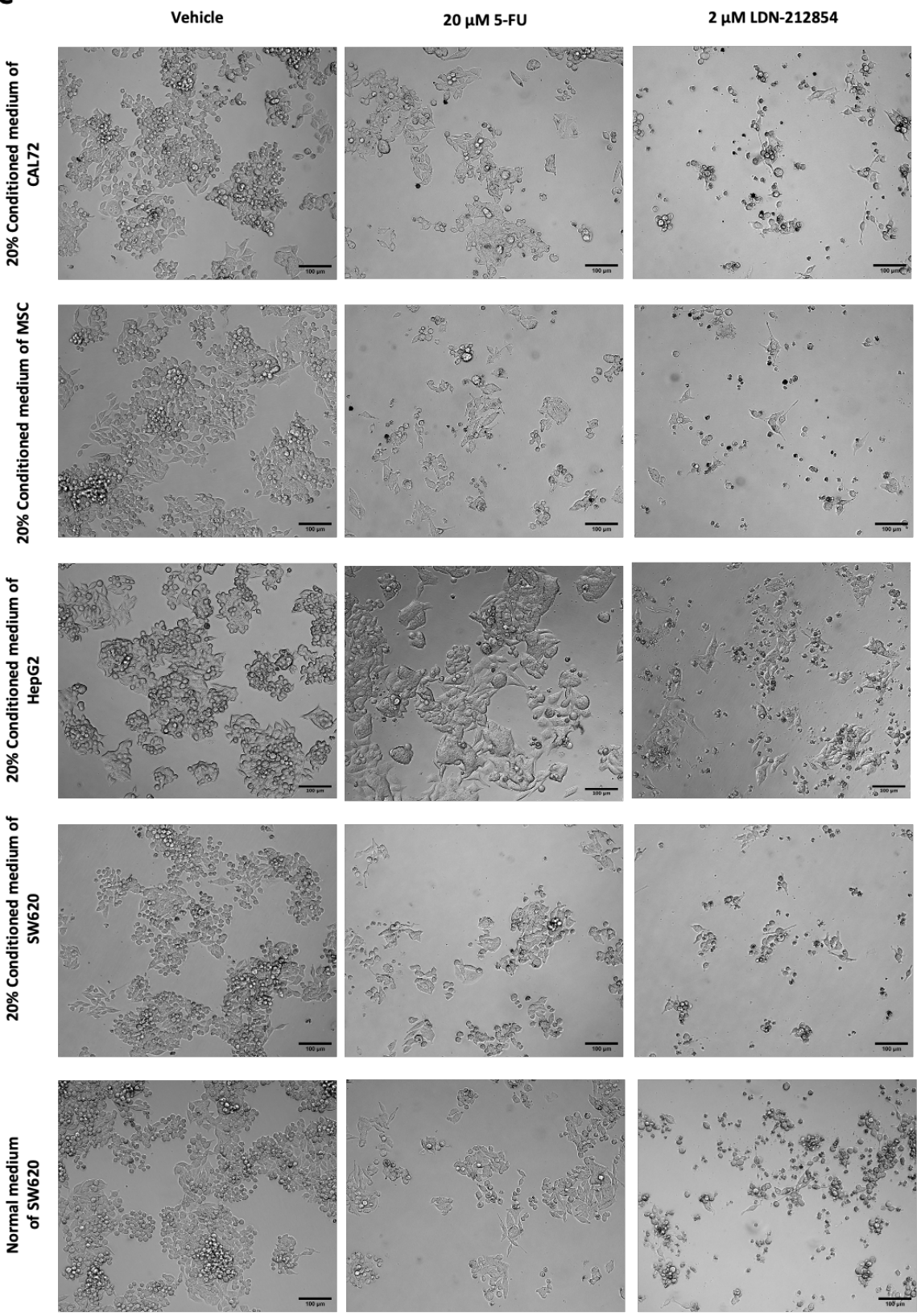

Vehicle
20 μM 5-FU
2 μM LDN-212854
20% Conditioned medium of CAL72
20% Conditioned medium of MSC
20% Conditioned medium of HepG2
20% Conditioned medium of SW620
Normal medium of SW620

Figure 6-16 Inhibitory effects on cell viability of LDN-212854 and 5-FU treatment in SW620 cells, which were cultured in 20% conditioned media from BM and HCC cells

The graphs showed that the inhibitory effects of 5-FU (A) and LDN-212854 (B) on cell viability were evaluated at 72 hours using the WST-1 assay, with several IC50 values determined. Images illustrated the morphology of SW620 cells treated with vehicle, 20 μM 5-FU, and 2 μM LDN-212854 after 72 hours in 20% conditioned media from CAL72, MSC, HepG2, and SW620, as well as in standard medium (C). The scale bar = 100 μm in Figure C.

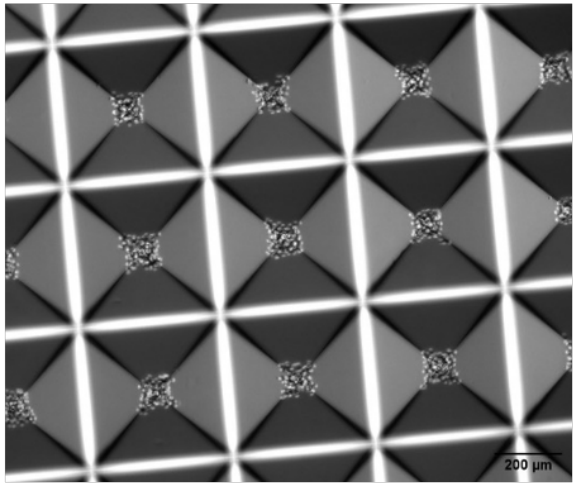
6.5.4 Formation of CRC spheroids on the Aggrewell™ plate

The formation of spheroids is one of the challenges in developing a 3D model. Therefore, we optimised the conditions for forming CRC spheroids. CRC cell lines, DLD1 and SW620, were seeded at densities of 50 and 100 cells per microwell, and spheroid development was observed over 7 days as described in Section 2.2.7.1. Representative images illustrating the generation of DLD1 spheroids using the Aggrewell™ plate over 7 days and stained with Live/Dead staining assay on Days 3, 5, and 7. Results show that DLD1 cells formed a 3D structure within 3 days at both seeding densities. Culturing DLD1 spheroids for 7 days maintains cell viability, as indicated by the live cell stain (Calcein AM) and the reduced intensity of dead cells (EthD-1), with no apparent size differences between conditions (Figure 6-17A and B). Meanwhile, SW620 began forming detectable spheroids on day 5, regardless of seeding density. Spheroids from both conditions can survive for at least 7 days, and those seeded with 100 cells were larger than those with 50 cells (Figure 6-18A and B).

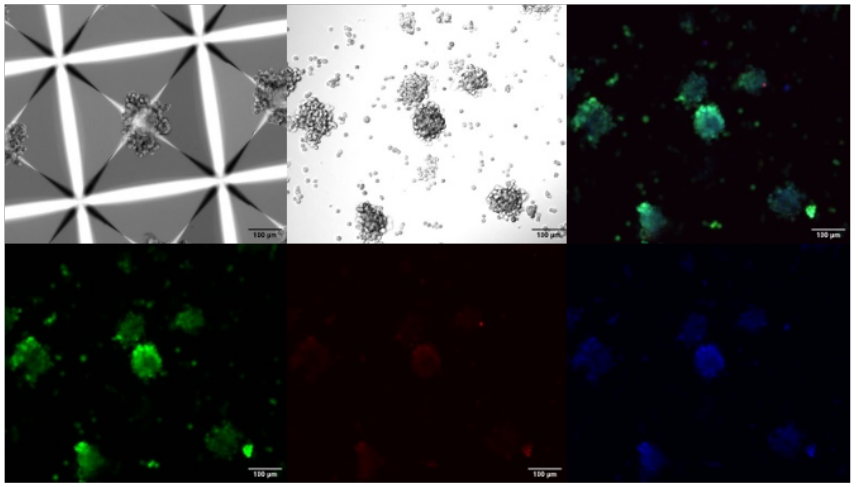
A

DLD1
50 cells/microwell

Day 0

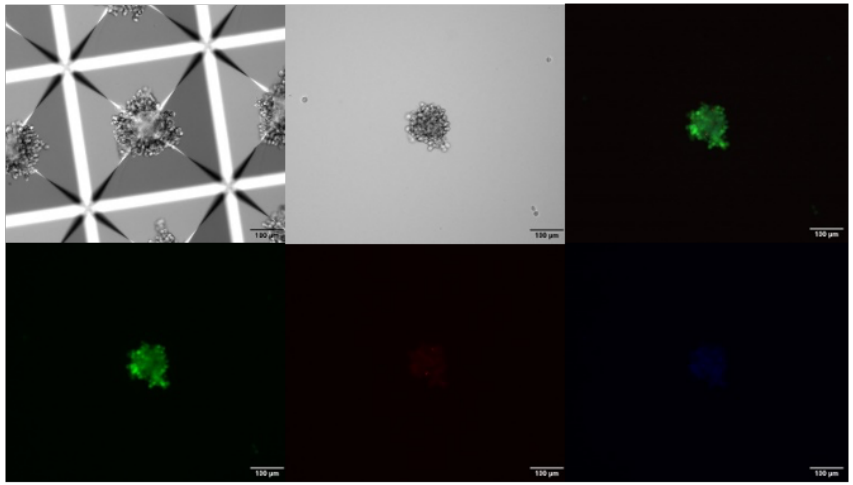


Day 3



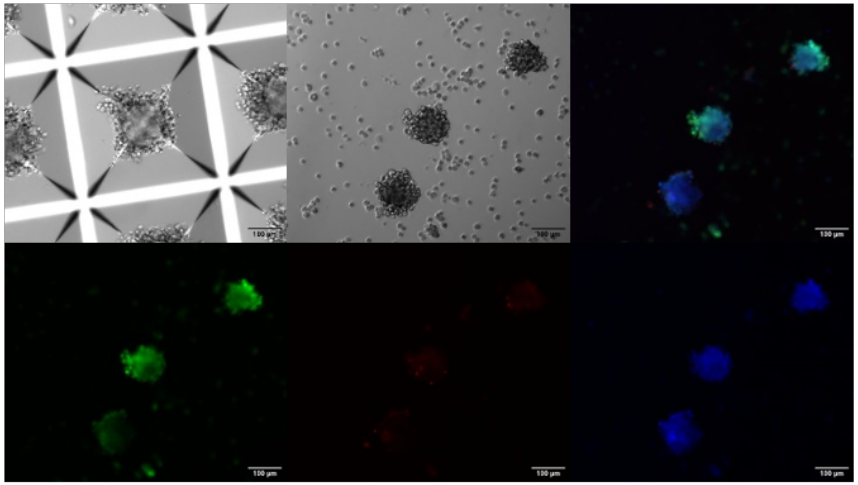
Calcein AM EthD-1 Hoechst 33342

Day 5



Calcein AM EthD-1 Hoechst 33342

Day 7

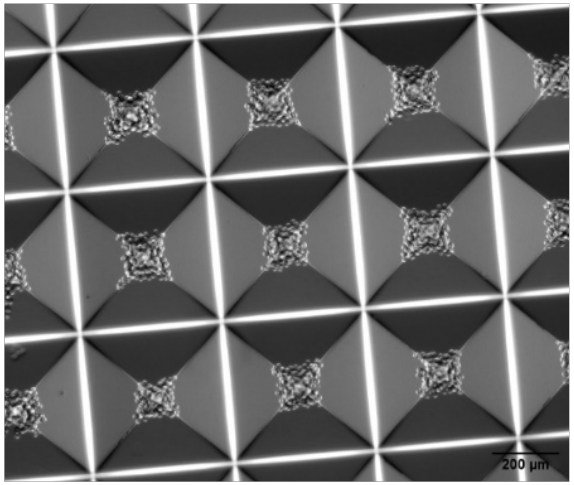


Calcein AM EthD-1 Hoechst 33342

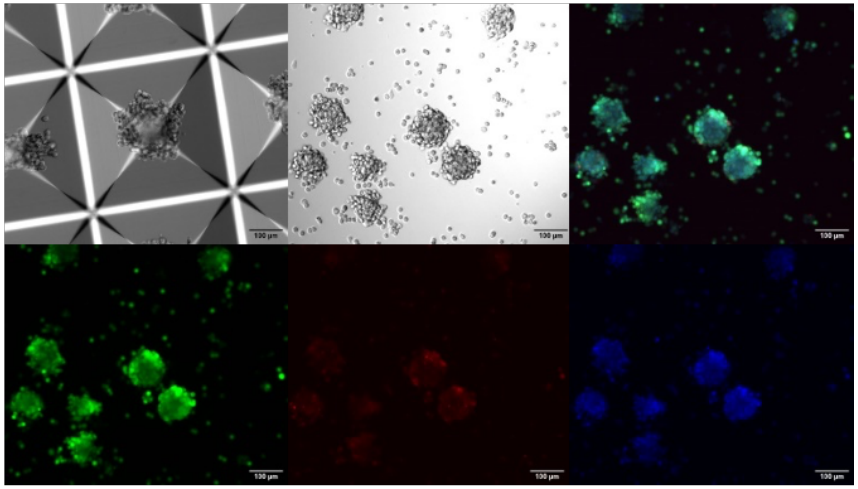
B

DLD1
100 cells/microwell

Day 0

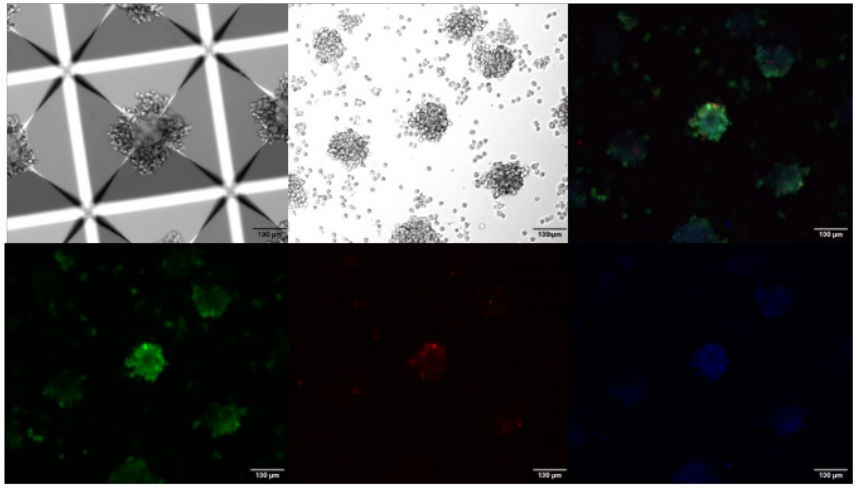


Day 3



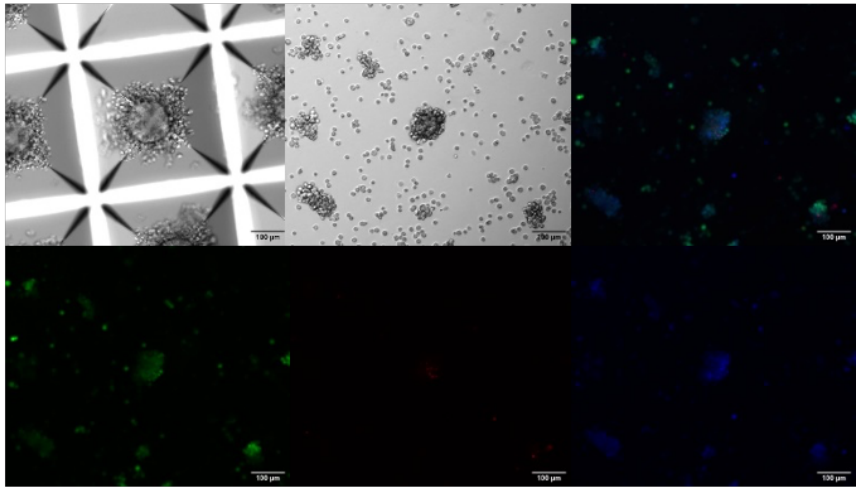
Calcein AM EthD-1 Hoechst 33342

Day 5



Calcein AM EthD-1 Hoechst 33342

Day 7



Calcein AM EthD-1 Hoechst 33342

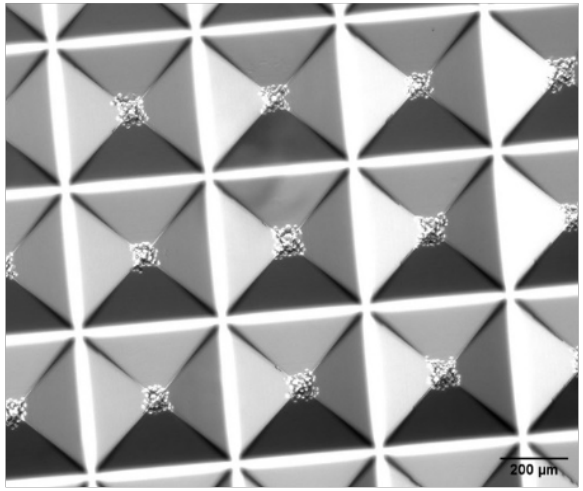
Figure 6-17 DLD1 spheroid formation during culture on an Aggrewell™ plate

Images showed DLD1 cells on an Aggrewell™ plate with seeding densities of 50 cells per microwell (A) and 100 cells per microwell (B) over days 0, 3, 5, and 7 of culture. Spheroid formation was observed on day 3. Calcein AM (green) indicated live cells, EthD-1 (red) indicated dead cells, and Hoechst 33342 (blue) stained nuclei. The scale bar = 200 μm in Figures for day 0. The scale bar = 100 μm in Figures for days 3, 5, and 7.

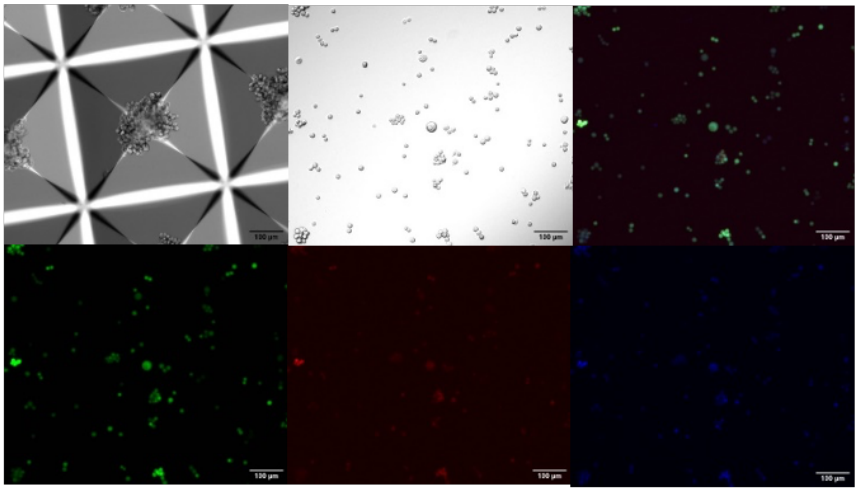
A

SW620
50 cells/microwell

Day 0

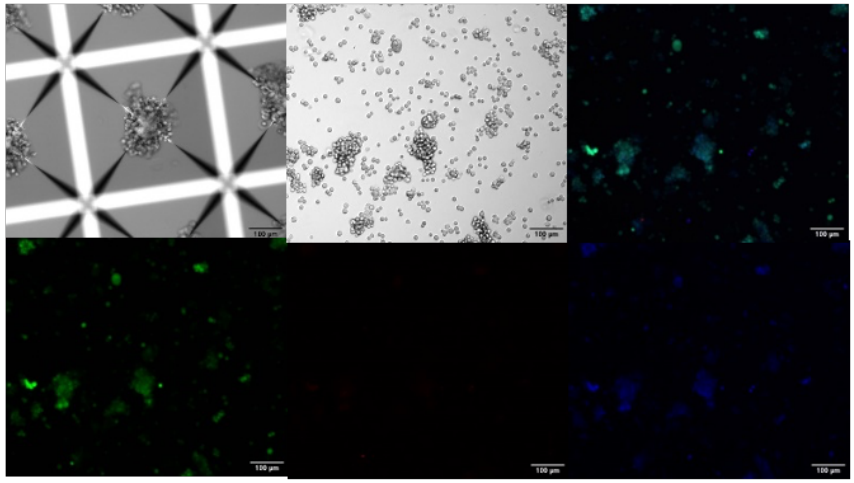


Day 3



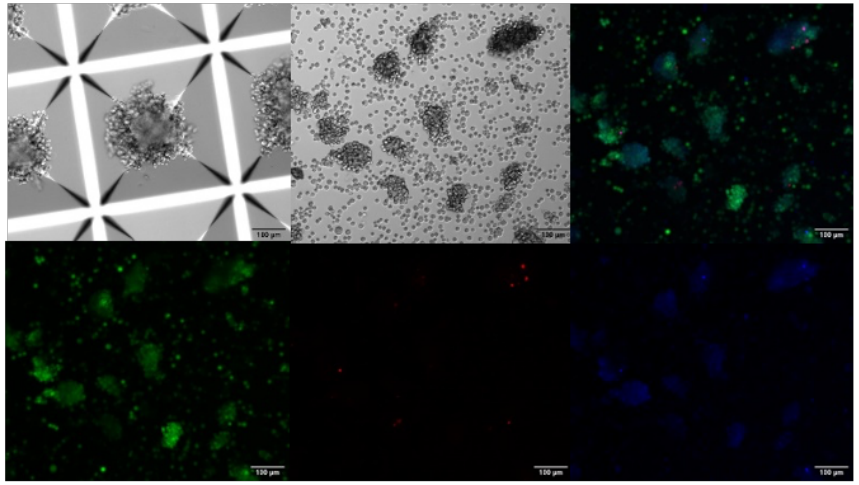
Calcein AM EthD-1 Hoechst 33342

Day 5



Calcein AM EthD-1 Hoechst 33342

Day 7

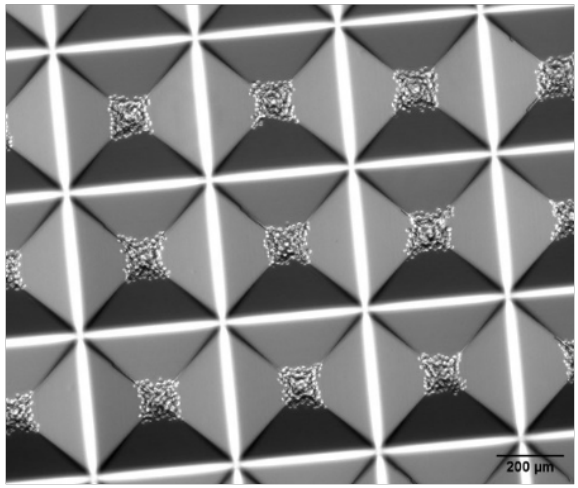


Calcein AM EthD-1 Hoechst 33342

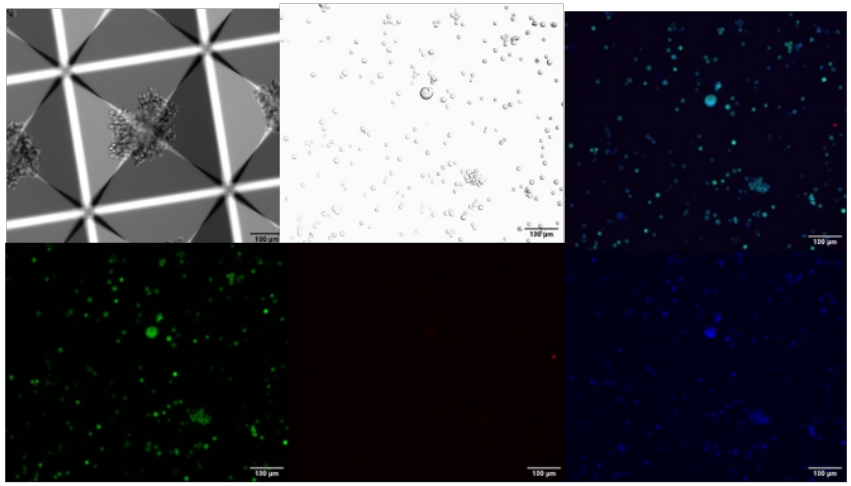
B

SW620
100 cells/microwell

Day 0

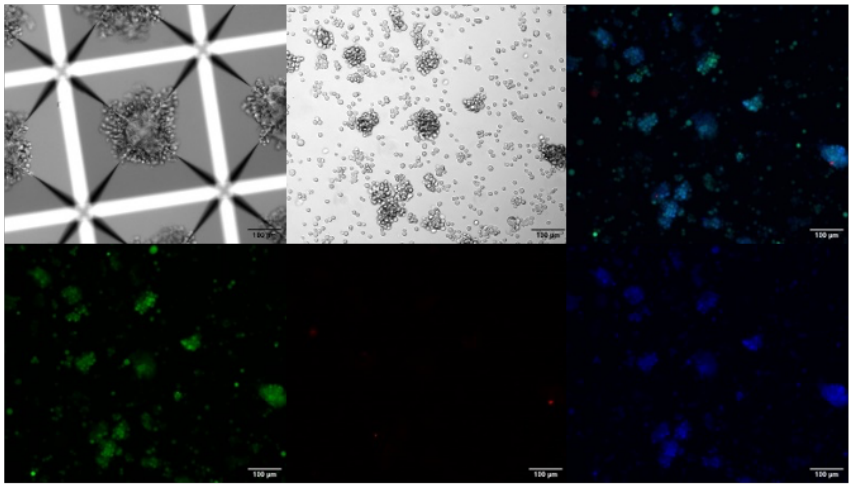


Day 3



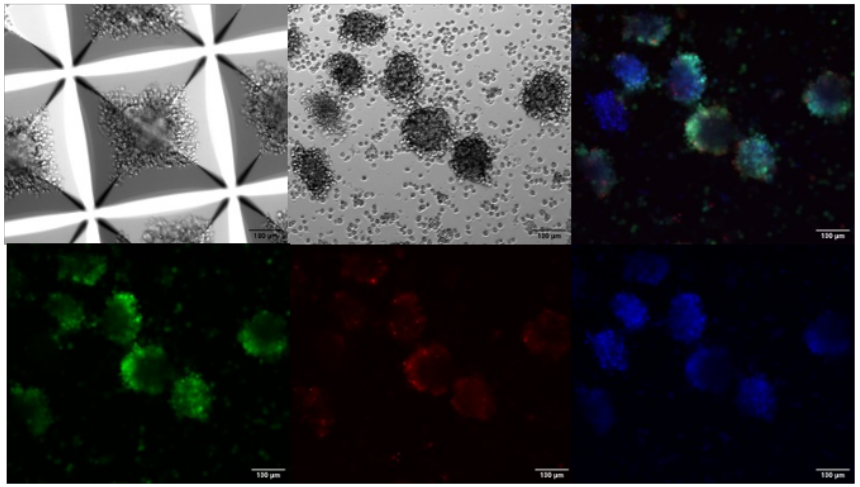
Calcein AM EthD-1 Hoechst 33342

Day 5



Calcein AM EthD-1 Hoechst 33342

Day 7



Calcein AM EthD-1 Hoechst 33342

Figure 6-18 SW620 spheroid formation during culture on an Aggrewell™ plate

Images showed SW620 cells on an Aggrewell™ plate with seeding densities of 50 cells per microwell (A) and 100 cells per microwell (B) over days 0, 3, 5, and 7 of culture. Spheroid formation was observed on day 5. Calcein AM (green) indicated live cells, EthD-1 (red) indicated dead cells, and Hoechst 33342 (blue) stained nuclei. The scale bar = 200 μm in Figures for day 0. The scale bar = 100 μm in Figures for days 3, 5, and 7.

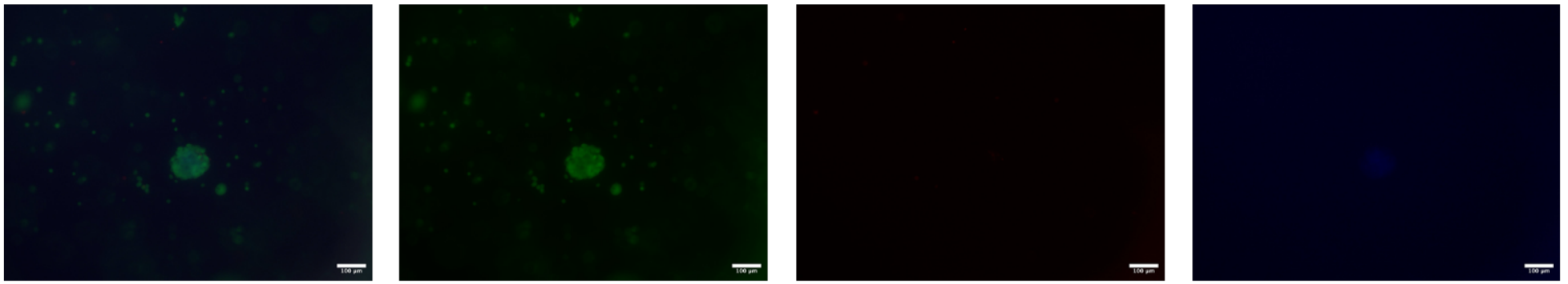
6.5.5 Cell viability of culturing SW620 spheroids in hydrogels over 7 days

Subsequently, to create a multicell 3D model of metastatic CRC with other cell types, we cultured SW620 spheroids in a hydrogel composed of 2% Alg and 8% Gel using either the hand-casting hydrogel method or 3D bioprinting for a larger batch of replicates, as described in Section 2.2.7.2 to 2.2.7.4. Briefly, SW620 spheroids were established 5 days prior to performing the hand-casting or 3D bioprinting hydrogel. The spheroids were harvested, mixed with 2% Alg and 8% Gel, then a droplet of the hydrogel/spheroid mixture was cast or printed, and the hydrogels were crosslinked with 150 mM $CaCl_2$. The spheroids were cultured in hydrogels with the standard medium for 7 days and stained with Live/Dead to evaluate their viability and morphology on days 0, 1, 3, 5, and 7 after casting, as described in Section 2.2.7.6.

Representative images of SW620 spheroids within a 2% Alg and 8% Gel hydrogel showed changes in spheroid morphology and an increase in size over 7 days. The spheroids expressed BFP and were stained using Live/Dead staining (Figure 6-19A to E). Histograms showing the percentages of the corrected total cell fluorescence (CTCF) relative to the total CTCF for live and dead spheroids indicated a decline in spheroid survival rates within the hydrogels over the 7-day period. However, the average percentage of dead cells on day 7 remained below 15%, indicating that at least 85% of the spheroids were viable in physiological 3D cultures throughout the time frame (Figure 6-19F). An increase in BFP levels was observed as the spheroids grew within the hydrogels over 7 days (Figure 6-19G). Similarly, spheroid size, measured by area, expanded over time (Figure 6-19H). Regarding aspect ratio, a value of 1 denotes an isotropic sphere, while the spheroids in hydrogels had ratios above 1, indicating a prolate shape, with no significant change observed over the culture period (Figure 6-19I). Additionally, the roundness of the spheroids ranged from 0.5 to 1 over 7 days, reflecting typical biological characteristics of spheroids. Differences in spheroid shapes were detected on day 3 compared to day 0. Overall, the spheroids displayed less circularity and compactness compared to perfect, highly circular, and compact spheroids, which have a roundness of 1 (Figure 6-19J).

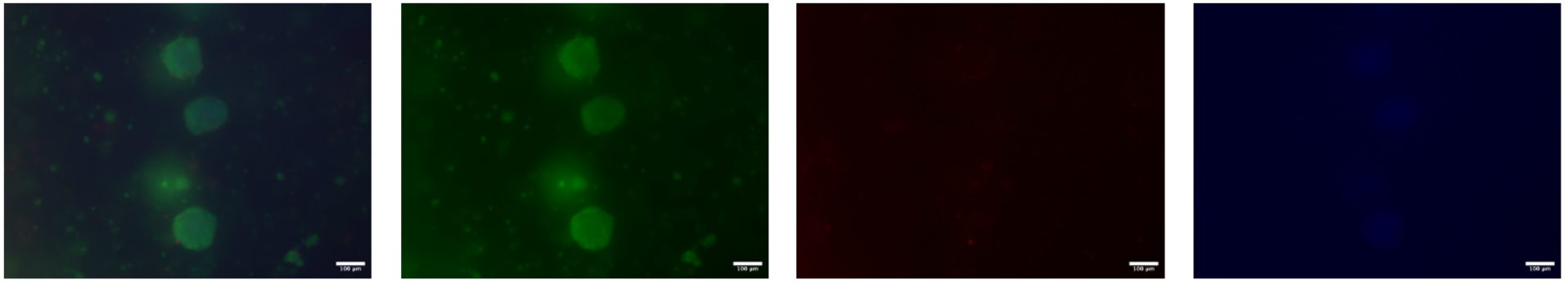
A

Day 0


Merge
Calcein AM
EthD-1
Hoechst 33342

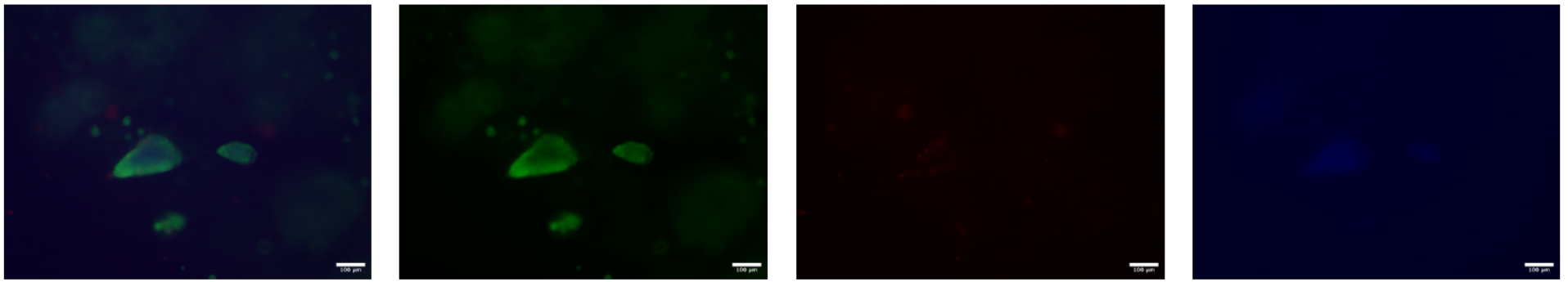
B

Day 1


Merge
Calcein AM
EthD-1
BFP

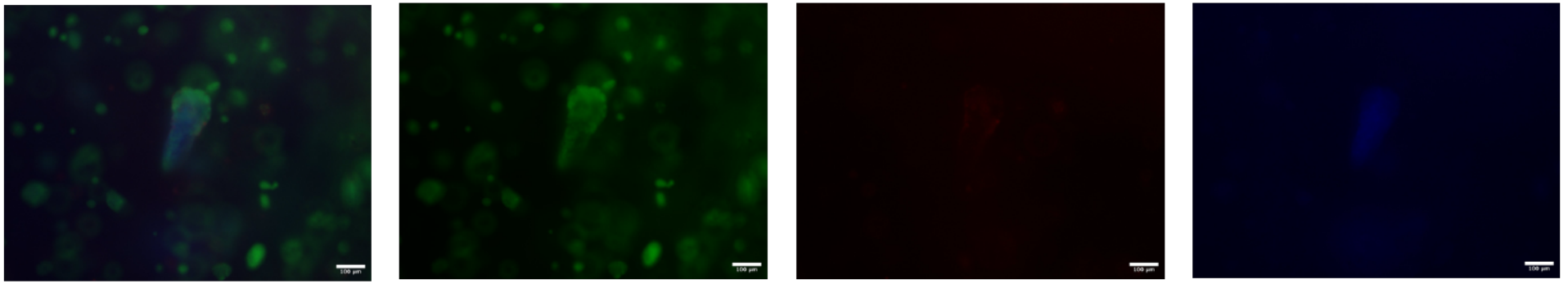
C

Day 3



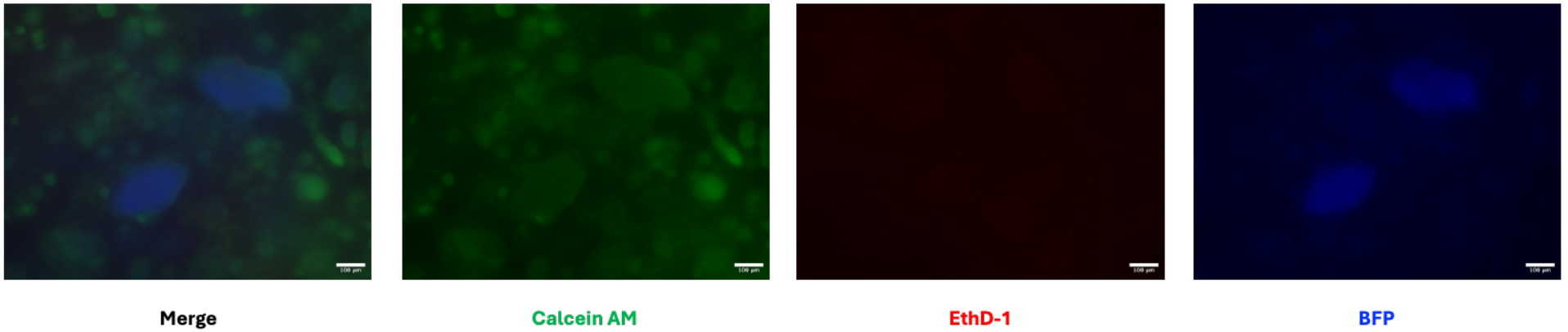
D

Day 5



E

Day 7



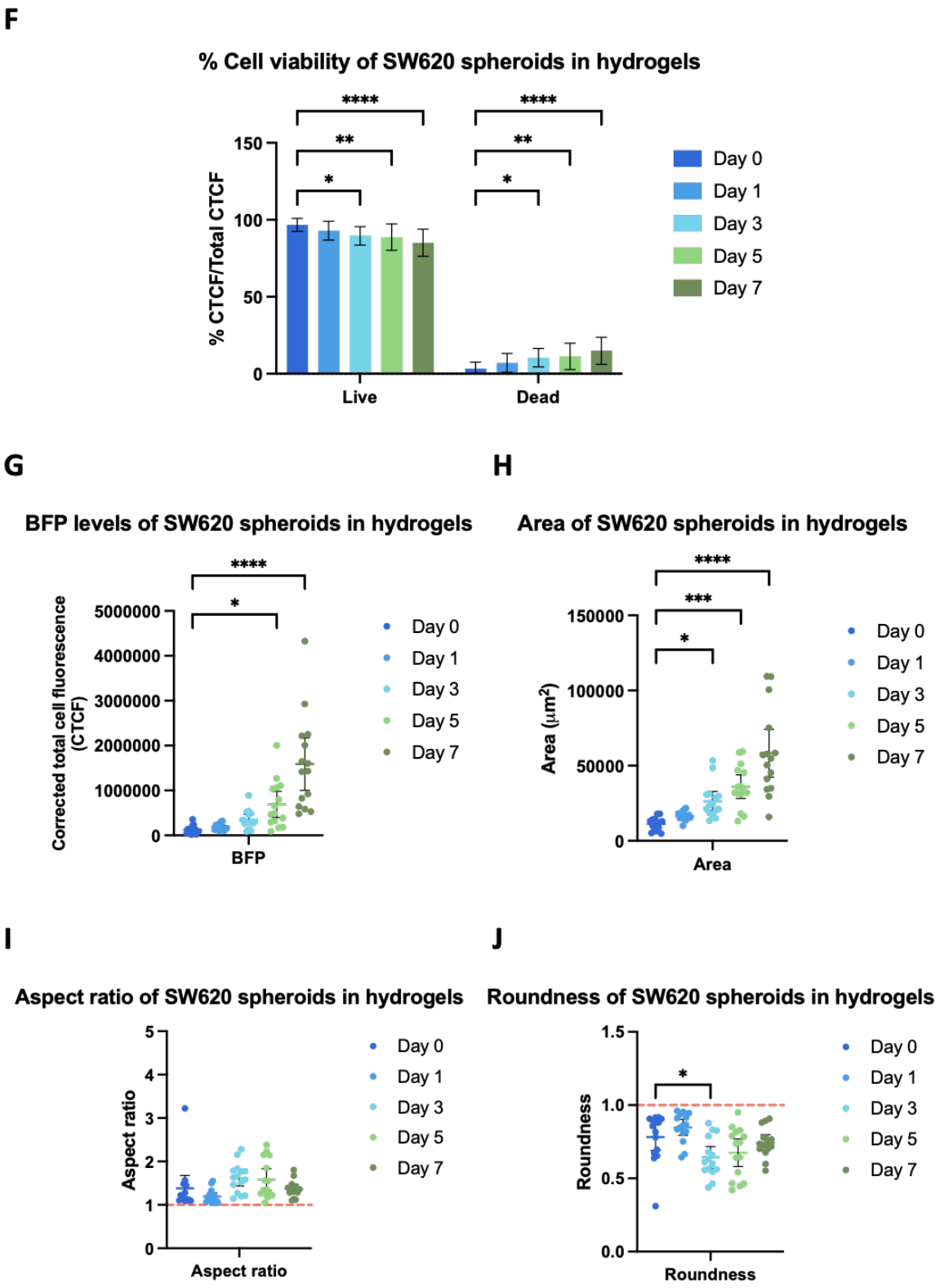


Figure 6-19 Cell viability of SW620 spheroids within hydrogels over 7 days

Images showed the morphology of SW620 spheroids in 3D-printed 2% Alginate 8% Gelatin hydrogels, which expressed BFP and were stained using Live/Dead staining over days 0 (A), 1 (B), 3 (C), 5 (D), and 7 (E) of culture. Calcein AM (green) indicated live cells, EthD-1 (red) indicated dead cells, and fluorescent BFP (Blue). The histogram comparing the percentages of the corrected total cell fluorescence (CTCF) relative to the total CTCF for live and dead spheroids showed a decrease in cell viability over 7 days (F). The histograms also showed increasing BFP levels (G) and the area of spheroids over time (H). The histograms showed that the aspect ratio from all conditions was

above 1, indicating a prolate shape (I), and the roundness from all conditions was below 1, indicating lower circularity (J). Aspect ratio = 1 (an isotropic sphere). Roundness = 1 (a highly circular and compact spheroid). Dunnett's or Sidak's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, 15 spheroids per condition, mean ± SD). The scale bar = 100 µm in Figures A-E.

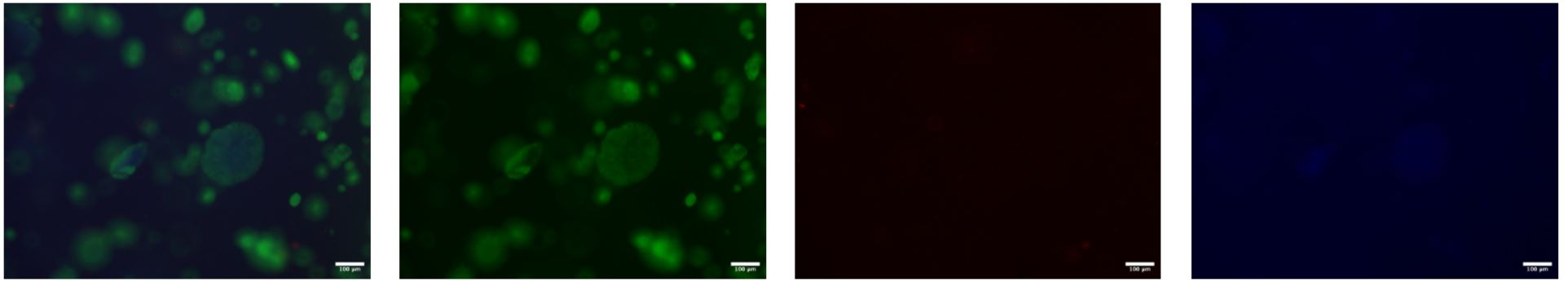
6.5.6 Inhibition effects of the drug treatment alone and in combination between LDN212854 and 5-FU on cell death of SW620 spheroids in hydrogels on days 4 and 7

To evaluate the inhibitory effects of targeting TGF-ß/BMP signalling pathways, both alone and in combination with chemotherapy, in 3D cultures of CRC metastatic cells. We performed 3D printing of SW620 spheroids in hydrogels as previously described. After 24 hours of casting spheroids in hydrogels, the drug treatments—20 µM of 5-FU, 2 µM of LDN-212854, and a combination of both drugs (20 µM of 5-FU and 2 µM of LDN-212854)—were applied, and the results were analysed using Live/Dead staining on day 4 and day 7, as described in Section 2.2.7.5 and 2.2.7.6.

On day 4 of drug treatment of SW620 spheroids in hydrogels, representative images showed variability in spheroid morphology and size across different conditions. These included cultures with standard medium, vehicle, 20 µM of 5-FU, 2 µM of LDN-212854, and a combination of 20 µM of 5-FU and 2 µM of LDN-212854, which expressed BFP and were stained with Live/Dead dyes (Figure 6-20A to E). Histograms comparing the percentage of live and dead spheroids indicated a significant reduction in cell viability with the drug combination compared to the vehicle, with an average of approximately 43% dead cells (Figure 6-20F). However, no differences were observed in BFP expression, spheroid size, aspect ratio, or roundness (Figure 6-20G to J).

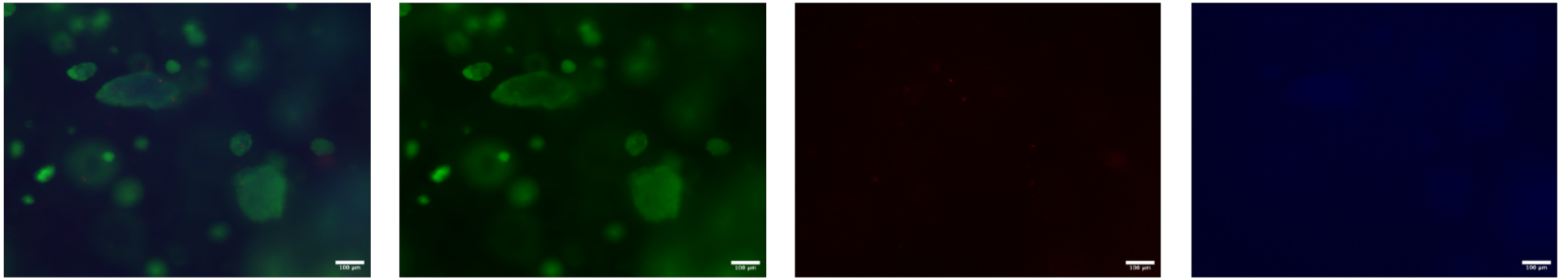
A Normal media

Treatment on Day 4



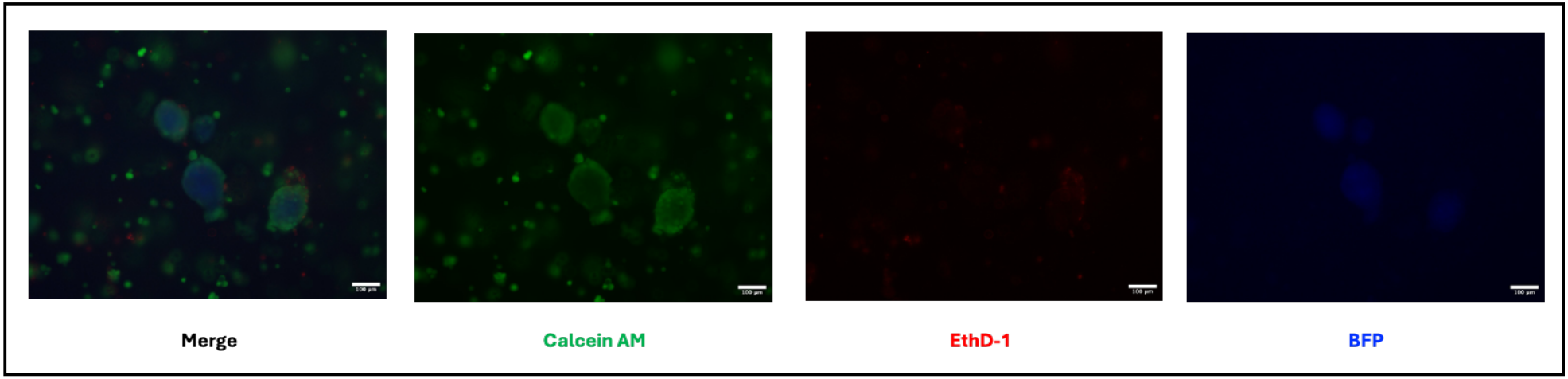
Merge | Calcein AM | EthD-1 | BFP

B Vehicle

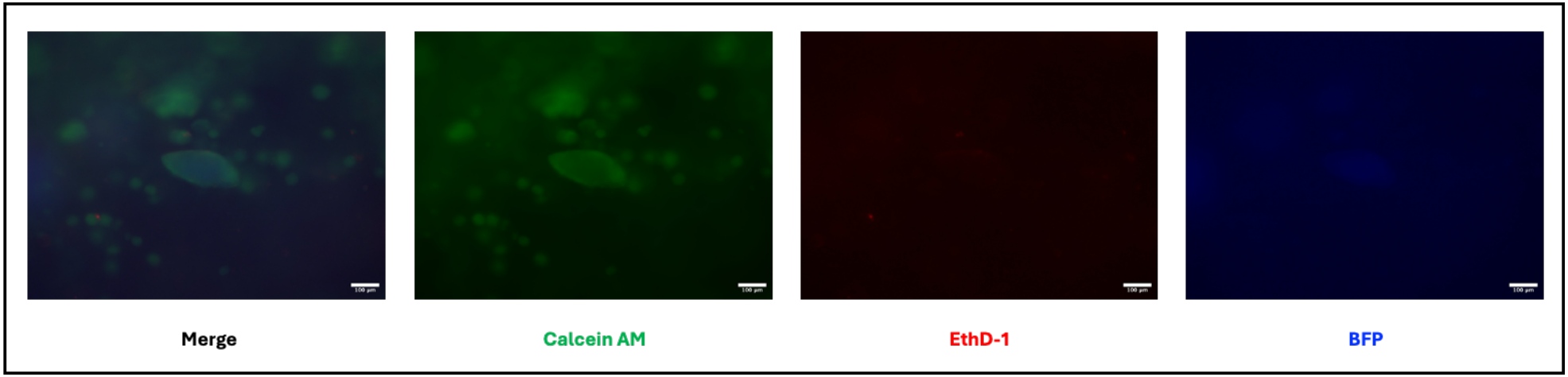


Merge | Calcein AM | EthD-1 | BFP

C 20 µM 5-FU

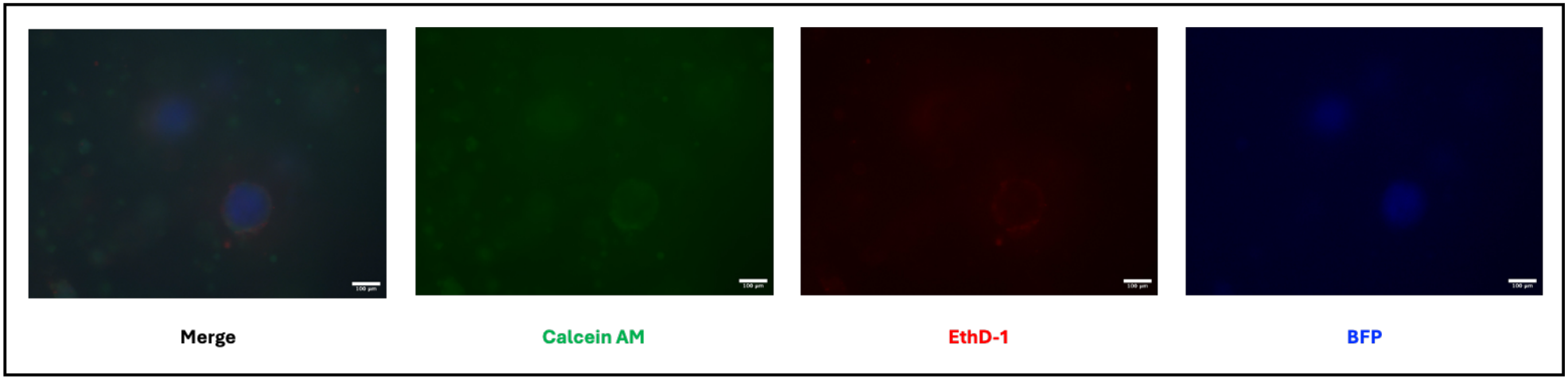


D 2 µM LDN-212854



E

Combo



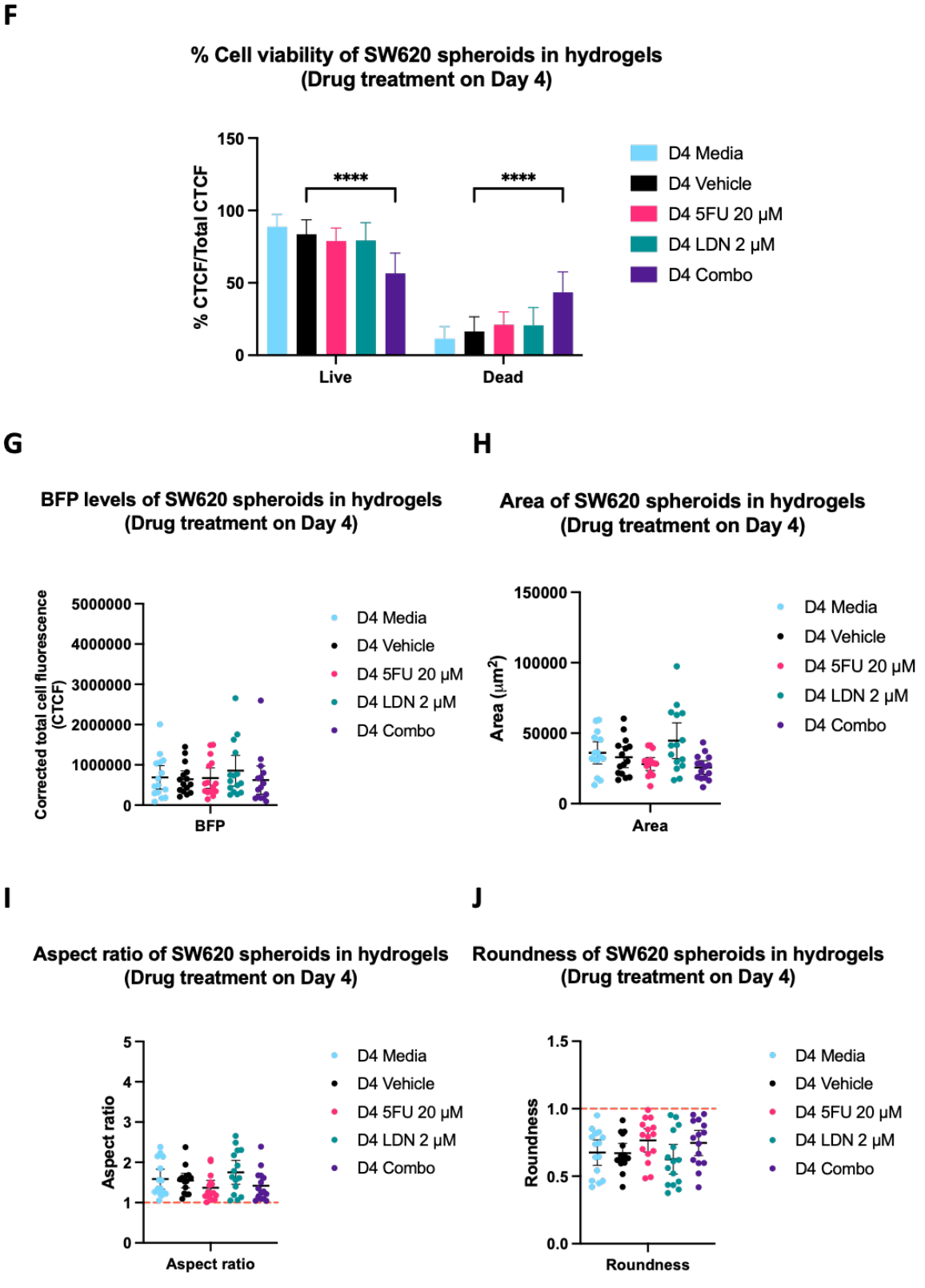


Figure 6-20 Inhibitory effects on cell viability of SW620 spheroids in hydrogels with LDN-212854 and 5-FU treatment on day 4

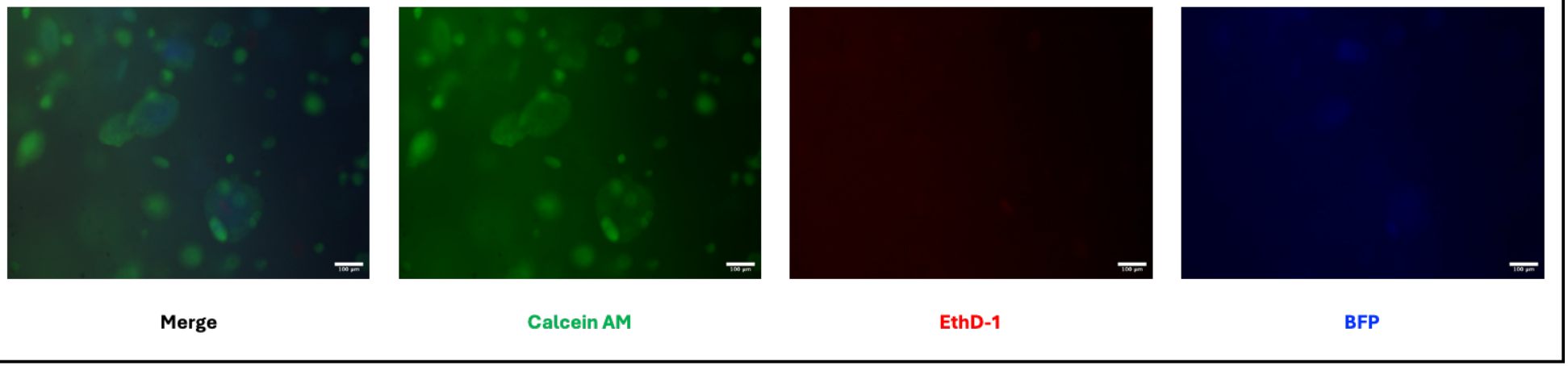
Images showed the morphology of SW620 spheroids in the hydrogels, which were treated with normal media (A), vehicle (B), 20 μM of 5-FU (C), 2 μM of LDN-212854 (D), and a combination of 20 μM of 5-FU and 2 μM of LDN-212854 (E) on day 4. Calcein AM (green) indicated live cells, EthD-1 (red) indicated dead cells, and fluorescent BFP (Blue). The histogram comparing the percentages of the corrected total cell fluorescence (CTCF) relative to the total CTCF for live and dead spheroids showed a significant decrease in cell viability in the combined drug treatment compared to the vehicle

(F). The histograms showed no difference in BFP levels (G), spheroid area (H), aspect ratio (I), and roundness (J) across treatment conditions. Aspect ratio = 1 (an isotropic sphere). Roundness = 1 (a highly circular and compact spheroid). Dunnett's or Sidak's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, 15 spheroids per condition, mean ± SD). The scale bar = 100 μm in Figures A-E.

Furthermore, the results from the drug treatment on day 7 showed that changes in spheroid appearance, due to a decrease in Calcein AM, exhibited lower intensity in the drug treatment groups, such as 5-FU and the drug combination (Figure 6-21A to E). Although it appeared to detect more cell death in the 5-FU treatment condition, only the drug combination significantly reduced cell viability, resulting in approximately 58% cell death compared to the vehicle (Figure 6-21F). A similar outcome to the Day 4 results—BFP expression, spheroid size, aspect ratio, or roundness across treatment conditions—was found to be not different (Figure 6-21G to J).

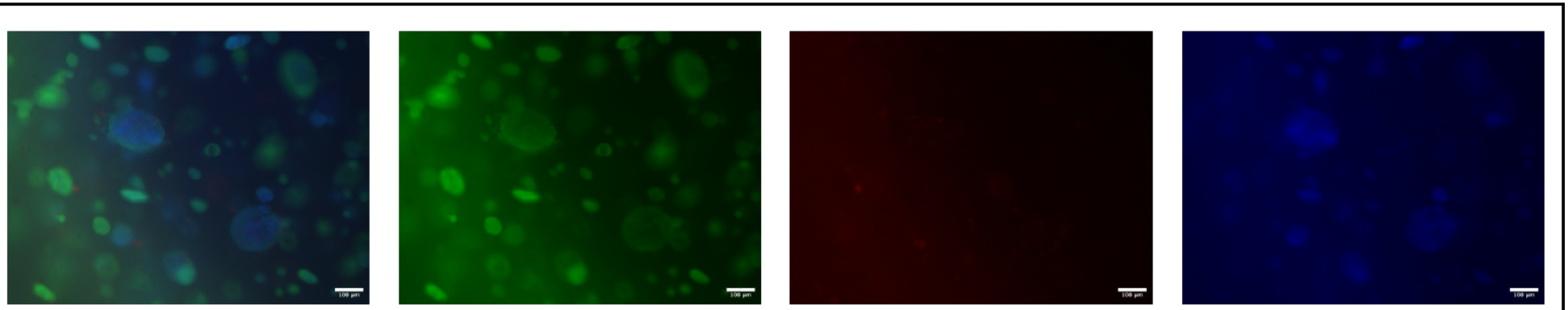
A Normal media

Treatment on Day 7



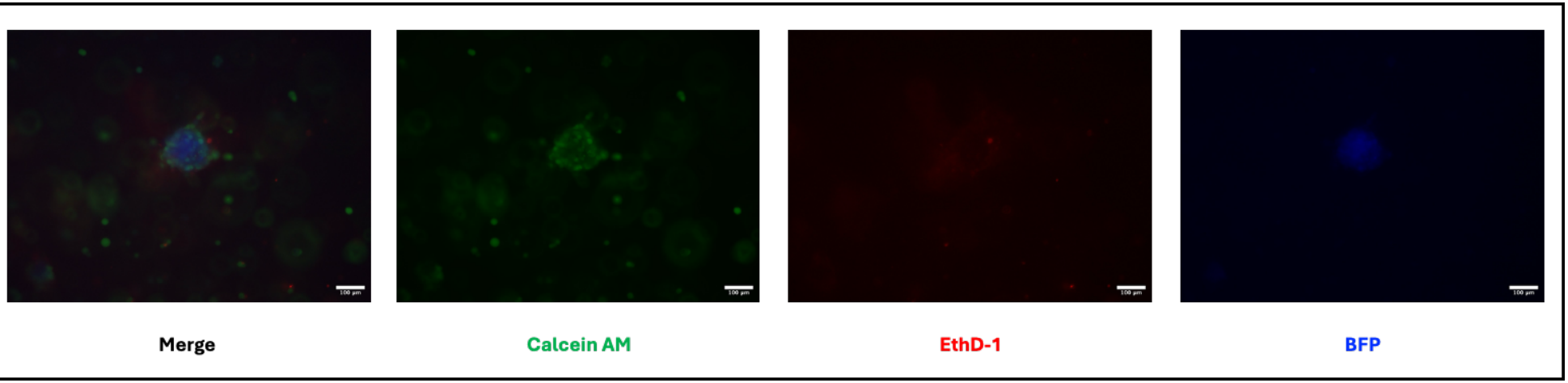
Merge | Calcein AM | EthD-1 | BFP

B Vehicle

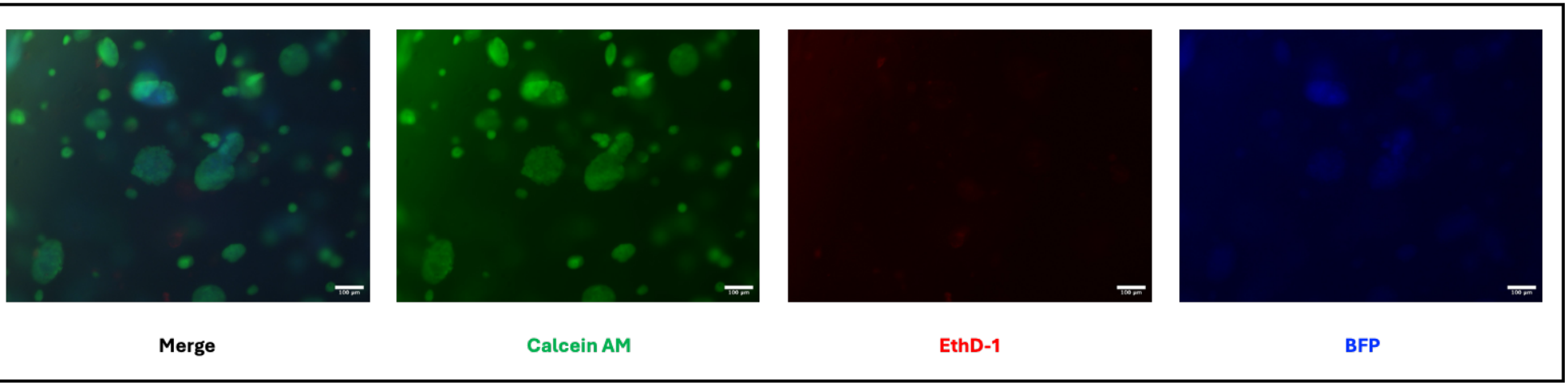


Merge | Calcein AM | EthD-1 | BFP

C 20 µM 5-FU

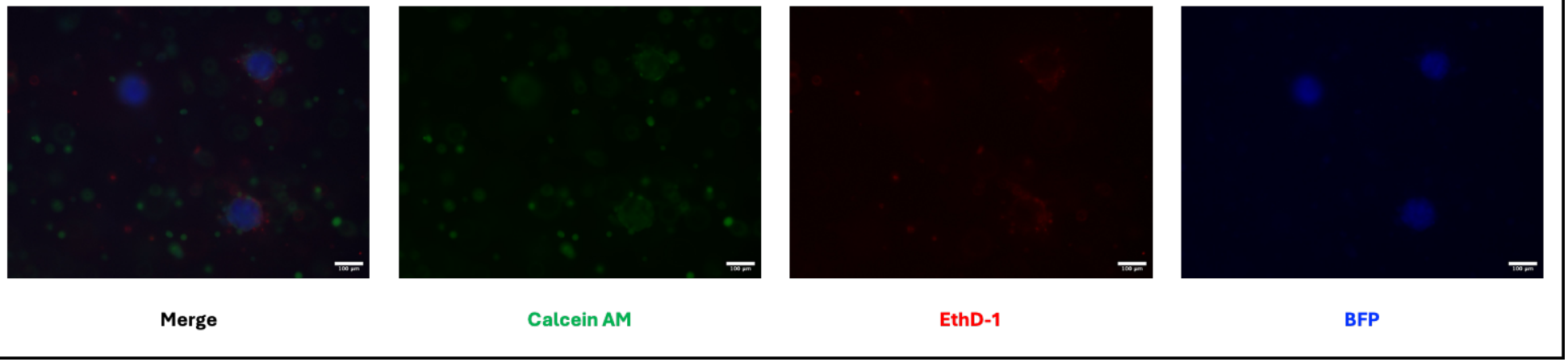


D 2 µM LDN-212854



E

Combo



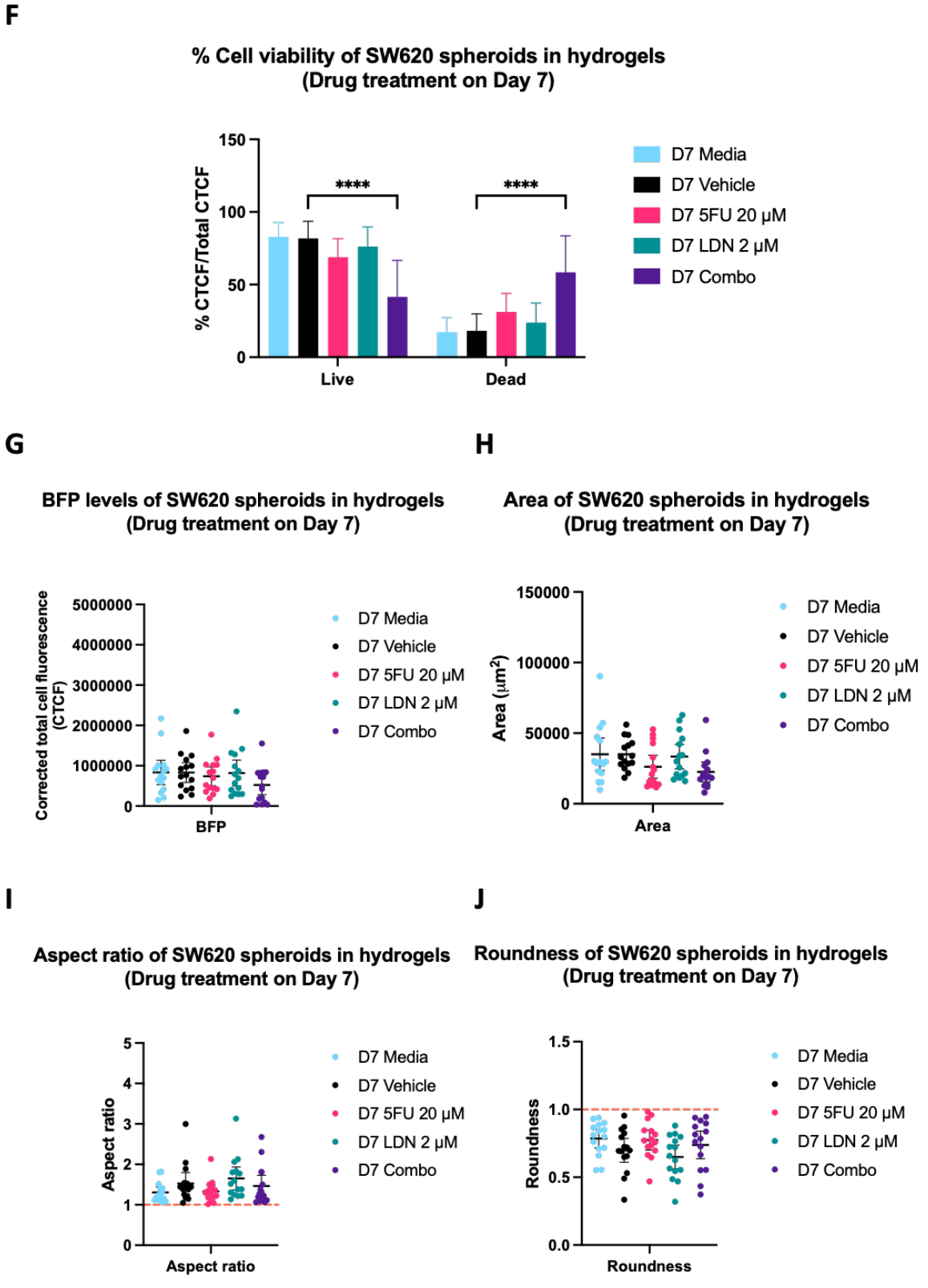


Figure 6-21 Inhibitory effects on cell viability of SW620 spheroids in hydrogels with LDN-212854 and 5-FU treatment on day 7

Images showed the morphology of SW620 spheroids in the hydrogels, which were treated with normal media (A), vehicle (B), 20 µM of 5-FU (C), 2 µM of LDN-212854 (D), and a combination of 20 µM of 5-FU and 2 µM of LDN-212854 (E) on day 7. Calcein AM (green) indicated live cells, EthD-1 (red) indicated dead cells, and fluorescent BFP (Blue). The histogram comparing the percentages of the corrected total cell fluorescence (CTCF) relative to the total CTCF for live and dead spheroids showed a significant decrease in cell viability in the combined drug treatment compared to the vehicle

(F). The histograms showed no difference in BFP levels (G), spheroid area (H), aspect ratio (I), and roundness (J) across treatment conditions. Aspect ratio = 1 (an isotropic sphere). Roundness = 1 (a highly circular and compact spheroid). Dunnett's or Sidak's multiple comparisons test * ($p \leq 0.05$), ** ($p \leq 0.01$), *** ($p \leq 0.001$), **** ($p \leq 0.0001$) compared with vehicle (n = 3, 15 spheroids per condition, mean ± SD). The scale bar = 100 μm in Figures A-E.

6.6 Discussion

Despite the fact that dysregulation of SMAD4 has prognostic significance in CRC and is associated with its aggressive features (237, 247), the interactions between TGF-β, BMP, and other signalling pathways form a complex mechanism that promotes progression to advanced CRC stages, facilitating invasion and metastasis to other tissues (2, 17, 18, 124, 125). The mechanisms involving phosphorylation of FGF/MAPK and Wnt/GSK3 signalling are highlighted, as they influence alterations in the tumour suppressor role of SMAD4 through polyubiquitination mediated by the E3 ligase β-TrCP (256). Our investigation into gene expression levels of various components in MAPK/ERK, Wnt/β-catenin, TGF-β/BMP, and PI3K signalling in CRC revealed diverse expression patterns at both gene and protein levels, as well as phosphorylation activity, across CRC cell lines. Notably, increased activation of the TGF-β/BMP and Wnt/β-catenin signalling pathways, along with elevated β-TrCP levels, was observed in SW620 cells with low SMAD4 expression. Additionally, these cells were stimulated individually with 20 ng/mL BMP4 and 20 ng/mL EGF for 4 hours. However, no change in SMAD4 protein levels was detected. A similar outcome was seen in DLD1 cells under these stimulation conditions. Interestingly, when the SW620 cells were under serum-starvation conditions, active β-catenin decreased, while p-ERK levels increased. Conversely, lower phosphorylation of MEK and p-ERK was observed in DLD1 cells. Furthermore, the cells were treated with bortezomib at concentrations of 10 and 30 μM for 24 hours, resulting in increased β-TrCP activity; however, this did not influence the regulation of SMAD4 levels in CRC. Nonetheless, this study provided preliminary results from one replicate, and other mechanisms may be involved in regulating SMAD4 activity, such as the regulation of SMAD4 Sumoylation by protein inhibitor of activated STAT1 (PIAS1) (257), or mediating SMAD4 feedback regulation of signalling pathways through targeting SMAD4-inhibited miRNAs (258). Therefore, further investigation is necessary to elucidate the intricate crosstalk mechanisms in CRC.

Considering the high heterogeneity of CRC tumours, it has been observed that patients with CRC in the CMS4 subtypes, characterised by prominent TGF-β activation and stromal infiltration (165), as well as those with loss of SMAD4 in their tumours (17, 124, 164-166), were predicted to have relatively worse survival outcomes and a poorer response to standard therapy. Consequently, we examined the inhibitory effects of novel inhibitors against BMPR1 in CRC cells, with the potential for therapeutic intervention. Among the three inhibitors (Dor, K02288, and LDN-212854), LDN-212854 was found to be the most potent inhibitor in reducing cell viability in DLD1 and SW620 cell lines. Treatment with a standard chemotherapeutic agent (5-FU) was validated, revealing that it required a high dose over a prolonged period to decrease cell viability in both cell types. Moreover, the combination of LDN-212854 and 5-FU showed synergistic effects in reducing cell viability in both DLD1 and SW620 cell lines after 48 to 72 hours. However, fewer nonviable cells were detected under treatment conditions. The drug treatment of 20 μM 5-FU and 3 μM LDN-212854 was performed in DLD1 cells, whereas SW620 cells were treated with 20 μM 5-FU and 2 μM LDN-212854, either alone or in combination. The results from Western blots confirmed that exposing cells to LDN-212854 led to the dephosphorylation of SMAD1/5/9 proteins between 4 and 24 hours. A cytotoxic effect of 5-FU treatment was demonstrated by the detection of H2AX phosphorylation and a trend towards increased cleaved caspase-3 after 48 hours.

Aiming to treat and prevent CRC metastasis, we further investigated the effects of the drug alone and in combination on apoptosis, the cell cycle, cell proliferation, and whether it was reversible by investigating the recovery in SW620 cells. Aside from a small number of apoptotic cells induced by these drug treatments, there was a consistently significant decrease in cell viability across all treatment conditions. Treating with LDN-212854 caused the cells to enter the G1 and G2 phases, while 5-FU-treated cells, both alone and in combination, arrested in the S phase after 48 hours of treatment. Within 96 hours, viable cells from LDN-212854 treatment were observed in several stages of division, with the majority in division 3. 5-FU-treated cells were predominantly in stage 2 of cell division. The combination treatment led to delayed cell proliferation as the treated cells remained in cell division 1. Additionally, monitoring cell recovery after drug treatment in a small number of SW620 cells cultured on a pre-coated

1% BME plate for 96 hours showed that the cells had reduced colony-forming ability within two weeks, especially with higher drug doses, and this effect was further enhanced with the drug combination. Further investigation will be necessary to explore the therapeutic impact on downstream signalling and more complex cellular mechanisms. For instance, it has been demonstrated that LDN-212854 exhibits therapeutic potential in inhibiting heterotopic ossification in an inducible transgenic mutant ALK2 mouse model of fibrodysplasia ossificans progressiva (FOP) (176). Overall, our results suggested that targeting the TGF-β/BMP signalling pathway in combination with standard chemotherapy is more effective than chemotherapy alone in preventing cell proliferation and survival of metastatic CRC cell lines.

Although preclinical drug screening and biological studies are often performed using 2D culture models for initial investigations, *in vitro* 3D culture models offer greater advantages for predicting drug efficacy and enabling comprehensive investigations as alternative tools that better reflect physiological diversity and complex interactions among different cell types (179, 180). Therefore, we proposed developing a 3D model of metastatic CRC that can incorporate other cell types, such as BM cells or hepatic cells. The ultimate aim being to develop a 3D multicellular model, constructed using spheroids of CRC cells and stromal cells embedded in an extracellular matrix (ECM), together with a monolayer of osteocytes expressing different fluorescent markers. Initially, fluorescent CRC cell lines, including DLD1 and SW620 cells, were established. Subsequently, the cell viability of CRC cell lines in conditioned media from BM cells (CAL72 and MSC) and hepatocellular carcinoma (HCC) cells (HepG2) was assessed. Our preliminary findings suggested that 20% conditioned media could alter growth patterns. The conditioned media of MSC and HepG2 induced a slightly higher proliferation rate of DLD1 cells. Additionally, SW620 cells also exhibited increased proliferation when cultured with conditioned media from MSC. In contrast, conditioned media from HepG2 and CAL72 reduced the proliferation of SW620 cells. When treated with LDN212854 and 5-FU, the CRC cells cultured in conditioned media, a worse response to the drug treatments was observed in DLD1 cells cultured in conditioned media from HepG2. Surprisingly, an improved response to the drug treatments was observed, especially when cultured with conditioned media from MSC. A previous study highlighted that AMPK/mTOR-mediated NF-κB was activated

as a result of MSC promoting CRC progression (259). Further investigation is necessary to gain a deeper understanding of these findings.

Nevertheless, developing a model for CRC metastasis remains challenging (179, 180). Here, we demonstrated an opportunity to produce batches of CRC spheroids using the Aggrewell™ plate. DLD1 cells seeded at 50-100 cells per microwell could form spheroids within 3 days, whereas SW620 cells required at least 5 days. Furthermore, culturing 3D bioprinted SW620 spheroids in 2% alginate and 8% gelatin hydrogels showed that the spheroids could survive at least at 85% viability in physiological 3D cultures for a week. An increase in spheroid size and higher fluorescence levels were observed over time. No significant morphological changes were noted, apart from a prolate shape. Only reduced circularity was observed on day 3 compared to day 0. Based on the benefits of the model used for drug screening, drug treatment alone and combined with LDN212854 and 5-FU were applied to SW620 spheroids in hydrogels. The results showed that, between days 4 and 7, the drug combination could induce around 43% to 58% cell death, leading to a reduction in cell viability. According to biomedical applications from using Alginate/gelatin (Alg-Gel) hydrogels, such as promoting cell adhesion and differentiation, guiding bone tissue formation involving MSCs, as well as modulating osteogenesis from the dental pulp stem cell (DPSC) (260), our results also suggested that these hydrogels could be applied in growing of CRC in 3D structure and use for drug screening. Future steps involve developing this model to incorporate other cell types to produce co-culture models of the BM and hepatic niches. This will enable us to explore cell-cell interactions, changes in cell behaviour, and responses to drug treatments.

In conclusion, this chapter demonstrated that diverse expression patterns of complex pathways, including MAPK/ERK signalling, Wnt/β-catenin signalling, TGF-β/BMP signalling, and PI3K signalling, were observed in CRC cell lines. However, the mechanisms regulating SMAD4 activity in CRC remain unclear. The drug combination of LDN-212854 and 5-FU exhibited synergistic effects in reducing CRC cell viability, arresting the cell cycle in the S phase, delaying cell proliferation, and inhibiting cell recovery. Several factors from the specific microenvironment were suggested to influence different alterations in CRC behaviour. The 3D bioprinted SW620 spheroids in 2% alginate and 8% gelatin hydrogel could serve as an alternative tool for studying metastatic CRC and drug

screening, as it provides more physiological interactions and supports tumour growth. The concordance of the inhibitory effects of the drug combination on decreasing cell viability in CRC was observed between the 3D model and the 2D model results. Developing a multicell 3D model of metastatic CRC cultured with other cell types will be the focus of future research, which would provide an alternative method for use in CRC research and drug discovery.

Chapter 7 Discussion

7.1 Understanding the different expression patterns of the TGF-β and BMP signalling pathways in CRC cell lines and cells within the bone marrow to identify dysregulation associated with CRC prognosis

TGF-β and BMP signalling pathways are recognised as essential mechanisms in controlling embryonic development, tissue homeostasis, and regeneration (162). These pathways facilitate cell proliferation and differentiation during crypt formation in the intestine (11-15). Dysregulation of these pathways in intestinal cells can impair their tumour-suppressing functions and promote tumour development (16-18). Mutations in the TGF-β and BMP pathways, such as those in *TGFBR2* and *SMAD4*, have been identified as contributing to CRC carcinogenesis, in both sporadic and hereditary cases (2, 3, 19-21). Other signalling components, such as *BMPR1A*, have also been found mutated in the premalignant and malignant stages of CRC (22, 23). The aim in this study was to determine which components of TGF-β and BMP signalling are associated with dysregulation in CRC prognosis, identifying potential targets for subsequent studies in prognostic research and drug discovery. The expression levels of TGF-β and BMP signalling components were analysed for any dysregulation in CRC cell lines using mRNA data from the DepMap database, and key proteins and their phosphorylation statuses were examined, as described in Chapter 3. The baseline gene expression of each TGF-β superfamily component was compared, revealing changes throughout CRC progression across several cell lines. Although ligand activation occurs at various levels, a notable activation of upstream receptors, such as TGFBR2, ACVR1B, and BMPR1A, was observed, alongside increased expression of signal transduction genes, including SMAD1, SMAD2, SMAD3, and SMAD5. Interestingly, while phosphorylation of SMAD1/5/8 was detected in SW620 cell lines (a metastatic colorectal cancer cell line), a downregulation of SMAD4 was identified.

Defining dysregulated components in TGF-β and BMP signalling pathways in CRC cell lines not only provides a fundamental understanding of their role in CRC during carcinogenesis but also serves as an informative tool for utilising the intrinsic expression data in subsequent studies. It is well established that CRC is a complex and highly heterogeneous disease (17). Several mutations and dysregulations occur

in tumours, contributing to their heterogeneity (18). Additionally, influences from the tumour microenvironment, which are involved in various signalling pathways, regulate biological processes that promote aggressive features. TGF-β and BMP signalling pathways are among the key mechanisms involved in CRC pathogenesis (9, 12). Among the CMS subtypes based on gene expression signatures, the activation of TGF-β, combined with stromal infiltration, EMT, and angiogenesis, is characteristic of CMS4 patients and is associated with poorer survival outcomes compared to other CMS groups (19). Furthermore, inactivation of SMAD4 has been shown to be associated with later stages of malignancy, CRC metastasis, and to predict poor prognosis (8, 20-23). However, a lack of a comprehensive understanding of these signalling pathways in CRC metastasis requires further investigation.

Metastatic CRC occurs in approximately 20-25% of patients with CRC and is associated with decreased survival rates. It can spread to various tissues, including lymph nodes, liver, lungs, peritoneum, bones, and the central nervous system (7). Among these sites, CRC with bone metastasis is rare—around 3-7%—but patients often experience poorer survival outcomes and limited treatment options (8-10). Additionally, restricted access to patient sample resources presents a challenge. Investigating how CRC spreads to bone appears to be complex (7, 131, 132). Several signalling pathways, including the TGF-β and BMP pathways, are crucial in maintaining bone homeostasis and are believed to regulate biological processes that promote metastatic CRC development in the bone marrow (8-10). This study demonstrated that high levels of TGF-β and BMP signalling pathways were observed across disease-free BM cells, as indicated by data from the Stemformatics database, which could imply the maintenance of homeostasis and serve as a reference for considering dysregulations related to disease conditions.

Although there was a considerable gap in understanding bone metastasis in CRC, other studies on bone metastasis of breast cancer have demonstrated that remodelling of the BM vascular microenvironment supports metastatic breast cancer (228). Additionally, hypoxia and TGF-β were observed in tumour cells and the bone microenvironment in promoting the spread of breast cancer to the bone. However, the heterogeneity of metastasis in the bone marrow, involving immune suppression and bone remodelling mechanisms, was identified through single-cell profiling of bone metastasis across multiple cancer types (227). Further

investigations examining the dysregulation of TGF-β and BMP signalling pathways in a co-culture model between metastatic CRC and BM cells could provide preclinical evidence to explain the specific mechanisms contributing to CRC metastasis and the bone microenvironment that facilitates invasion and therapeutic resistance, providing new avenues for development of targeted therapies in the future.

7.2 Low SMAD4 expression in colorectal adenomatous polyps predicted the development of metachronous polyps, particularly in tubulovillous polyps

During CRC tumourigenesis, adenomatous polyps are recognised as premalignant lesions in CRC. Patients with advanced adenomatous polyps are identified as having an increased risk of developing CRC (4-6). Surveillance colonoscopy in high-risk patients is recommended by the BSG2020 guideline, aiming to improve the prediction of metachronous polyps, which enables monitoring and removal of the polyps to prevent CRC development (6, 40, 46). However, it has been shown to be less effective in prognostic outcomes with a small number of metachronous polyp detections (49). Accordingly, the role of SMAD4 as a key mediator of the TGF-β and BMP signalling pathways in regulating tissue homeostasis has been demonstrated, while dysfunction of SMAD4 has been shown to be linked to colorectal polyps (232, 235). Analysis of epithelial SMAD4 expression in colorectal polyps using IHC and digital weighted histoscoring via QuPath provided a reliable method for assessing the relationship between dysregulated SMAD4 in adenomatous polyps and its clinical implications.

The findings in Chapter 4 revealed the prognostic significance of epithelial SMAD4 expression in colorectal adenomatous polyps in predicting the development of metachronous polyps. Overall, a low SMAD4 level in the polyps was associated with a higher likelihood of developing metachronous polyps or CRC in a Scottish cohort. Additionally, higher grades of dysplasia, histological subtypes, and the occurrence of metachronous polyps or CRC were found to correlate with decreased SMAD4 expression, even though the results were not linked with BSG2020 risk stratification. However, these could provide more informative and beneficial results for prospective clinical applications. A better understanding of changes in SMAD4 at the protein level in colorectal polyps was established, thereby enhancing

the observation of genetic mutations in *SMAD4* and in patients with hereditary conditions, as previously investigated (231-235). Moreover, it was first highlighted that decreased SMAD4 expression in adenomatous polyps occurs in the tubulovillous polyp subtype, which is associated with unfavourable development of metachronous lesions, while there were no prognostic differences in tubular and villous polyps. In comparison, other studies also highlighted the predictive value of SMAD4 expression in premalignant lesions. For instance, a combination of high SMAD4 expression with a high grade of dysplasia predicts a higher risk of malignant transformation in patients with oral leukoplakia, a premalignant lesion associated with oral squamous cell carcinoma (261). Whereas a low SMAD4 expression, combined with high lymphocyte infiltration, correlated with high-risk malignant transformation (262).

Transcriptomic analysis of tubulovillous polyps examining differential gene expression between groups with low and high SMAD4 levels revealed an upregulation of genes involved in protein deubiquitination in polyps with low SMAD4 levels. It also indicated a likely enrichment in tyrosine metabolism, PPAR signalling, arginine and proline metabolism, leukocyte transendothelial migration, and basal cell carcinoma. These findings suggested dysregulation mechanisms linked to a decrease in SMAD4 protein in tubulovillous polyps, which could be explored further in larger and more diverse cohorts. The integration of multi-omics analysis would offer a more comprehensive understanding of these changes, along with their application in clinical practice. Recent studies have integrated data from various omics platforms—including genomics, transcriptomics, and methylomics—in the analysis of colorectal polyps and the mechanisms involved in CRC development (241, 242). Furthermore, utilising image-based technology and deep learning in digital pathology (263) to incorporate data analysis from other candidate genes or relevant prognostic factors with SMAD4 expression in polyp tissues would broaden understanding of disease-specific features and improve prognostic prediction.

7.3 Low SMAD4 expression in colorectal cancer predicted worse survival outcomes, involving significant predictors in the right-sided colon and tumours with high stroma

CRC remains one of the leading aggressive diseases with a high rate of cancer-related deaths (31). The high heterogeneity and complexity of CRC tumours and their microenvironment are characterised by various mutations and dysregulations that occur during carcinogenesis, which are crucial for tumour progression and metastasis (2, 3, 19-21, 245). A specific tumour microenvironment has been identified for each location of CRC, which is associated with differences in tumour characteristics and clinical outcomes. Microsatellite instability and increased immune activity, as well as poorer outcomes, are associated with right-sided tumours, whereas classic tumours with activation of WNT and MYC signalling are linked to the left side (26, 55). Additionally, increased tumour-associated stroma has been shown to have prognostic value in predicting poorer survival outcomes in CRC (65). Besides the prominent TGF-β/BMP signalling activation in CMS4 with stromal invasion (57), these pathways were also found to be dysregulated throughout carcinogenesis (17, 124, 164), as they can switch functions between tumour suppression and tumour promotion (16-18). Furthermore, *SMAD4* mutation and a reduction of SMAD4 protein have been observed in association with an advanced stage of CRC and a shorter survival time (166, 237, 243, 244). Therefore, to better understand the relationship between SMAD4 and clinical characteristics in predicting CRC patient outcomes, two CRC cohorts were studied: the early CRC cohort and the GRI cohort of stages I-III.

As highlighted in Chapter 5, lower expression of tumour epithelial SMAD4 was associated with higher tumour stages, a finding consistent with other studies (237, 247). However, a contrasting result was observed, with a weak correlation found between SMAD4 expression and tumour locations in the early CRC cohort, as well as between SMAD4 expression and histological differentiation grade in the GRI cohort (237, 247). Interestingly, it was found that a significant correlation existed between decreased SMAD4 and higher tumour stroma in the GRI cohort. The prognostic significance of SMAD4 expression in CRC tumours showed that a low SMAD4 level predicted decreased cancer-specific survival in CRC patients, which is consistent with other reports (165, 237, 243, 247, 248). When considering

tumour locations, worse survival outcomes were predicted significantly only in tumours arising in the right colon with low SMAD4, whereas no difference in survival outcomes was observed between low and high SMAD4 levels in tumours arising in the left colon and rectum. Moreover, the worst survival outcome was associated with low SMAD4 expression in tumours and high stroma, compared to tumours with either low SMAD4 expression and low stroma or high SMAD4 expression with both low and high stroma. Another study also confirmed that the combination of tumour SMAD4 expression levels and stromal content has prognostic value (251).

A subsequent transcriptomic analysis focusing on CRC patients with high stromal content revealed that upregulated *SOD3* was present in tumours with higher SMAD4 expression. Additionally, gene sets such as the ECM, growth factor activity, and PI3K-Akt signalling were enriched in this group. In contrast, tumours with low SMAD4 expression are associated with increased aminopeptidase activity. These findings demonstrated changes in biological processes related to the dysregulation of SMAD4 in tumours with high stromal content, suggesting potential targets for further investigation to enhance understanding. Regarding the prognostic value of SOD3 expression in CRC tumours, a high level of SOD3 was predicted to be associated with better survival outcomes for CRC patients. In liver cancer, an increase in SOD3 was shown to decrease Ki-67 expression, suggesting a reduced rate of liver metastasis in an *in vivo* model (252). Consequently, our future work will focus on multiplex staining using either immunohistochemistry (mIHC) or immunofluorescence (mIF) (264) to examine the relationship between SMAD4, SOD3, and Ki-67 expression in CRC tumours, as well as their association with the tumour microenvironment. Additionally, another signalling component of the TGF-β/BMP signalling pathways, such as p-SMAD3, is currently being evaluated for its prognostic significance in CRC by members of our research group. Integrating these data would benefit clinical applications. Nonetheless, limitations of this study include a lack of a retrospective CRC cohort of metastatic CRC patients for validation of SMAD4 expression in cases of metastatic CRC. Therefore, tissues from patients with metastatic CRC would be valuable for further investigation.

7.4 Studying signalling mechanism alterations related to SMAD4 levels in CRC cells, revealing synergistic effects of TGF-β and BMP signalling inhibition in combination with standard chemotherapy in a metastatic CRC cell, and developing a model of metastasis

The complex mechanisms of various cell signalling pathways, including TGF-ß and BMP signalling, were highlighted in the regulation of CRC carcinogenesis, promoting development, progression, and ultimately metastasis (2, 17, 18, 124, 125). Decreases in SMAD4 protein in colorectal tissues have been associated with poorer outcomes at different stages of colorectal polyps and cancer, as previously described. In regulating SMAD4 functions, phosphorylation of FGF/MAPK and WNT/GSK3 signalling can prime SMAD4, leading to polyubiquitination mediated by the E3 ligase ß-TrCP (256). Our preliminary findings in Chapter 6 focused on changes in signalling mechanisms related to SMAD4 levels in CRC cell lines, which showed diverse expression patterns at both gene and protein levels, including phosphorylation activity, across the MAPK/ERK, WNT/ß-catenin, TGF-ß/BMP, and PI3K signalling pathways. For the SW620 cells, a metastatic CRC cell line with low SMAD4 expression, there was prominent activation of the TGF-ß/BMP and WNT/ß-catenin signalling pathways, along with elevated ß-TrCP levels. Stimulation with 20 ng/mL BMP4 or EGF for 4 hours did not increase SMAD4 protein expression. Additionally, decreased active ß-catenin and increased p-ERK levels were observed in SW620 cells under serum-starvation conditions. Increased ß-TrCP activity was detected following bortezomib treatment at 10 and 30 μM for 24 hours, without a change in SMAD4 levels. In contrast, under the same stimulation conditions, no significant changes in SMAD4 protein were seen in DLD1 cells, which are a high SMAD4 CRC cell line. Different regulatory activity was associated with lower phosphorylation of MEK and p-ERK in DLD1 cells under serum-starvation conditions. A similar result was observed with high ß-TrCP activity after bortezomib treatment. Due to time constraints in this study, multiple replicates would improve reliability and consistency with these findings, and further investigation of other regulatory mechanisms involved in SMAD4 activity, such as PIAS1 as a SUMO E3 Ligase in SMAD4 Sumoylation (257) and targeting SMAD4-inhibited miRNAs in mediating SMAD4 feedback regulation of signalling pathways

(258), would enhance understanding of the underlying mechanisms related to dysregulation of SMAD4 in CRC.

During the development and metastasis of CRC, high activation of TGF-ß/BMP signalling, along with the loss of SMAD4 in CRC patients, has been linked to shorter survival and a reduced response to standard treatments (17, 124, 164-166). In particular, metastatic CRC requires more effective therapy to prolong patients' lifespans (7). Therefore, novel inhibitors targeting BMPR1, including Dor, K02288, and LDN-212854, were evaluated in a 2D *in vitro* model to assess their effectiveness in treating CRC. The results indicated that LDN-212854 was the most potent inhibitor, as it decreased the cell viability of both DLD1 and SW620 cell lines. Furthermore, the combination of LDN-212854 with standard therapy (5-FU) demonstrated synergistic effects in reducing cell viability. Dephosphorylation of SMAD1/5/9 resulting from LDN-212854 treatment was observed, while 5-FU showed a cytotoxic effect through H2AX phosphorylation and cleaved caspase-3 detection. Moreover, the inhibitory effects of the drug alone and in combination were studied in a metastatic CRC cell line, revealing that the drug combination treatment of 2 µM LDN-212854 and 20 µM 5-FU resulted in greater inhibition of cell proliferation and survival compared to the drug alone. After treating SW620 cells with the drug combination, a significant decrease in cell viability, cell cycle arrest in the S phase, delayed cell division, and a greater reduction in colony-forming ability were observed, indicating a markedly reduced rate of cell survival. This drug combination was then further investigated for its therapeutic impact on downstream signalling and more complex cellular mechanisms to gain a deeper understanding of the underlying mechanisms.

Currently, *in vitro* 3D culture models have demonstrated greater advantages for preclinical drug screening and biological research, as they enable physiological diversity and complex interactions among different cell types, supporting comprehensive investigations (179, 180). In this study, we aimed to develop a 3D model of metastatic CRC with other cell types, such as BM cells or hepatic cells, as this would be useful for examining cell behaviour in a physiological environment and for testing novel drugs for curative and protective effects. For example, the proposed 3D multicellular model of metastatic CRC to bone marrow includes spheroids of fluorescent CRC cells and stromal cells embedded in an extracellular matrix (ECM), along with a monolayer of osteocytes expressing different

fluorescent markers. Establishing fluorescent CRC cell lines, including DLD1 and SW620 cells, was performed for use in the 3D model.

Additionally, preliminary results from 2D culturing of both CRC cell lines in conditioned media from BM cells (CAL72 and MSC) and hepatocellular carcinoma (HCC) cells (HepG2) showed that 20% conditioned media could alter growth patterns. Conditioned media from MSCs stimulated an increase in the proliferation of DLD1 and SW620 cells. Conditioned media from CAL72 reduced the proliferation rate of SW620 cells but had no effect on DLD1 cell proliferation. When culturing the cells with conditioned media from HepG2, the results showed contrasting effects on cell proliferation, with an increase in DLD1 cells and a decrease in SW620 cells. Furthermore, initial results from treating CRC cells cultured in conditioned media with LDN212854 and 5-FU indicated that, overall, there were no significant differences in inhibitory effects compared to standard treatment conditions, indicating the secretome in the conditioned media had minimal protective effects on the CRC cells. However, a better response to the drug treatments was observed when culturing CRC cells with conditioned media from MSC. A poorer response to LDN212854 treatment was seen in DLD1 cells cultured in conditioned media from HepG2 cells. In subsequent steps, repeating this examination with more biological replicates would provide stronger evidence, and investigating secretory cytokines in these conditioned media, along with changes in cell signalling pathways in CRC cells, would uncover unknown mechanisms related to culturing CRC with conditioned media from other cells and inform observations in a co-culture model.

Finally, we demonstrated the development of a 3D model of CRC metastasis by culturing SW620 spheroids in hydrogels to mimic the mechanophysical properties of the microenvironment and evaluate the inhibitory effects of TGF-β and BMP signalling inhibition in combination with standard chemotherapy. SW620 spheroids were formed using an Aggrewell™ plate over 5 days before being embedded in 2% alginate and 8% gelatin hydrogels, with a droplet of the gel cast using a 3D bioprinter. The 3D bioprinted SW620 spheroids in hydrogels maintained cell viability at >85%, showed an increase in spheroid size, higher fluorescence levels, and retained a prolate shape in physiological 3D cultures over a week. Moreover, an investigation into the efficacy of treating SW620 spheroids in hydrogels with LDN212854 and 5-FU alone and in combination revealed that the drug combination

produced a superior inhibitory effect, inducing cell death of approximately 43% to 58% between days 4 and 7. These findings suggested a high potential for growing CRC spheroids in a 3D structure, utilisation in drug screening, and further development of a 3D multicellular model. Due to the benefits of alginate-based hydrogels, such as supporting the structure of bone marrow, wound dressings, and drug delivery vehicles for cancer treatment (260, 265), incorporating other cell types to produce co-culture models of the BM and hepatic niches involving metastatic CRC will be explored in future studies. These would provide valuable resources for the comprehensive study of cell-cell interactions, changes in cell behaviour, and responses to drug treatments in metastatic CRC.

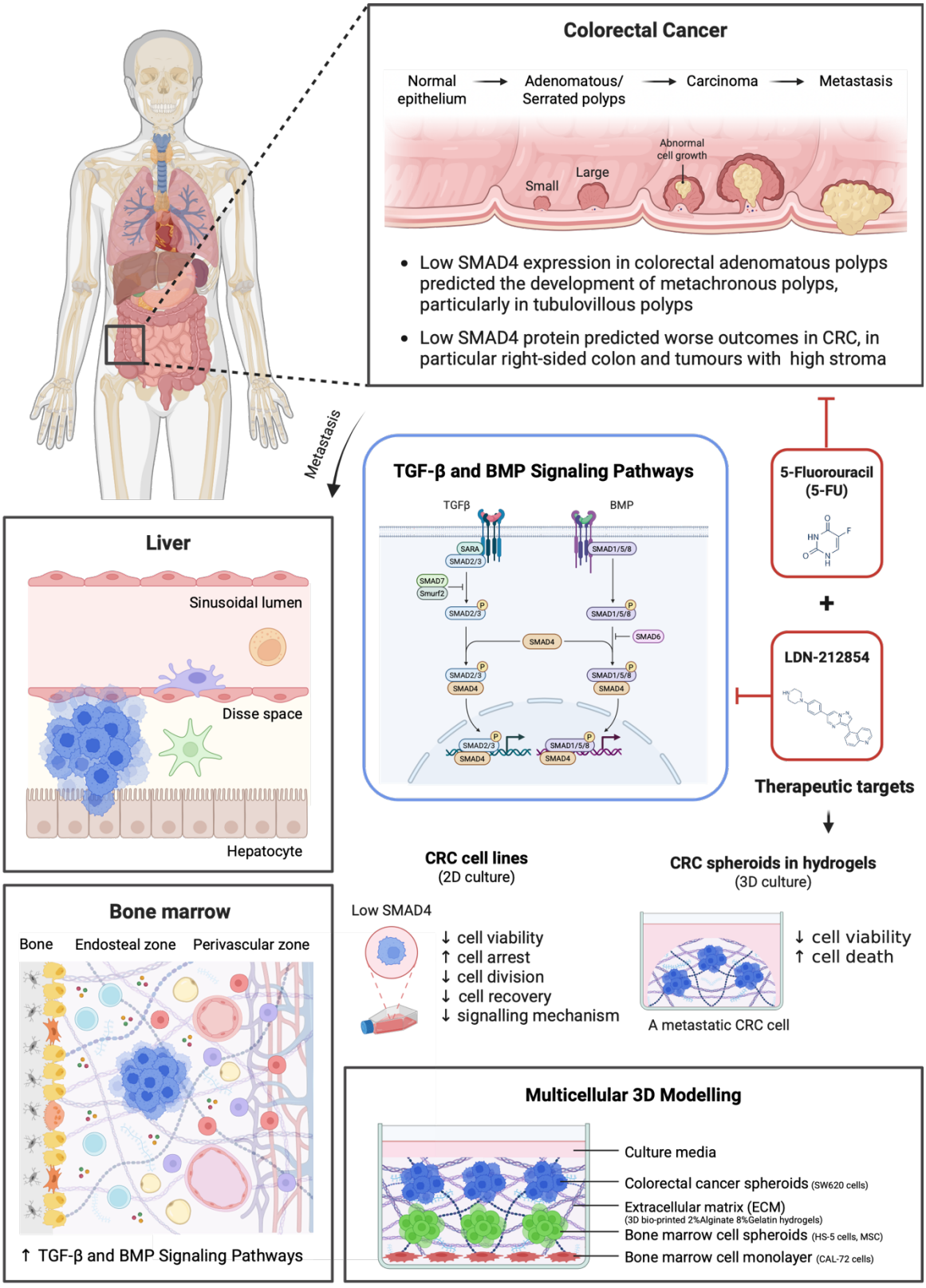


Figure 7-1 A schematic overview of the key findings from this thesis in elucidating the role of the TGF beta superfamily in metastatic spread of colorectal cancer

7.5 Future direction

This study revealed dysregulated expression patterns of the TGF-ß and BMP signalling pathways in CRC cell lines and provides valuable data for future research on CRC progression. Limited information on other metastatic CRC sites remains a challenge. Therefore, understanding the intrinsic expression of these pathways in disease-free conditions of other tissue sites would be informative, providing baseline expression levels in each specific tissue. We used the Stemformatics database to explore TGF-ß and BMP signalling pathways in several cell types within the bone marrow. However, the gene expression of TGF-ß and BMP signalling pathways in hepatic tissues requires further investigation, as it could offer insights into a co-culture model of metastatic CRC in the hepatic niche. Thus, public gene expression data from previous studies on normal liver tissues will be assessed. In the colorectal polyp study, which showed low SMAD4 expression, particularly in the tubulovillous polyp subtype, it was associated with an increased development of metachronous lesions. Transcriptomic analysis of these premalignant lesion subtypes indicated that protein degradation and other signalling mechanisms were involved. Incorporating genomics data or other relevant factors, such as dysplasia grades, into future analyses could enhance understanding of SMAD4 expression in colorectal polyps. Furthermore, validation of SMAD4 expression in other cohorts would strengthen its prognostic significance.

In the colorectal tissue study, the prognostic significance of SMAD4 expression was found to be associated with poorer survival outcomes, particularly in right-sided colon tumours and tumours with abundant stroma. Increased *SOD3* expression was observed in tumours with higher SMAD4 levels, as identified through transcriptomic analysis focusing on CRC patients with extensive stromal tissue surrounding the tumours. Multiplex staining of SMAD4, SOD3, and Ki-67 in CRC tumours is planned to be optimised to evaluate their relationships. Furthermore, combining the analysis of SMAD4 and p-SMAD3 expression may offer deeper insights into the mechanisms of the TGF-ß and BMP signalling pathways in CRC and their potential clinical importance. The regulatory mechanisms of SMAD4 activity in CRC still remain unclear. Conditions for stimulation will be further optimised, or other mechanisms such as PIAS1 may be investigated to uncover these complex processes. Additionally, the drug combination of LDN-212854 and 5-FU showed synergistic effects on cellular functions and activation of the TGF-ß and BMP

signalling pathways. Gene expression of downstream signalling will be studied to better understand the effects of the drug treatments. Regarding a 3D model of metastatic CRC development, it appears that CRC cell proliferation and response to drug treatments can be influenced by factors secreted from BM and hepatic cells. Further investigation into these cytokines and their effects on differential gene expression in CRC cells will be a focus. The 3D bioprinted SW620 spheroids in 2% alginate and 8% gelatin hydrogel present a promising method for culturing metastatic CRC within a physiological structure and are useful for drug screening. A multicellular 3D model of metastatic CRC cultured with BM and hepatic niches will be created for future CRC research and drug discovery. In addition, this *in vitro* system has the potential to be further developed to investigate patient derived organoids in the future.

7.6 Conclusion

Our *in silico* analysis revealed differential gene expression patterns of the TGF-β and BMP signalling pathways in CRC cell lines and cells within the BM, confirming active roles for TGF-β and BMP signalling in CRC progression and in the BM's maintenance of homeostasis. The absence of SMAD4 was also notable in metastatic CRC. Assessing IHC SMAD4 expression through digital weighted histoscoring provided prognostic value in predicting the development of metachronous polyps in patients with low SMAD4 expression in polyps, especially in tubulovillous polyps. Additionally, low SMAD4 expression in colorectal cancer was associated with poorer survival outcomes, highlighting its importance as a predictor in right-sided colon tumours and tumours with high stroma. Nevertheless, the mechanisms controlling SMAD4 activity in CRC remain unclear. Various expression patterns of complex pathways, including MAPK/ERK signalling, WNT/β-catenin signalling, TGF-β/BMP signalling, and PI3K signalling, were observed in CRC cell lines. The drug combination of LDN-212854 and 5-FU showed synergistic effects on CRC cell viability, cell cycle arrest, cell proliferation, and cell recovery in an *in vitro* 2D study. Several factors from BM and hepatic cells induced alterations in CRC behaviour. The 3D bioprinted SW620 spheroids in 2% alginate and 8% gelatin hydrogel provided physiological interactions, supported tumour growth, and served as a valuable tool for drug screening. Developing an *in vitro* multicell 3D model of metastatic CRC cultured with BM and hepatic niches would be beneficial in future CRC research and drug discovery.

List of References

1. Hossain MS, Karuniawati H, Jairoun AA, Urbi Z, Ooi J, John A, et al. Colorectal Cancer: A Review of Carcinogenesis, Global Epidemiology, Current Challenges, Risk Factors, Preventive and Treatment Strategies. Cancers (Basel). 2022;14(7).

2. Li Q, Geng S, Luo H, Wang W, Mo YQ, Luo Q, et al. Signaling pathways involved in colorectal cancer: pathogenesis and targeted therapy. Signal Transduct Target Ther. 2024;9(1):266.

3. Testa U, Pelosi E, Castelli G. Colorectal cancer: genetic abnormalities, tumor progression, tumor heterogeneity, clonal evolution and tumor-initiating cells. Med Sci (Basel). 2018;6(2).

4. Mohamed M, Schofield JB. The pathology of colorectal polyps and cancers (including biopsy). Surgery (Oxford). 2014;32(4):165-71.

5. Pickhardt PJ, Pooler BD, Kim DH, Hassan C, Matkowskyj KA, Halberg RB. The Natural History of Colorectal Polyps: Overview of Predictive Static and Dynamic Features. Gastroenterol Clin North Am. 2018;47(3):515-36.

6. Rutter MD, East J, Rees CJ, Cripps N, Docherty J, Dolwani S, et al. British Society of Gastroenterology/Association of Coloproctology of Great Britain and Ireland/Public Health England post-polypectomy and post-colorectal cancer resection surveillance guidelines. Gut. 2020;69(2):201-23.

7. Biller LH, Schrag D. Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review. JAMA. 2021;325(7):669-85.

8. Holladay L, Luu J, Balendra V, Kmetz K. Current and potential treatment of colorectal cancer metastasis to bone. Cancer Treat Res Commun. 2023;37:100763.

9. Gigante I, Tutino V, De Nunzio V, Notarnicola M. Colorectal Cancer and Bone Tissue: Fantastic Relations and Where to Find Them. Cancers (Basel). 2020;12(8).

10. Christensen TD, Jensen SG, Larsen FO, Nielsen DL. Systematic review: Incidence, risk factors, survival and treatment of bone metastases from colorectal cancer. J Bone Oncol. 2018;13:97-105.

11. Beumer J, Puschhof J, Yengej FY, Zhao L, Martinez-Silgado A, Blotenburg M, et al. BMP gradient along the intestinal villus axis controls zonated enterocyte and goblet cell states. Cell Rep. 2022;38(9):110438.

12. Beumer J, Clevers H. Cell fate specification and differentiation in the adult mammalian intestine. Nat Rev Mol Cell Biol. 2021;22(1):39-53.

13. Wang RN, Green J, Wang Z, Deng Y, Qiao M, Peabody M, et al. Bone Morphogenetic Protein (BMP) signaling in development and human diseases. Genes Dis. 2014;1(1):87-105.

14. Baulies A, Angelis N, Li VSW. Hallmarks of intestinal stem cells. Development. 2020;147(15).

15. Malijauskaite S, Connolly S, Newport D, McGourty K. Gradients in the *in vivo* intestinal stem cell compartment and their *in vitro* recapitulation in mimetic platforms. Cytokine Growth Factor Rev. 2021;60:76-88.

16. Akhurst RJ, Derynck R. TGF-beta signaling in cancer--a double-edged sword. Trends Cell Biol. 2001;11(11):S44-51.

17. Itatani Y, Kawada K, Sakai Y. Transforming Growth Factor-beta Signaling Pathway in Colorectal Cancer and Its Tumor Microenvironment. Int J Mol Sci. 2019;20(23).

18. Yeh HW, Lee SS, Chang CY, Lang YD, Jou YS. A New Switch for TGFbeta in Cancer. Cancer Res. 2019;79(15):3797-805.

19. Nguyen LH, Goel A, Chung DC. Pathways of Colorectal Carcinogenesis. Gastroenterology. 2020;158(2):291-302.

20. Hinoi T. Cancer Genomic Profiling in Colorectal Cancer: Current Challenges in Subtyping Colorectal Cancers Based on Somatic and Germline Variants. J Anus Rectum Colon. 2021;5(3):213-28.

21. Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. N Engl J Med. 2009;361(25):2449-60.

22. Nieminen TT, Abdel-Rahman WM, Ristimaki A, Lappalainen M, Lahermo P, Mecklin JP, et al. BMPR1A mutations in hereditary nonpolyposis colorectal cancer without mismatch repair deficiency. Gastroenterology. 2011;141(1):e23-6.

23. Poaty H, Batamba Bouya L, Ondima IPL, Mongo-Onkouo A, Lumaka Zola A, Peko JF, et al. BMPR1A and SMAD4 mutations in juvenile polyposis syndrome: clinicopathological and genetic data from two congolese patients. Gene Reports. 2021;23.

24. Schneider A, Feussner H. Anatomy, Physiology, and Selected Pathologies of the Gastrointestinal Tract. Biomedical Engineering in Gastrointestinal Surgery2017. p. 11-39.

25. Kishore S, Chong SKF. Intestine: Normal Development, Structure and Function. Textbook of Clinical Pediatrics2012. p. 1811-6.

26. Puccini A, Marshall JL, Salem ME. Molecular Variances Between Right- and Left-sided Colon Cancers. Current Colorectal Cancer Reports. 2018;14(5):152-8.

27. Dekker E, Tanis PJ, Vleugels JLA, Kasi PM, Wallace MB. Colorectal cancer. Lancet. 2019;394(10207):1467-80.

28. Humphries A, Wright NA. Colonic crypt organization and tumorigenesis. Nat Rev Cancer. 2008;8(6):415-24.

29. Thursby E, Juge N. Introduction to the human gut microbiota. Biochem J. 2017;474(11):1823-36.

30. Shussman N, Wexner SD. Colorectal polyps and polyposis syndromes. Gastroenterol Rep (Oxf). 2014;2(1):1-15.

31. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-63.

32. Bowel cancer statistics: Cancer Research UK; [Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-Zero.

33. Clark GR, Anderson AS, Godfrey TG, Strachan JA, Fraser CG, Steele RJ. Variation in changes in the incidence of colorectal cancer by age and association with screening uptake: an observational study. BMJ Open. 2020;10(9):e037925.

34. Sawicki T, Ruszkowska M, Danielewicz A, Niedzwiedzka E, Arlukowicz T, Przybylowicz KE. A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis. Cancers (Basel). 2021;13(9).

35. Rawla P, Sunkara T, Barsouk A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. Prz Gastroenterol. 2019;14(2):89-103.

36. Granados-Romero JJ, Valderrama-Treviño AI, Contreras-Flores EH, Barrera-Mera B, Herrera Enríquez M, Uriarte-Ruíz K, et al. Colorectal cancer: a review. International Journal of Research in Medical Sciences. 2017;5(11).

37. Ladabaum U, Dominitz JA, Kahi C, Schoen RE. Strategies for Colorectal Cancer Screening. Gastroenterology. 2020;158(2):418-32.

38. Steele RJ, McDonald PJ, Digby J, Brownlee L, Strachan JA, Libby G, et al. Clinical outcomes using a faecal immunochemical test for haemoglobin as a first-

line test in a national programme constrained by colonoscopy capacity. United European Gastroenterol J. 2013;1(3):198-205.

39. Clark G, Strachan JA, Carey FA, Godfrey T, Irvine A, McPherson A, et al. Transition to quantitative faecal immunochemical testing from guaiac faecal occult blood testing in a fully rolled-out population-based national bowel screening programme. Gut. 2021;70(1):106-13.

40. Winter J, Clark G, Steele R, Thornton M. Post-colonoscopy cancer rates in Scotland from 2012 to 2018: A population-based cohort study. Colorectal Dis. 2025;27(1):e17298.

41. Zhou H, Zhu L, Song J, Wang G, Li P, Li W, et al. Liquid biopsy at the frontier of detection, prognosis and progression monitoring in colorectal cancer. Mol Cancer. 2022;21(1):86.

42. Fadlallah H, El Masri J, Fakhereddine H, Youssef J, Chemaly C, Doughan S, et al. Colorectal cancer: Recent advances in management and treatment. World J Clin Oncol. 2024;15(9):1136-56.

43. McMillan DC. An inflammation-based prognostic score and its role in the nutrition-based management of patients with cancer. Proc Nutr Soc. 2008;67(3):257-62.

44. McMillan DC, Crozier JE, Canna K, Angerson WJ, McArdle CS. Evaluation of an inflammation-based prognostic score (GPS) in patients undergoing resection for colon and rectal cancer. Int J Colorectal Dis. 2007;22(8):881-6.

45. Proctor MJ, Morrison DS, Talwar D, Balmer SM, O'Reilly DS, Foulis AK, et al. An inflammation-based prognostic score (mGPS) predicts cancer survival independent of tumour site: a Glasgow Inflammation Outcome Study. Br J Cancer. 2011;104(4):726-34.

46. Quyn AJ, Fraser CG, Stanners G, Carey FA, Rees CJ, Moores B, et al. Scottish Bowel Screening Programme colonoscopy quality - scope for improvement? Colorectal Dis. 2018;20(9):O277-O83.

47. Haggitt RC, Glotzbach RE, Soffer EE, Wruble LD. Prognostic factors in colorectal carcinomas arising in adenomas: implications for lesions removed by endoscopic polypectomy. Gastroenterology. 1985;89(2):328-36.

48. Kikuchi R, Takano M, Takagi K, Fujimoto N, Nozaki R, Fujiyoshi T, et al. Management of early invasive colorectal cancer. Risk of recurrence and clinical guidelines. Dis Colon Rectum. 1995;38(12):1286-95.

49. Johnstone MS, Stoops R, Lynch G, Hay J, Jawny J, Sloan W, et al. Risk stratification for the detection of metachronous polyps after bowel screening polypectomy: clinical outcomes from the Integrated Technologies for Improved Polyp Surveillance (INCISE) study cohort. BJS Open. 2023;7(3).

50. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more "personalized" approach to cancer staging. CA Cancer J Clin. 2017;67(2):93-9.

51. Colorectal Cancer Stages: American Cancer Society; [updated January 29, 2024. Available from: https://www.cancer.org/cancer/types/colon-rectal-cancer/detection-diagnosis-staging/staged.html.

52. Dukes CE. The classification of cancer of the rectum. The Journal of Pathology and Bacteriology. 2005;35(3):323-32.

53. Tarazi M, Guest K, Cook AJ, Balasubramaniam D, Bailey CMH. Two and five year survival for colorectal cancer after resection with curative intent: A retrospective cohort study. Int J Surg. 2018;55:152-5.

54. Morris EJ, Maughan NJ, Forman D, Quirke P. Who to treat with adjuvant therapy in Dukes B/stage II colorectal cancer? The need for high quality pathology. Gut. 2007;56(10):1419-25.

55. Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz HJ. Understanding the role of primary tumour localisation in colorectal cancer treatment and outcomes. Eur J Cancer. 2017;84:69-80.

56. Paschke S, Jafarov S, Staib L, Kreuser ED, Maulbecker-Armstrong C, Roitman M, et al. Are Colon and Rectal Cancer Two Different Tumor Entities? A Proposal to Abandon the Term Colorectal Cancer. Int J Mol Sci. 2018;19(9).

57. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21(11):1350-6.

58. Petersen VC, Baxter KJ, Love SB, Shepherd NA. Identification of objective pathological prognostic determinants and models of prognosis in Dukes' B colon cancer. Gut. 2002;51(1):65-9.

59. Liebig C, Ayala G, Wilks J, Verstovsek G, Liu H, Agarwal N, et al. Perineural invasion is an independent predictor of outcome in colorectal cancer. J Clin Oncol. 2009;27(31):5131-7.

60. Lugli A, Kirsch R, Ajioka Y, Bosman F, Cathomas G, Dawson H, et al. Recommendations for reporting tumor budding in colorectal cancer based on the International Tumor Budding Consensus Conference (ITBCC) 2016. Mod Pathol. 2017;30(9):1299-311.

61. Hatthakarnkul P, Pennel K, Alexander P, van Wyk H, Roseweir A, Inthagard J, et al. Histopathological tumour microenvironment score independently predicts outcome in primary operable colorectal cancer. J Pathol Clin Res. 2024;10(3):e12374.

62. Koelzer VH, Zlobec I, Lugli A. Tumor budding in colorectal cancer--ready for diagnostic practice? Hum Pathol. 2016;47(1):4-19.

63. Klintrup K, Makinen JM, Kauppila S, Vare PO, Melkko J, Tuominen H, et al. Inflammation and prognosis in colorectal cancer. Eur J Cancer. 2005;41(17):2645-54.

64. Vayrynen JP, Tuomisto A, Klintrup K, Makela J, Karttunen TJ, Makinen MJ. Detailed analysis of inflammatory cell infiltration in colorectal cancer. Br J Cancer. 2013;109(7):1839-47.

65. Park JH, Richards CH, McMillan DC, Horgan PG, Roxburgh CSD. The relationship between tumour stroma percentage, the tumour microenvironment and survival in patients with primary operable colorectal cancer. Ann Oncol. 2014;25(3):644-51.

66. Park JH, McMillan DC, Powell AG, Richards CH, Horgan PG, Edwards J, et al. Evaluation of a tumor microenvironment-based prognostic score in primary operable colorectal cancer. Clin Cancer Res. 2015;21(4):882-8.

67. Alexander PG, Roseweir AK, Pennel KAF, van Wyk HC, Powell A, McMillan DC, et al. The Glasgow Microenvironment Score associates with prognosis and adjuvant chemotherapy response in colorectal cancer. Br J Cancer. 2021;124(4):786-96.

68. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. Gastroenterology. 2010;138(6):2059-72.

69. Tariq K, Ghias K. Colorectal cancer carcinogenesis: a review of mechanisms. Cancer Biol Med. 2016;13(1):120-35.

70. Jen J, Kim H, Piantadosi S, Liu ZF, Levitt RC, Sistonen P, et al. Allelic loss of chromosome 18q and prognosis in colorectal cancer. N Engl J Med. 1994;331(4):213-21.

71. Evrard C, Tachon G, Randrian V, Karayan-Tapon L, Tougeron D. Microsatellite Instability: Diagnosis, Heterogeneity, Discordance, and Clinical Impact in Colorectal Cancer. Cancers (Basel). 2019;11(10).

72. Nazemalhosseini Mojarad E, Kuppen PJ, Aghdaei HA, Zali MR. The CpG island methylator phenotype (CIMP) in colorectal cancer. Gastroenterol Hepatol Bed Bench. 2013;6(3):120-8.

73. Barault L, Charon-Barra C, Jooste V, de la Vega MF, Martin L, Roignot P, et al. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. Cancer Res. 2008;68(20):8541-6.

74. Ogino S, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS. CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. J Mol Diagn. 2006;8(5):582-8.

75. Ogino S, Cantor M, Kawasaki T, Brahmandam M, Kirkner GJ, Weisenberger DJ, et al. CpG island methylator phenotype (CIMP) of colorectal cancer is best characterised by quantitative DNA methylation analysis and prospective cohort studies. Gut. 2006;55(7):1000-6.

76. Ogino S, Meyerhardt JA, Kawasaki T, Clark JW, Ryan DP, Kulke MH, et al. CpG island methylation, response to combination chemotherapy, and patient survival in advanced microsatellite stable colorectal carcinoma. Virchows Arch. 2007;450(5):529-37.

77. Ogino S, Kawasaki T, Kirkner GJ, Kraft P, Loda M, Fuchs CS. Evaluation of markers for CpG island methylator phenotype (CIMP) in colorectal cancer by a large population-based sample. J Mol Diagn. 2007;9(3):305-14.

78. Lee S, Cho NY, Yoo EJ, Kim JH, Kang GH. CpG island methylator phenotype in colorectal cancers: comparison of the new and classic CpG island methylator phenotype marker panels. Arch Pathol Lab Med. 2008;132(10):1657-65.

79. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012;487(7407):330-7.

80. Dunne PD, Arends MJ. Molecular pathological classification of colorectal cancer-an update. Virchows Arch. 2024;484(2):273-85.

81. Sirinukunwattana K, Domingo E, Richman SD, Redmond KL, Blake A, Verrill C, et al. Image-based consensus molecular subtype (imCMS) classification of colorectal cancer using deep learning. Gut. 2021;70(3):544-54.

82. Joanito I, Wirapati P, Zhao N, Nawaz Z, Yeo G, Lee F, et al. Single-cell and bulk transcriptome sequencing identifies two epithelial tumor cell states and

refines the consensus molecular classification of colorectal cancer. Nat Genet. 2022;54(7):963-75.

83. Isella C, Brundu F, Bellomo SE, Galimi F, Zanella E, Porporato R, et al. Selective analysis of cancer-cell intrinsic transcriptional traits defines novel clinically relevant subtypes of colorectal cancer. Nat Commun. 2017;8:15107.

84. Chen B, Scurrah CR, McKinley ET, Simmons AJ, Ramirez-Solano MA, Zhu X, et al. Differential pre-malignant programs and microenvironment chart distinct paths to malignancy in human colorectal polyps. Cell. 2021;184(26):6262-80 e26.

85. Becker WR, Nevins SA, Chen DC, Chiu R, Horning AM, Guha TK, et al. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. Nat Genet. 2022;54(7):985-95.

86. Fumagalli A, Oost KC, Kester L, Morgner J, Bornes L, Bruens L, et al. Plasticity of Lgr5-Negative Cancer Cells Drives Metastasis in Colorectal Cancer. Cell Stem Cell. 2020;26(4):569-78 e7.

87. Vasquez EG, Nasreddin N, Valbuena GN, Mulholland EJ, Belnoue-Davis HL, Eggington HR, et al. Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. Cell Stem Cell. 2022;29(8):1213-28 e8.

88. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1(6):417-25.

89. Malla SB, Byrne RM, Lafarge MW, Corry SM, Fisher NC, Tsantoulis PK, et al. Pathway level subtyping identifies a slow-cycling biological phenotype associated with poor clinical outcomes in colorectal cancer. Nat Genet. 2024;56(3):458-72.

90. Carethers JM, Jung BH. Genetics and Genetic Biomarkers in Sporadic Colorectal Cancer. Gastroenterology. 2015;149(5):1177-90 e3.

91. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61(5):759-67.

92. Fleming M, Ravula S, Tatishchev SF, Wang HL. Colorectal carcinoma: Pathologic aspects. J Gastrointest Oncol. 2012;3(3):153-73.

93. Zhao H, Ming T, Tang S, Ren S, Yang H, Liu M, et al. Wnt signaling in colorectal cancer: pathogenic role and therapeutic target. Mol Cancer. 2022;21(1):144.

94. Fang JY, Richardson BC. The MAPK signalling pathways and colorectal cancer. Lancet Oncol. 2005;6(5):322-7.

95. Leiphrakpam PD, Are C. PI3K/Akt/mTOR Signaling Pathway as a Target for Colorectal Cancer Treatment. Int J Mol Sci. 2024;25(6).

96. Qiao L, Wong BC. Role of Notch signaling in colorectal cancer. Carcinogenesis. 2009;30(12):1979-86.

97. Geyer N, Gerling M. Hedgehog Signaling in Colorectal Cancer: All in the Stroma? Int J Mol Sci. 2021;22(3).

98. Mohammadpour S, Torshizi Esfahani A, Sarpash S, Vakili F, Zafarjafarzadeh N, Mashaollahi A, et al. Hippo Signaling Pathway in Colorectal Cancer: Modulation by Various Signals and Therapeutic Potential. Anal Cell Pathol (Amst). 2024;2024:5767535.

99. Luo C, Zhang H. The Role of Proinflammatory Pathways in the Pathogenesis of Colitis-Associated Colorectal Cancer. Mediators Inflamm. 2017;2017:5126048.

100. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128(4):683-92.

101. Spit M, Koo BK, Maurice MM. Tales from the crypt: intestinal niche signals in tissue renewal, plasticity and cancer. Open Biol. 2018;8(9).

102. Liu J, Xiao Q, Xiao J, Niu C, Li Y, Zhang X, et al. Wnt/beta-catenin signalling: function, biological mechanisms, and therapeutic opportunities. Signal Transduct Target Ther. 2022;7(1):3.

103. Zhou B, Lin W, Long Y, Yang Y, Zhang H, Wu K, et al. Notch signaling pathway: architecture, disease, and therapeutics. Signal Transduct Target Ther. 2022;7(1):95.

104. Abud HE, Chan WH, Jarde T. Source and Impact of the EGF Family of Ligands on Intestinal Stem Cells. Front Cell Dev Biol. 2021;9:685665.

105. Tang X, Liu H, Yang S, Li Z, Zhong J, Fang R. Epidermal Growth Factor and Intestinal Barrier Function. Mediators Inflamm. 2016;2016:1927348.

106. Wee P, Wang Z. Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. Cancers (Basel). 2017;9(5).

107. Perez White BE, Getsios S. Eph receptor and ephrin function in breast, gut, and skin epithelia. Cell Adh Migr. 2014;8(4):327-38.

108. Chen SM, Guo BJ, Feng AQ, Wang XL, Zhang SL, Miao CY. Pathways regulating intestinal stem cells and potential therapeutic targets for radiation enteropathy. Mol Biomed. 2024;5(1):46.

109. Hong AW, Meng Z, Guan KL. The Hippo pathway in intestinal regeneration and disease. Nat Rev Gastroenterol Hepatol. 2016;13(6):324-37.

110. Deng F, Wu Z, Zou F, Wang S, Wang X. The Hippo-YAP/TAZ Signaling Pathway in Intestinal Self-Renewal and Regeneration After Injury. Front Cell Dev Biol. 2022;10:894737.

111. Zhang W, Liu HT. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. Cell Res. 2002;12(1):9-18.

112. Osaki LH, Gama P. MAPKs and signal transduction in the control of gastrointestinal epithelial cell proliferation and differentiation. Int J Mol Sci. 2013;14(5):10143-61.

113. Danielsen SA, Eide PW, Nesbakken A, Guren T, Leithe E, Lothe RA. Portrait of the PI3K/AKT pathway in colorectal cancer. Biochim Biophys Acta. 2015;1855(1):104-21.

114. He Y, Sun MM, Zhang GG, Yang J, Chen KS, Xu WW, et al. Targeting PI3K/Akt signal transduction for cancer therapy. Signal Transduct Target Ther. 2021;6(1):425.

115. Vanhaesebroeck B, Stephens L, Hawkins P. PI3K signalling: the path to discovery and understanding. Nat Rev Mol Cell Biol. 2012;13(3):195-203.

116. Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. Signal Transduct Target Ther. 2021;6(1):402.

117. Guo Q, Jin Y, Chen X, Ye X, Shen X, Lin M, et al. NF-kappaB in biology and targeted therapy: new insights and translational implications. Signal Transduct Target Ther. 2024;9(1):53.

118. Bertrand FE, Angus CW, Partis WJ, Sigounas G. Developmental pathways in colon cancer: crosstalk between WNT, BMP, Hedgehog and Notch. Cell Cycle. 2012;11(23):4344-51.

119. Rodilla V, Villanueva A, Obrador-Hevia A, Robert-Moreno A, Fernandez-Majada V, Grilli A, et al. Jagged1 is the pathological link between Wnt and Notch pathways in colorectal cancer. Proc Natl Acad Sci U S A. 2009;106(15):6315-20.

120. Koveitypour Z, Panahi F, Vakilian M, Peymani M, Seyed Forootan F, Nasr Esfahani MH, et al. Signaling pathways involved in colorectal cancer progression. Cell Biosci. 2019;9:97.

121. Jeong WJ, Ro EJ, Choi KY. Interaction between Wnt/beta-catenin and RAS-ERK pathways and an anti-cancer strategy via degradations of beta-catenin and RAS by targeting the Wnt/beta-catenin pathway. NPJ Precis Oncol. 2018;2(1):5.

122. Song Y, Chen M, Wei Y, Ma X, Shi H. Signaling pathways in colorectal cancer implications for the target therapies. Mol Biomed. 2024;5(1):21.

123. Bahar ME, Kim HJ, Kim DR. Targeting the RAS/RAF/MAPK pathway for cancer therapy: from mechanism to clinical studies. Signal Transduct Target Ther. 2023;8(1):455.

124. Lampropoulos P, Zizi-Sermpetzoglou A, Rizos S, Kostakis A, Nikiteas N, Papavassiliou AG. TGF-beta signalling in colon carcinogenesis. Cancer Lett. 2012;314(1):1-7.

125. Guo X, Wang XF. Signaling cross-talk between TGF-beta/BMP and other pathways. Cell Res. 2009;19(1):71-88.

126. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420-8.

127. Gonzalez DM, Medici D. Signaling mechanisms of the epithelial-mesenchymal transition. Sci Signal. 2014;7(344):re8.

128. Venook AP, Niedzwiecki D, Innocenti F, Fruth B, Greene C, O'Neil BH, et al. Impact of primary (1°) tumor location on overall survival (OS) and progression-free survival (PFS) in patients (pts) with metastatic colorectal cancer (mCRC): Analysis of CALGB/SWOG 80405 (Alliance). Journal of Clinical Oncology. 2016;34(15_suppl):3504-.

129. Riihimaki M, Hemminki A, Sundquist J, Hemminki K. Patterns of metastasis in colon and rectal cancer. Sci Rep. 2016;6:29765.

130. Qiu M, Hu J, Yang D, Cosgrove DP, Xu R. Pattern of distant metastases in colorectal cancer: a SEER based study. Oncotarget. 2015;6(36):38658-66.

131. Shin AE, Giancotti FG, Rustgi AK. Metastatic colorectal cancer: mechanisms and emerging therapeutics. Trends Pharmacol Sci. 2023;44(4):222-36.

132. Fares J, Fares MY, Khachfe HH, Salhab HA, Fares Y. Molecular principles of metastasis: a hallmark of cancer revisited. Signal Transduct Target Ther. 2020;5(1):28.

133. Massague J, Obenauf AC. Metastatic colonization by circulating tumour cells. Nature. 2016;529(7586):298-306.

134. Martin J, Petrillo A, Smyth EC, Shaida N, Khwaja S, Cheow HK, et al. Colorectal liver metastases: Current management and future perspectives. World J Clin Oncol. 2020;11(10):761-808.

135. Zaharia C, Veen T, Lea D, Kanani A, Alexeeva M, Soreide K. Histopathological Growth Pattern in Colorectal Liver Metastasis and The Tumor Immune Microenvironment. Cancers (Basel). 2022;15(1).

136. Vekemans K, Braet F. Structural and functional aspects of the liver and liver sinusoidal cells in relation to colon carcinoma metastasis. World J Gastroenterol. 2005;11(33):5095-102.

137. Zhou H, Liu Z, Wang Y, Wen X, Amador EH, Yuan L, et al. Colorectal liver metastasis: molecular mechanism and interventional therapy. Signal Transduct Target Ther. 2022;7(1):70.

138. Mutsaers SE. Mesothelial cells: their structure, function and role in serosal repair. Respirology. 2002;7(3):171-91.

139. Demuytere J, Ceelen W, Van Dorpe J, Hoorens A. The role of the peritoneal microenvironment in the pathogenesis of colorectal peritoneal carcinomatosis. Exp Mol Pathol. 2020;115:104442.

140. Franko J, Shi Q, Meyers JP, Maughan TS, Adams RA, Seymour MT, et al. Prognosis of patients with peritoneal metastatic colorectal cancer given systemic therapy: an analysis of individual patient data from prospective randomised trials from the Analysis and Research in Cancers of the Digestive System (ARCAD) database. Lancet Oncol. 2016;17(12):1709-19.

141. Baaten I, West NP, Quyn AJ, Seymour MT, Seligmann JF. Colorectal cancer peritoneal metastases: Biology, treatment and next steps. Eur J Surg Oncol. 2020;46(4 Pt A):675-83.

142. Lemoine L, Sugarbaker P, Van der Speeten K. Pathophysiology of colorectal peritoneal carcinomatosis: Role of the peritoneum. World J Gastroenterol. 2016;22(34):7692-707.

143. Kranenburg O, van der Speeten K, de Hingh I. Peritoneal Metastases From Colorectal Cancer: Defining and Addressing the Challenges. Front Oncol. 2021;11:650098.

144. Knudsen L, Ochs M. The micromechanics of lung alveoli: structure and function of surfactant and tissue components. Histochem Cell Biol. 2018;150(6):661-76.

145. Meng X, Cui G, Peng G. Lung development and regeneration: newly defined cell types and progenitor status. Cell Regen. 2023;12(1):5.

146. Chiappetta M, Salvatore L, Congedo MT, Bensi M, De Luca V, Petracca Ciavarella L, et al. Management of single pulmonary metastases from colorectal cancer: State of the art. World J Gastrointest Oncol. 2022;14(4):820-32.

147. Mitry E, Guiu B, Cosconea S, Jooste V, Faivre J, Bouvier AM. Epidemiology, management and prognosis of colorectal cancer with lung metastases: a 30-year population-based study. Gut. 2010;59(10):1383-8.

148. Shishido Y, Ishii M, Maeda T, Kokado Y, Masuya D, Kusama T, et al. Survival outcomes of lung metastases from colorectal cancer treated with pulmonary metastasectomy or modern systemic chemotherapy: a single institution experience. J Cardiothorac Surg. 2023;18(1):327.

149. Beckers P, Berzenji L, Yogeswaran SK, Lauwers P, Bilotta G, Shkarpa N, et al. Pulmonary metastasectomy in colorectal carcinoma. J Thorac Dis. 2021;13(4):2628-35.

150. Busch C, Wheadon H. Bone marrow niche crosses paths with BMPs: a road to protection and persistence in CML. Biochem Soc Trans. 2019;47(5):1307-25.

151. Al-Drees MA, Yeo JH, Boumelhem BB, Antas VI, Brigden KW, Colonne CK, et al. Making Blood: The Haematopoietic Niche throughout Ontogeny. Stem Cells Int. 2015;2015:571893.

152. Frantz C, Stewart KM, Weaver VM. The extracellular matrix at a glance. J Cell Sci. 2010;123(Pt 24):4195-200.

153. Massague J. TGF-beta signal transduction. Annu Rev Biochem. 1998;67:753-91.

154. Benson AB, Venook AP, Adam M, Chang G, Chen YJ, Ciombor KK, et al. Colon Cancer, Version 3.2024, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2024;22(2 D).

155. Chang GJ, Kaiser AM, Mills S, Rafferty JF, Buie WD. Practice Parameters for the Management of Colon Cancer. Diseases of the Colon & Rectum. 2012;55(8):831-43.

156. George J, Rockall T. Surgery for colorectal cancer. Surgery (Oxford). 2023;41(1):35-40.

157. Monson JRT, Weiser MR, Buie WD, Chang GJ, Rafferty JF, Buie WD, et al. Practice Parameters for the Management of Rectal Cancer (Revised). Diseases of the Colon & Rectum. 2013;56(5):535-50.

158. Kumar A, Gautam V, Sandhu A, Rawat K, Sharma A, Saha L. Current and emerging therapeutic approaches for colorectal cancer: A comprehensive review. World J Gastrointest Surg. 2023;15(4):495-519.

159. Rentsch M, Schiergens T, Khandoga A, Werner J. Surgery for Colorectal Cancer - Trends, Developments, and Future Perspectives. Visc Med. 2016;32(3):184-91.

160. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. Signal Transduct Target Ther. 2020;5(1):22.

161. Shimu AS, Wei HX, Li Q, Zheng X, Li B. The new progress in cancer immunotherapy. Clin Exp Med. 2023;23(3):553-67.

162. Massague J, Sheppard D. TGF-beta signaling in health and disease. Cell. 2023;186(19):4007-37.

163. Deng Z, Fan T, Xiao C, Tian H, Zheng Y, Li C, et al. TGF-beta signaling in health, disease, and therapeutics. Signal Transduct Target Ther. 2024;9(1):61.

164. Takayama T, Miyanishi K, Hayashi T, Sato Y, Niitsu Y. Colorectal cancer: genetics of development and metastasis. J Gastroenterol. 2006;41(3):185-92.

165. Kozak MM, von Eyben R, Pai J, Vossler SR, Limaye M, Jayachandran P, et al. Smad4 inactivation predicts for worse prognosis and response to fluorouracil-based treatment in colorectal cancer. J Clin Pathol. 2015;68(5):341-5.

166. Xie W, Rimm DL, Lin Y, Shih WJ, Reiss M. Loss of Smad signaling in human colorectal cancer is associated with advanced disease and poor prognosis. Cancer J. 2003;9(4):302-12.

167. Fan Y, Guo L, Zheng H, Ji C, Wang W, Sun H. BMP-9 is a novel marker for colorectal tumorigenesis undergoing the normal mucosa-adenoma-adenocarcinoma sequence and is associated with colorectal cancer prognosis. Oncol Lett. 2020;19(1):271-82.

168. Yokoyama Y, Watanabe T, Tamura Y, Hashizume Y, Miyazono K, Ehata S. Autocrine BMP-4 Signaling Is a Therapeutic Target in Colorectal Cancer. Cancer Res. 2017;77(15):4026-38.

169. Voorneveld PW, Kodach LL, Jacobs RJ, Liv N, Zonnevylle AC, Hoogenboom JP, et al. Loss of SMAD4 alters BMP signaling to promote colorectal cancer cell metastasis via activation of Rho and ROCK. Gastroenterology. 2014;147(1):196-208 e13.

170. Ouahoud S, Voorneveld PW, van der Burg LRA, de Jonge-Muller ESM, Schoonderwoerd MJA, Paauwe M, et al. Bidirectional tumor/stroma crosstalk promotes metastasis in mesenchymal colorectal cancer. Oncogene. 2020;39(12):2453-66.

171. Ehata S, Miyazono K. Bone Morphogenetic Protein Signaling in Cancer; Some Topics in the Recent 10 Years. Front Cell Dev Biol. 2022;10:883523.

172. Liu S, Ren J, Ten Dijke P. Targeting TGFbeta signal transduction for cancer therapy. Signal Transduct Target Ther. 2021;6(1):8.

173. Akhurst RJ, Hata A. Targeting the TGFbeta signalling pathway in disease. Nat Rev Drug Discov. 2012;11(10):790-811.

174. Huang CY, Chung CL, Hu TH, Chen JJ, Liu PF, Chen CL. Recent progress in TGF-beta inhibitors for cancer therapy. Biomed Pharmacother. 2021;134:111046.

175. Lowery JW, Brookshire B, Rosen V. A Survey of Strategies to Modulate the Bone Morphogenetic Protein Signaling Pathway: Current and Future Perspectives. Stem Cells Int. 2016;2016:7290686.

176. Mohedas AH, Xing X, Armstrong KA, Bullock AN, Cuny GD, Yu PB. Development of an ALK2-biased BMP type I receptor kinase inhibitor. ACS Chem Biol. 2013;8(6):1291-302.

177. Yu PB, Hong CC, Sachidanandan C, Babitt JL, Deng DY, Hoyng SA, et al. Dorsomorphin inhibits BMP signals required for embryogenesis and iron metabolism. Nat Chem Biol. 2008;4(1):33-41.

178. Sanvitale CE, Kerr G, Chaikuad A, Ramel MC, Mohedas AH, Reichert S, et al. A new class of small molecule inhibitor of BMP signaling. PLoS One. 2013;8(4):e62721.

179. Liu L, Yan Q, Chen Z, Wei X, Li L, Tang D, et al. Overview of research progress and application of experimental models of colorectal cancer. Front Pharmacol. 2023;14:1193213.

180. Vitale S, Calapa F, Colonna F, Luongo F, Biffoni M, De Maria R, et al. Advancements in 3D *In Vitro* Models for Colorectal Cancer. Adv Sci (Weinh). 2024;11(32):e2405084.

181. Preetam S, Nahak BK, Patra S, Toncu DC, Park S, Syväjärvi M, et al. Emergence of microfluidics for next generation biomedical devices. Biosensors and Bioelectronics: X. 2022;10.

182. Whitesides GM. The origins and the future of microfluidics. Nature. 2006;442(7101):368-73.

183. Neto I, Rocha J, Gaspar MM, Reis CP. Experimental Murine Models for Colorectal Cancer Research. Cancers (Basel). 2023;15(9).

184. Mittal VK, Bhullar JS, Jayant K. Animal models of human colorectal cancer: Current status, uses and limitations. World J Gastroenterol. 2015;21(41):11854-61.

185. Lubbock ALR, Lopez CF. Programmatic modeling for biological systems. Curr Opin Syst Biol. 2021;27.

186. Suleman M, Ali S, Rashid F, Khan FI. Editorial: Experimental and computational methods in the development of diagnostics and therapeutics for colon cancer. Front Mol Biosci. 2025;12:1568721.

187. Vallejo Morales E, Suarez Guerrero G, Hoyos Palacio LM. Computational Simulation of Colorectal Cancer Biomarker Particle Mobility in a 3D Model. Molecules. 2023;28(2).

188. Sommariva S, Caviglia G, Ravera S, Frassoni F, Benvenuto F, Tortolina L, et al. Computational quantification of global effects induced by mutations and drugs in signaling networks of colorectal cancer cells. Sci Rep. 2021;11(1):19602.

189. Paterson C, Clevers H, Bozic I. Mathematical model of colorectal cancer initiation. Proc Natl Acad Sci U S A. 2020;117(34):20681-8.

190. Ahmad K, Shaikh S, Khan FI, Khan ME. Editorial: Computational drug discovery of medicinal compounds for cancer management, volume II. Front Chem. 2024;12:1446510.

191. Zitter R, Chugh RM, Saha S. Patient Derived Ex-Vivo Cancer Models in Drug Development, Personalized Medicine, and Radiotherapy. Cancers (Basel). 2022;14(12).

192. Hariton E, Locascio JJ. Randomised controlled trials - the gold standard for effectiveness research: Study design: randomised controlled trials. BJOG. 2018;125(13):1716.

193. Inadomi JM, Vijan S, Janz NK, Fagerlin A, Thomas JP, Lin YV, et al. Adherence to colorectal cancer screening: a randomized clinical trial of competing strategies. Arch Intern Med. 2012;172(7):575-82.

194. Advani S, Kopetz S. Ongoing and future directions in the management of metastatic colorectal cancer: Update on clinical trials. J Surg Oncol. 2019;119(5):642-52.

195. Venook AP, Niedzwiecki D, Lenz HJ, Innocenti F, Fruth B, Meyerhardt JA, et al. Effect of First-Line Chemotherapy Combined With Cetuximab or Bevacizumab on Overall Survival in Patients With KRAS Wild-Type Advanced or

Metastatic Colorectal Cancer: A Randomized Clinical Trial. JAMA. 2017;317(23):2392-401.

196. Hoeben A, Joosten EAJ, van den Beuken-van Everdingen MHJ. Personalized Medicine: Recent Progress in Cancer Therapy. Cancers (Basel). 2021;13(2).

197. Riedesser JE, Ebert MP, Betge J. Precision medicine for metastatic colorectal cancer in clinical practice. Ther Adv Med Oncol. 2022;14:17588359211072703.

198. Mansouri D, McMillan DC, Grant Y, Crighton EM, Horgan PG. The impact of age, sex and socioeconomic deprivation on outcomes in a colorectal cancer screening programme. PLoS One. 2013;8(6):e66063.

199. Arafeh R, Shibue T, Dempster JM, Hahn WC, Vazquez F. The present and future of the Cancer Dependency Map. Nat Rev Cancer. 2025;25(1):59-73.

200. Killian T, Gatto L. Exploiting the DepMap cancer dependency data using the depmap R package. F1000Research. 2021;10.

201. Webster JD, Solon M, Gibson-Corley KN. Validating Immunohistochemistry Assay Specificity in Investigative Studies: Considerations for a Weight of Evidence Approach. Vet Pathol. 2021;58(5):829-40.

202. Kirkegaard T, Edwards J, Tovey S, McGlynn LM, Krishna SN, Mukherjee R, et al. Observer variation in immunohistochemical analysis of protein expression, time for a change? Histopathology. 2006;48(7):787-94.

203. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878.

204. Yeakley JM, Shepard PJ, Goyena DE, VanSteenhouse HC, McComb JD, Seligmann BE. A trichostatin A expression signature identified by TempO-Seq targeted whole transcriptome profiling. PLoS One. 2017;12(5):e0178302.

205. Hatthakarnkul P, Ammar A, Pennel KAF, Officer-Jones L, Cusumano S, Quinn JA, et al. Protein expression of S100A2 reveals it association with patient prognosis and immune infiltration profile in colorectal cancer. J Cancer. 2023;14(10):1837-47.

206. Strober W. Trypan Blue Exclusion Test of Cell Viability. Curr Protoc Immunol. 2015;111:A3 B 1-A3 B

207. Adan A, Kiraz Y, Baran Y. Cell Proliferation and Cytotoxicity Assays. Curr Pharm Biotechnol. 2016;17(14):1213-21.

208. Chou TC. Drug combination studies and their synergy quantification using the Chou-Talalay method. Cancer Res. 2010;70(2):440-6.

209. Elwakeel A, Soudan H, Eldoksh A, Shalaby M, Eldemellawy M, Ghareeb D, et al. Implementation of the Chou-Talalay method for studying the *in vitro* pharmacodynamic interactions of binary and ternary drug combinations on MDA-MB-231 triple negative breast cancer cells. Synergy. 2019;8.

210. Franken NA, Rodermond HM, Stap J, Haveman J, van Bree C. Clonogenic assay of cells i*n vitro*. Nat Protoc. 2006;1(5):2315-9.

211. Elegheert J, Behiels E, Bishop B, Scott S, Woolley RE, Griffiths SC, et al. Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins. Nat Protoc. 2018;13(12):2991-3017.

212. Dowling P, Clynes M. Conditioned media from cell lines: a complementary model to clinical specimens for the discovery of disease-specific biomarkers. Proteomics. 2011;11(4):794-804.

213. Vermes I, Haanen C, Steffens-Nakken H, Reutelingsperger C. A novel assay for apoptosis. Flow cytometric detection of phosphatidylserine expression on early apoptotic cells using fluorescein labelled Annexin V. J Immunol Methods. 1995;184(1):39-51.

214. Wallberg F, Tenev T, Meier P. Analysis of Apoptosis and Necroptosis by Fluorescence-Activated Cell Sorting. Cold Spring Harb Protoc. 2016;2016(4):pdb prot087387.

215. Kim KH, Sederstrom JM. Assaying Cell Cycle Status Using Flow Cytometry. Curr Protoc Mol Biol. 2015;111:28 6 1- 6 11.

216. Lemieszek MB, Findlay SD, Siegers GM. CellTrace Violet Flow Cytometric Assay to Assess Cell Proliferation. Methods Mol Biol. 2022;2508:101-14.

217. Sablina AA, Chumakov PM, Levine AJ, Kopnin BP. p53 activation in response to microtubule disruption is mediated by integrin-Erk signaling. Oncogene. 2001;20(8):899-909.

218. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82.

219. Guerrieri AN, Montesi M, Sprio S, Laranga R, Mercatali L, Tampieri A, et al. Innovative Options for Bone Metastasis Treatment: An Extensive Analysis on Biomaterials-Based Strategies for Orthopedic Surgeons. Front Bioeng Biotechnol. 2020;8:589964.

220. Cai YD, Zhang Q, Zhang YH, Chen L, Huang T. Identification of Genes Associated with Breast Cancer Metastasis to Bone on a Protein-Protein Interaction Network with a Shortest Path Algorithm. J Proteome Res. 2017;16(2):1027-38.

221. Howe JR, Roth S, Ringold JC, Summers RW, Jarvinen HJ, Sistonen P, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. Science. 1998;280(5366):1086-8.

222. Aasebo E, Brenner AK, Hernandez-Valladares M, Birkeland E, Berven FS, Selheim F, et al. Proteomic Comparison of Bone Marrow Derived Osteoblasts and Mesenchymal Stem Cells. Int J Mol Sci. 2021;22(11).

223. Hennrich ML, Romanov N, Horn P, Jaeger S, Eckstein V, Steeples V, et al. Cell-specific proteome analyses of human bone marrow reveal molecular features of age-dependent functional decline. Nat Commun. 2018;9(1):4004.

224. Lovero D, D'Oronzo S, Palmirotta R, Cafforio P, Brown J, Wood S, et al. Correlation between targeted RNAseq signature of breast cancer CTCs and onset of bone-only metastases. Br J Cancer. 2022;126(3):419-29.

225. Wang S, Ao L, Lin H, Wei H, Wu Z, Lu S, et al. Single-cell transcriptomic analysis of the senescent microenvironment in bone metastasis. Cell Prolif. 2025;58(1):e13743.

226. Poudel BH, Koks S. The whole transcriptome analysis using FFPE and fresh tissue samples identifies the molecular fingerprint of osteosarcoma. Exp Biol Med (Maywood). 2024;249:10161.

227. Liu F, Ding Y, Xu Z, Hao X, Pan T, Miles G, et al. Single-cell profiling of bone metastasis ecosystems from multiple cancer types reveals convergent and divergent mechanisms of bone colonization. Cell Genom. 2025:100888.

228. Yip RKH, Rimes JS, Capaldo BD, Vaillant F, Mouchemore KA, Pal B, et al. Mammary tumour cells remodel the bone marrow vascular microenvironment to support metastasis. Nat Commun. 2021;12(1):6920.

229. Zhou Y, Tao L, Qiu J, Xu J, Yang X, Zhang Y, et al. Tumor biomarkers for diagnosis, prognosis and targeted therapy. Signal Transduct Target Ther. 2024;9(1):132.

230. de Back TR, van Hooff SR, Sommeijer DW, Vermeulen L. Transcriptomic subtyping of gastrointestinal malignancies. Trends Cancer. 2024;10(9):842-56.

231. Cao K, Plazzer JP, Macrae F. SMAD4 variants and its genotype-phenotype correlations to juvenile polyposis syndrome. Hered Cancer Clin Pract. 2023;21(1):27.

232. Langeveld D, van Hattem WA, de Leng WW, Morsink FH, Ten Kate FJ, Giardiello FM, et al. SMAD4 immunohistochemistry reflects genetic status in juvenile polyposis syndrome. Clin Cancer Res. 2010;16(16):4126-34.

233. Howe JR, Sayed MG, Ahmed AF, Ringold J, Larsen-Haidle J, Merg A, et al. The prevalence of MADH4 and BMPR1A mutations in juvenile polyposis and absence of BMPR2, BMPR1B, and ACVR1 mutations. J Med Genet. 2004;41(7):484-91.

234. Vautier S, Mauillon J, Parodi N, Bou J, Kasper E, Manase S, et al. SMAD4 mosaicism in juvenile polyposis: Essential contribution of somatic analysis in diagnosis. Am J Med Genet A. 2024;194(9):e63648.

235. Pearson M, McGowan R, Greene P, Lam W, Miedzybrodzka Z, Berg J. Outcomes of patients with Juvenile Polyposis-Hereditary Haemorrhagic Telangiectasia caused by pathogenic SMAD4 variants in a pan-Scotland cohort. Eur J Hum Genet. 2024;32(6):731-5.

236. Wilentz RE, Su GH, Dai JL, Sparks AB, Argani P, Sohn TA, et al. Immunohistochemical labeling for dpc4 mirrors genetic status in pancreatic adenocarcinomas : a new marker of DPC4 inactivation. Am J Pathol. 2000;156(1):37-43.

237. Wasserman I, Lee LH, Ogino S, Marco MR, Wu C, Chen X, et al. SMAD4 Loss in Colorectal Cancer Patients Correlates with Recurrence, Loss of Immune Infiltrate, and Chemoresistance. Clin Cancer Res. 2019;25(6):1948-56.

238. Ali S, Cohen C, Little JV, Sequeira JH, Mosunjac MB, Siddiqui MT. The utility of SMAD4 as a diagnostic immunohistochemical marker for pancreatic adenocarcinoma, and its expression in other solid tumors. Diagn Cytopathol. 2007;35(10):644-8.

239. Ohki D, Yamamichi N, Sakaguchi Y, Takahashi Y, Kageyama-Yahara N, Yamamichi M, et al. Transcriptome of sessile serrated adenoma/polyps is associated with MSI-high colorectal cancer and decreased expression of CDX2. Cancer Med. 2022;11(24):5066-78.

240. Sabates-Bellver J, Van der Flier LG, de Palo M, Cattaneo E, Maake C, Rehrauer H, et al. Transcriptome profile of human colorectal adenomas. Mol Cancer Res. 2007;5(12):1263-75.

241. Druliner BR, Wang P, Bae T, Baheti S, Slettedahl S, Mahoney D, et al. Molecular characterization of colorectal adenomas with and without malignancy reveals distinguishing genome, transcriptome and methylome alterations. Sci Rep. 2018;8(1):3161.

242. Pan P, Li J, Wang B, Tan X, Yin H, Han Y, et al. Molecular characterization of colorectal adenoma and colorectal cancer via integrated genomic transcriptomic analysis. Front Oncol. 2023;13:1067849.

243. Fang T, Liang T, Wang Y, Wu H, Liu S, Xie L, et al. Prognostic role and clinicopathological features of SMAD4 gene mutation in colorectal cancer: a systematic review and meta-analysis. BMC Gastroenterol. 2021;21(1):297.

244. Mizuno T, Cloyd JM, Vicente D, Omichi K, Chun YS, Kopetz SE, et al. SMAD4 gene mutation predicts poor prognosis in patients undergoing resection for colorectal liver metastases. Eur J Surg Oncol. 2018;44(5):684-92.

245. de Visser KE, Joyce JA. The evolving tumor microenvironment: From cancer initiation to metastatic outgrowth. Cancer Cell. 2023;41(3):374-403.

246. Mehrvarz Sarshekeh A, Advani S, Overman MJ, Manyam G, Kee BK, Fogelman DR, et al. Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. PLoS One. 2017;12(3):e0173345.

247. Yoo SY, Lee JA, Shin Y, Cho NY, Bae JM, Kang GH. Clinicopathological Characterization and Prognostic Implication of SMAD4 Expression in Colorectal Carcinoma. J Pathol Transl Med. 2019;53(5):289-97.

248. Isaksson-Mettavainio M, Palmqvist R, Forssell J, Stenling R, Oberg A. SMAD4/DPC4 expression and prognosis in human colorectal cancer. Anticancer Res. 2006;26(1B):507-10.

249. Patel M, McSorley ST, Park JH, Roxburgh CSD, Edwards J, Horgan PG, et al. The relationship between right-sided tumour location, tumour microenvironment, systemic inflammation, adjuvant therapy and survival in patients undergoing surgery for colon and rectal cancer. Br J Cancer. 2018;118(5):705-12.

250. Saunders MP, Iype R, Kelly C, Crosby J, Kerr R, Harkin A, et al. SCOT: Tumor Sidedness and the Influence of Adjuvant Chemotherapy Duration on Disease Free Survival (DFS). Clin Colorectal Cancer. 2023;22(2):231-7.

251. Mesker WE, Liefers GJ, Junggeburt JM, van Pelt GW, Alberici P, Kuppen PJ, et al. Presence of a high amount of stroma and downregulation of SMAD4 predict for worse survival for stage I-II colon cancer patients. Cell Oncol. 2009;31(3):169-78.

252. Wang D, Chen M, Tao Z, Du J, Tian K, Chen Z, et al. Overexpression of Extracellular Superoxide Dismutase 3 Inhibits Cancer Cell Growth and Migration in Colorectal Cancer. Turk J Gastroenterol. 2024;35(6):465-74.

253. Sanz B, Perez I, Beitia M, Errarte P, Fernandez A, Blanco L, et al. Aminopeptidase N activity predicts 5-year survival in colorectal cancer patients. J Investig Med. 2015;63(5):740-6.

254. Szeglin BC, Wu C, Marco MR, Park HS, Zhang Z, Zhang B, et al. A SMAD4-modulated gene profile predicts disease-free survival in stage II and III colorectal cancer. Cancer Rep (Hoboken). 2022;5(1):e1423.

255. Zhao D, Qu Y, Gao N, Wu T. Integrated multi-omics characterization of SMAD4 mutant colorectal cancer. Discov Oncol. 2024;15(1):386.

256. Demagny H, Araki T, De Robertis EM. The tumor suppressor Smad4/DPC4 is regulated by phosphorylations that integrate FGF, Wnt, and TGF-beta signaling. Cell Rep. 2014;9(2):688-700.

257. Liang M, Melchior F, Feng XH, Lin X. Regulation of Smad4 sumoylation and transforming growth factor-beta signaling by protein inhibitor of activated STAT1. J Biol Chem. 2004;279(22):22857-65.

258. Du X, Pan Z, Li Q, Liu H, Li Q. SMAD4 feedback regulates the canonical TGF-beta signaling pathway to control granulosa cell apoptosis. Cell Death Dis. 2018;9(2):151.

259. Wu XB, Liu Y, Wang GH, Xu X, Cai Y, Wang HY, et al. Mesenchymal stem cells promote colorectal cancer progression through AMPK/mTOR-mediated NF-kappaB activation. Sci Rep. 2016;6:21420.

260. Ferjaoui Z, Lopez-Munoz R, Akbari S, Chandad F, Mantovani D, Rouabhia M, et al. Design of Alginate/Gelatin Hydrogels for Biomedical Applications: Fine-Tuning Osteogenesis in Dental Pulp Stem Cells While Preserving Other Cell Behaviors. Biomedicines. 2024;12(7).

261. Xia RH, Song XM, Wang XJ, Li J, Mao L. The combination of SMAD4 expression and histological grade of dysplasia is a better predictor for the malignant transformation of oral leukoplakia. PLoS One. 2013;8(6):e66794.

262. Sakata J, Yoshida R, Matsuoka Y, Nagata M, Hirosue A, Kawahara K, et al. Predictive value of the combination of SMAD4 expression and lymphocyte infiltration in malignant transformation of oral leukoplakia. Cancer Med. 2017;6(4):730-8.

263. Madabhushi A, Lee G. Image analysis and machine learning in digital pathology: Challenges and opportunities. Med Image Anal. 2016;33:170-5.

264. Harms PW, Frankel TL, Moutafi M, Rao A, Rimm DL, Taube JM, et al. Multiplex Immunohistochemistry and Immunofluorescence: A Practical Update for Pathologists. Mod Pathol. 2023;36(7):100197.

265. Abasalizadeh F, Moghaddam SV, Alizadeh E, Akbari E, Kashani E, Fazljou SMB, et al. Alginate-based hydrogels as drug delivery vehicles in cancer treatment and their applications in wound dressing and 3D bioprinting. J Biol Eng. 2020;14:8.