Characterisation of the effect of *FBXW7* mutations in normal and malignant colonic tissue

Dedrick Kok Hong Chan

Linacre College



A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2023

For Beatrice and Edith

Acknowledgements

There is a famous saying by the Chinese philosopher Lao Tzu which goes, “千里之行，始于足下.” Translated, it suggests that even the longest journeys begin with a first step. This time doing my DPhil has often felt like an impossible journey. I found it difficult to see beyond every failed step along the way, often asking myself if this was the end of the road - would this be as far along my DPhil as I would go? In those early days marked by the solitude and bleakness of the COVID lockdowns, it was easy to imagine this to be a solo journey with nary a soul to seek solace from. This could not be further from the truth.

To my friends May Sin Ke, Nikhil Lal, Aik Seng Ng, and Jack Flower, named in order of when I met you during my DPhil, I am grateful for your lasting friendship. Together, we have shared our difficulties unravelling Science, gossiped and fellowshipped. When I ate dinner in the atrium with two make-shift pipette chopsticks, you bought me a real wooden pair. When you realized I liked skimmed milk with my coffee, you brought a pint each time. We suffered Science together, and that kept me on track.

To my supervisor Professor Simon Buczacki, thank you for taking a chance on me. I hadn’t done much Science when I first met you in Cambridge, but yet, you somehow took me on. Over the last three years, you’ve given me space to mature as a Scientist, and challenged me to tackle the great and difficult problems, rather than settle for the easier ones. I will always remember our scientific discussions, whether in your wife’s car before you left for holiday, in the middle of a park, or when I came to your operating theatre for commiseration because of yet another failed experiment. Wherever my journey in Science takes me, I will always be made in the Buczacki Lab.

To Beansprout, even though you were meant for something greater than this world, you brought warmth and light into my bleak midwinter. In you, I will always have a reason to return into the Oxfordshire countryside. To Edith, you gave me such joy and happiness in this

past six months that any disappointment magically transformed into hope, and despair into empowerment. Your smile gives me strength to try again and again.

And finally, to my wife Beatrice. You endured so much during the past three years, sacrificing so that I could lead an almost monastic existence to carry out my research. I regret I was not always there for you. Yet, you endured miles of walking alone, even when you had to carry loads which grew more cumbersome over time. You were also harassed by the miasma of the shapeless London night. Even when we were together, you suffered every twist and turn that marked my DPhil journey. I hope there were still tender times for I think those memories forged will last our lifetimes. You gave me a reason to work hard and to keep trying, as I wanted to make sure the pound of flesh exacted from you was worth it. I hope I did you proud.

Although my time on this DPhil draws to a close, I wish to acknowledge that every step was made possible only because of these people who have supported and suffered with me till this moment.

AD GLORIAM, NOLI TIMERE

Abstract

Colorectal cancer harbours significant intratumoral heterogeneity. Yet, while intratumoural heterogeneity may be readily observed, events which give rise to heterogeneity, as well as an understanding of the molecular effects arising from interactions between clonal populations in each cancer's unique milieu of heterogeneity, remain elusive. I use *FBXW7* as the starting point in two separate but related themes. In the first theme, colorectal cancer cell lines are used to model the effect of *FBXW7* mutations on neighbouring *FBXW7* wildtype cells. I demonstrate that wildtype cells may acquire phenotypic characteristics present in mutant cells owing from proteins arising from the secretome of mutant cells. In the second theme, I investigate the role of *FBXW7* present in phenotypically normal tissue using gene editing in patient derived human colon organoids. Here, I demonstrate that early *FBXW7* mutations result in a more restricted whole genome chromosome accessibility, abrogating the effects of later-acquired cancer driver mutations such as *APC*. In addition, I demonstrate that an early *FBXW7* mutation can result in fetal reprogramming. In this way, *FBXW7* mutations acquired in phenotypically normal tissue appear to be protective in a human colonic model of early colorectal cancer. Taken together, my experiments bear significant implications on understanding colorectal cancer intratumoural heterogeneity. These results suggest that characterising molecular interactions in intra-cancer clonal populations may lead to a deeper understanding of phenotypic characteristics relevant to that cancer. The results also suggest that early events could forestall the generation of heterogeneity and be potentially protective of cancer development. Here, I use *FBXW7* as an exemplar mutation to generate implications which may be validated using other genes, and in other types of cancer.

Table of Contents

Acknowledgements iii
Abstract v
Table of Contents vi
List of abbreviations x
List of Figures xx
List of Tables xxii

Chapter 1: Introduction 1
Colorectal cancer exerts a heavy burden of disease globally 1
The environmental and genetic basis for disease in CRC 2
Chromosomal instability (CIN) 2
Microsatellite instability (MSI) 5
Serrated neoplasia 6
Driver mutations play an important role in CRC carcinogenesis 8
Driver mutations have also been identified in normal colonic epithelium 10
FBXW7 10
STAG2 13
AXIN2 13
ERBB2 14
ERBB3 14
PIK3CA 15
Introduction to the two themes in this thesis 15
Theme 1: Biallelic *FBXW7* knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells 16
Intratumoral heterogeneity hints at a complex interplay of interactions between driver genes 16
Clonal interactions imbue tumours with unique phenotypes 19
Theme 2: Mutational order and epistasis determine the consequences of *FBXW7* mutations during colorectal cancer evolution 21
Phenotypically normal epithelia accumulate driver mutations 21
Intestine as a model to study stem cell dynamics 22
Molecular and functional characteristics of *FBXW7* 24

Organoids as models for studying early tumour evolution 27
Hypotheses of this thesis 29
Conclusion 29

Chapter 2: Methods for Theme 1 experiments 31
Cell culture 31
Generation of clonal CRISPR-Cas9 gene knockouts 31
Protein lysates and western blot 33
Quantitative reverse transcription-PCR (RT-PCR) 34
Live-cell microscopy 34
Flow cytometry 35
Extreme limiting dilution colony forming assay 35
Metabolic assay 36
Apoptosis assays 36
RNA sequencing and analysis 38
Transwell cell culture 38
Immunofluorescence and imaging 39
Single-cell gel electrophoresis assay 40
Proteomics analysis 41
Metabolomics analysis 42
Transient transfection of plasmid for AKAP8 overexpression 44
Validation of results from datasets 44
Statistical analyses 45

Chapter 3: Methods for Theme 2 experiments 46
Human material for organoid cultures 46
Organoid culture 46
CRISPR/Cas9 gene editing of human organoids 49
Protein lysates and western blot 51
Immunofluorescence and imaging 52
Co-culture of organoids 53
Bulk RNA sequencing and analysis 54
Single-cell RNA sequencing and analysis 55
Assay for transposase-accessible chromatin with sequencing and analysis 56

Validation of results from datasets 58
Statistical analyses 58

Chapter 4: Biallelic *FBXW7* knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells 59
Introduction 59
Results 62
FBXW7$^{-/-}$ cells are at a proliferative disadvantage compared to wildtype cells 62
FBXW7$^{-/-}$ demonstrate upregulation of oxidative phosphorylation 66
FBXW7$^{-/-}$ induce DNA damage in neighbouring wildtype cells 72
AKAP8 from *FBXW7*$^{-/-}$ mediates DNA damage on wildtype cells 79
Effects of *FBXW7* mutation are recapitulated in TCGA-COAD data 91
Discussion 93

Chapter 5: Mutational order and epistasis determine the consequences of FBXW7 mutations during colorectal cancer evolution 96
Introduction 96
Results 102
Mutant *FBXW7*$^{-/-}$ recapitulate known effects on downstream targets 102
FBXW7$^{-/-}$ mutants do not exhibit a competitive proliferative advantage over wildtype organoids 106
FBXW7 mutation provides evidence of epistasis 110
Order of *FBXW7* mutation determines its transcriptional effect 116
scRNAseq data suggests dedifferentiation towards a fetal state in *FBXW7* mutant organoids 119
Transcriptional effects of mutational order are driven by local and global changes in chromatin accessibility 127
Effects of our observed timing of mutation are recapitulated in patient data 131
Discussion 134

Chapter 6: Conclusion 140
FBXW7 mutations in different settings as an exemplar mutation to glean insights into intratumoral heterogeneity 141
Insights into the origin and functions of ITH 142
Epistasis and the timing of a mutation suggests one-off sequencing incompletely characterises the cell 145
Implications on screening for cancer 146

Conclusion.....147

References.....148

List of abbreviations

53BP1	P53 binding protein 1
7-AAD	7-Aminoactinomycin D
AAVS1	Adeno-associated virus integration site 1
ADAMTSL1	A disintegrin and metalloproteinase with thrombospondin motif-like 1
ADP	Adenosine diphosphate
AKAP8	A-kinase anchoring protein 8
AKT2	Protein kinase B
AMP	Adenosine monophosphate
ANXA1	Annexin A1
AP-1	Activator protein-1
APAF-1	Apoptotic protease activating factor-1
APC	Adenomatous polyposis coli
APC/C	Anaphase promoting complex/cyclosome
APOBEC	Apolipoprotein B mRNA-editing enzyme, catalytic polypeptide
ARID1A	AT-rich interactive domain-containing protein 1A
ASH1	Asymmetric synthesis of *HO*
ATACseq	Assay of transposase accessible chromatin sequencing
ATCC	American Type Culture Collection
ATM	Ataxia-telangiectasia mutated
ATP	Adenosine triphosphate
ATP5A	ATP synthase subunit alpha
AXIN2	Axis inhibition protein 2
BACH1	BTB and CNC homology 1

BAX2	BCL2-associated X protein
BCA	Bicinchoninic acid
BCL2	B-cell lymphoma 2
BIN3	Bridging integrator 3
BRAF	V-Raf murine sarcoma viral oncogene homolog B
BRCA	Breast cancer gene
BSA	Bovine serum albumin
BME	Basement membrane extract
BMP	Bone morphogenetic protein
BUB	Budding uninhibited by benzimidazoles
CBC	Crypt-base columnar
CCNE1	Cyclin E1
CDC4	Cell division control protein 4
cDNA	Complementary DNA
CHD8	Chromodomain-helicase-DNA binding protein 8
CIN	Chromosomal instability
cJUN	Cellular homolog of oncoprotein v-jun
CMTR2	Cap methyltransferase 2
cMYC	Cellular myelocytomatosis
COAD	Colorectal adenocarcinoma
COSMIC	Catalogue Of Somatic Mutations In Cancer
CPD	CDC4 phosphodegron
CRC	Colorectal cancer
CREB3	cAMP responsive element binding protein 3
CREBRF	CREB3 regulatory factor

CRISPR	Clustered regularly interspaced short palindromic repeats
CRUK	Cancer Research United Kingdom
CTCFL	CCCTC-binding factor like
CUL1	Cullin 1
DAPI	4′,6-diamidino-2-phenylindole
DAPK	Death associated protein kinase
DBS	Doublet base substitutions
DKK2	Dickkopf WNT signaling pathway inhibitor 2
DMEM	Dulbecco′s Modified Eagle′s Medium
DNA	Deoxyribonucleic acid
DNMT3B	DNA Methyltransferase 3 Beta
DPBS	Dulbecco’s PBS
DR5	Death-receptor 5
EB1	End-binding 1
ECAR	Extracellular acidification rate
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
ELISA	Enzyme-linked immunosorbent assay
EMT	Epithelial-mesenchymal transition
ERBB2	Erb-b2 receptor tyrosine kinase 2
ERBB3	Erb-b2 receptor tyrosine kinase 3
ERK	Extracellular signal-regulated kinase
EVPL	Envoplakin

FAP	Familial adenomatous polyposis
FAS	Fas cell surface death receptor
FAT1	FAT atypical cadherin 1
FBS	Fetal bovine serum
FBXW7	F-box and WD repeat domain containing 7
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
FOXO	Forkhead box O
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GDP	Guanosine diphosphate
GEMIN2	Gem nuclear organelle associated protein 2
GFP	Green fluorescent protein
GLUT1	Glucose transporter 1
GSEA	Gene set enrichment analysis
GSK3β	Glycogen synthase kinase 3 beta
GSTM1	Glutathione S-transferase mu 1
GTP	Guanosine triphosphate
H3K27ac	Histone H3 lysine-27 acetylation
H3K27me3	Histone H3 lysine-27 trimethylation
HBSS	Hank's balanced salt solution
HDAC2	Histone deacetylase 2
HDI	Human development index
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HLA-DQB1	Major histocompatibility complex, class II, DQ beta 1
HRAS	Harvey rat sarcoma virus

HSP60	Heat shock protein 60
ID	Indel
IF	Immunofluorescence
IFN-γ	Interferon gamma
IGF	Insulin-like growth factor
IHE	Institute for Health Economics
ITH	Intratumoural heterogeneity
JAK2	Janus kinase 2
KMT2D	Histone-lysine N-methyltransferase 2D
KRAS	Kirsten rat sarcoma virus
KRT17	Keratin 17
KRT20	Keratin 20
LEF	Lymphoid enhancer factor
LFQ	Label-free quantification
LGR5	Leucine-rich repeat-containing G-protein-coupled receptor 5
LMAN1	Lectin mannose binding 1
LMF1	Lipase maturation factor 1
LOH	Loss of heterozygosity
LRP5	Low density lipoprotein receptor-related protein 5
LRP6	Low density lipoprotein receptor-related protein 6
LUM	Lumican
MAD	Mitotic arrest deficient
MAFG	V-maf avian musculoaponeurotic fibrosarcoma oncogene homolog G
MCFD2	Multiple coagulation factor deficiency 2
MCL1	Myeloid cell leukemia 1

MDM2	Mouse double minute 2 homolog
MEK	Mitogen-activated protein kinase kinase
METTL15	Methyltransferase 15
MKI67	Marker of proliferation ki67
MLH	MutL homolog
MMP1	Matrix metallopeptidase 1
MMP3	Matrix metallopeptidase 3
MMP7	Matrix metallopeptidase 7
MMP10	Matrix metallopeptidase 10
MMP12	Matrix metallopeptidase 12
MMP19	Matrix metallopeptidase 19
MMR	Mismatched repair
MOI	Multiplicity of infection
MRE11	Meiotic recombination 11
MRS2	Mitochondrial RNA splicing 2
MSH	MutS homolog
MSI	Microsatellite instability
MSTFA	N-Methyl-N-trimethylsilyltrifluoroacetamide
MTBE	Tert-butyl methyl ether
mTOR	Mammalian target of rapamycin
MUC6	Mucin 6
MUC16	Mucin 16
MYC	Myelocytomatosis
NBS	Nijmegen breakage syndrome
NCAM	Neural cell adhesion molecule

NDRG1	N-myc downstream regulated 1
NES	Normalised enrichment score
NGS	Next generation sequencing
NICD	NOTCH intracellular domain
NIST	National Institute of Standards and Technology
nMYC	Neural myelocytomatosis
NOTCH	Neurogenic locus notch homolog
NOTCH1	Neurogenic locus notch homolog protein 1
NOTUM	Notum, Palmitoleoyl-Protein Carboxylesterase
OCR	Oxygen consumption rate
P4HA2	Prolyl 4-hydroxylase subunit alpha 2
PARP1	Poly(ADP-ribose) polymerase 1
PBS	Phosphate buffered saline
PCA	Principal component analysis
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PI3K	Phosphoinositide 3-kinase
PIDD	p53-induced protein with a death domain
PKA	Protein kinase A
PMS	Postmeiotic segregation increased
PMSF	Phenylmethanesulfonyl fluoride
POLD	DNA polymerase delta
POLE	DNA polymerase epsilon
POLR1A	RNA polymerase I subunit A
PPP1R12C	Protein phosphatase 1 regulatory subunit 12C

PUMA	p53 upregulated modulator of apoptosis
RAF	Rapidly accelerated fibrosarcoma
RASCAL	The Kirsten ras in-colorectal-cancer collaborative group
RB	Retinoblastoma
RBL2	RB transcriptional corepressor like 2
RBX1	Ring-box 1
RIPA	Radioimmunoprecipitation assay
RNA	Ribonucleic acid
RNF43	Ring finger protein 43
RNP	Ribonucleoprotein
RPA	Replicated protein A
RPS15	Ribosomal protein S15
RSC	Regenerative stem cell
RT-PCR	Reverse transcription-PCR
RUNX1	Runt-related transcription factor 1
RYR2	Ryanodine receptor 2
SBE	SMAD-binding DNA element
SBS	Single base substitution
SCA1	Stem cell antigen 1
SCF	Skp1-Cdc53/Cullin-F-box-protein
scRNAseq	Single-cell RNA sequencing
SCRT2	Scratch family transcriptional repressor 2
SERPINB9	Serine protease inhibitor B9
sgRNA	Single-guide RNA
SRRM2	Serine/arginine repetitive matrix 2

SKP1	S-phase kinase-associated protein 1
SLC2A1	Solute carrier family 2 member 1
SMAD2	Mothers against decapentaplegic homolog 2
SMAD3	Mothers against decapentaplegic homolog 3
SMAD4	Mothers against decapentaplegic homolog 4
SMARCA4	SWI/SNF-related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4
SNAI1	Snail family transcriptional repressor 1
SNP	Single nucleotide polymorphism
SOS	Son of sevenless
SSCP	Single-strand conformation polymorphism
ssDNA	Single-stranded DNA
STAG2	Stromal antigen 2
STAT3	Signal transducer and activation of transcription 3
STXBP2	Syntaxin binding protein 2
SWI/SNF	Switch/sucrose non-fermentable
SYP	Synaptophysin
TAGLN	Transgelin
TAZ	Transcriptional coactivator with PDZ-binding motif
TBP	TATA-box binding protein
TCGA	The Cancer Genome Atlas
TCF	T-cell factor
TEAD1	Transcriptional enhanced associate domain 1
TEAD2	Transcriptional enhanced associate domain 2
TERT	Telomerase reverse transcriptase
TET2	Tet methylcytosine dioxygenase 2

TGFα	Transforming growth factor alpha
TGFβ	Transforming growth factor beta
TIEG	TGF-β-inducible early response gene
TMB	3,3',5,5'-Tetramethylbenzidine
TMCS	Chlorotrimethylsilane
TMUB1	Transmembrane and ubiquitin like domain containing 1
TP53	Tumour protein p53
TTC2	Tetratricopeptide repeat domain-containing protein 2
QoL	Quality of life
QUASAR	Quick and Simple and Reliable
UK	United Kingdom
UMAP	Uniform manifold approximation and projection
UN	United Nations
UPRT	Uracil phosphoribosyltransferase homolog
UPS	Ubiquitin-proteosome system
VAF	Variant allele frequency
WES	Whole exome sequencing
WIF1	Wnt inhibitory factor 1
WNT	Wingless/integrated
WNT3A	WNT family member 3a
YAP	Yes-associated protein
ZNF449	Zinc finger protein 449
ZNRF3	Zinc and ring finger 3
β-TrCP	Beta-transducin repeats-containing proteins
γH2AX	Gamma H2A histone family member X

List of Figures

Figure 1: Dendrograms depicting phylogenetic trees in premalignant lesions and in cancer. 18

Figure 2: Validation of RKO and LS513 cell line models. 63

Figure 3: Comparison of proliferation and growth between FBXW7$^{-/-}$ and wildtype cells..... 64

Figure 4: Comparisons of apoptosis markers between FBXW7$^{-/-}$ and wildtype cells. 65

Figure 5: RNAseq results between wildtype and FBXW7$^{-/-}$ cells. 67

Figure 6: RNAseq results between wildtype and FBXW7$^{-/-}$ cells for MMP genes. 68

Figure 7: Analysis by GSEA reveals upregulation of oxidative phosphorylation in FBXW7 $^{-/-}$ cells..... 69

Figure 8: Validation of upregulated oxidative phosphorylation in FBXW7 $^{-/-}$ cells. 71

Figure 9: DNA damage was increased in wildtype cells co-cultured with FBXW7$^{-/-}$ cells. ... 75

Figure 10: Comet assay showed DNA damage was increased in wildtype cells co-cultured with FBXW7$^{-/-}$ cells..... 76

Figure 11: AAVS1 knockouts did not show increased DNA damage..... 78

Figure 12: Schematic showing mass spectrometry experiments performed 79

Figure 13: Proteomic analysis of WT and Mut cells. 81

Figure 14: Proteomic analysis of wildtype and CC-WT cells. 83

Figure 15: Proteomic and metabolic analysis of media from wildtype vs FBXW7$^{-/-}$ cells. 85

Figure 16: AKAP8 overexpression induces DNA damage in wildtype cells. 87

Figure 17: Loss of AKAP8 protein abrogates the DNA damage effect seen in neighbouring cells..... 90

Figure 18: Correlation of the FBXW7 cell line models with TCGA-COAD-derived patient samples. 92

Figure 19: Schematic showing the interaction between FBXW7$^{-/-}$ and wildtype cells. 95

Figure 20: Characterisation of W and F organoids..... 105

Figure 21: FBXW7$^{-/-}$ organoids do not influence growth of neighbouring wildtype organoids. 109

Figure 22: Effect of FBXW7 mutation is dependent on the mutational background. 113
Figure 23: Effect of epistasis not an artefact of CRISPR/Cas9 gene editing. 115
Figure 24: Effect of FBXW7 mutation is dependent on the order of mutation. 118
Figure 25: scRNAseq of W and F organoids..... 121
Figure 26: scRNAseq of W and F organoids demonstrated differences on gene signature and pathway analysis..... 122
Figure 27: Cluster 13 in F organoids was marked by genes involved in EMT. 123
Figure 28: GSEA of bulk RNAseq data of F vs W organoids using key pathways derived from scRNAseq analysis. 125
Figure 29: GSEA of bulk RNAseq data of FA vs W and AF vs W for the YAP pathway signature. 126
Figure 30: ATACseq analysis suggested cell state changes are due to alterations in chromatin accessibility. 130
Figure 31: Validation of results in patient data. 133

List of Tables

Table 1: sgRNA sequences for CRISPR-Cas9 gene editing used in Theme 1 32

Table 2: Sequencing primers to validate gene knockout in Theme 1 32

Table 3: Antibodies and dilutions used for western blotting in Theme 1 34

Table 4: Taqman probes used for RT-PCR 34

Table 5: Antibodies and dilutions used for immunofluorescence in Theme 1 40

Table 6: Reagents used to produce human organoid growth media 48

Table 7: sgRNA sequences for CRISPR-Cas9 gene editing used in Theme 2 50

Table 8: Sequencing primers to validate gene knockout in Theme 2 51

Table 9: Antibodies and dilutions used for western blotting in Theme 2 52

Table 10: Antibodies and dilutions used for immunofluorescence in Theme 2 53

Chapter 1: Introduction

Colorectal cancer exerts a heavy burden of disease globally

Colorectal cancer (CRC) is the third most prevalent cancer globally, and the second highest contributor to cancer-related deaths (Sung et al., 2021). In 2020 alone, there were approximately 2 million incident cases of CRC and 1 million deaths. In the United Kingdom (UK), data from Cancer Research UK (CRUK) observed that CRC accounted for 11% of all cancer diagnosed between 2016 – 2018, with 42,886 diagnoses in each of those years (CRUK). CRC is often regarded as a cancer associated with developed countries, and is a marker of socioeconomic development. Using the United Nation's (UN) four-tier Human Development Index (HDI) classification, countries with very high HDI were found to have four times higher incidence rate of CRC compared to countries with low HDI in both males (35.4 per 100,000 vs 6.7 per 100,000) and females (24.4 per 100,000 vs 6.7 per 100,000) (Morgan et al., 2023). A similar pattern was observed with regards to CRC-associated mortality. However, while screening and a greater awareness of healthy lifestyles has led to a decrease in the incidence of CRC in developed countries, the incidence burden of CRC in low and medium HDI countries is expected to double in the next two decades. Furthermore, inadequate access to healthcare infrastructure and insufficient resources for detection and treatment of CRC in low and medium HDI countries lead to an outsized impact on the mortality rates of CRC in those countries (Sung et al., 2021).

The disease burden of CRC must also encompass the significant healthcare costs to healthcare payers, which includes not only patients but healthcare systems. The Swedish Institute for Health Economics (IHE) computed the pan-European cost of CRC, and revealed that direct, indirect and informal costs amounted to €12.2 billion in Europe, and was almost twice as high as the next most costly gastrointestinal cancer (Hofmarcher T). In the UK, the per capita expenditure on CRC amounted to almost €30. Patients with CRC may also suffer from a reduction in their quality of life (QoL), which reflect the complex interplay of treatment

options patients undergo during their treatment process. QoL could be negatively affected by a range of factors including surgery, chemotherapy or radiotherapy side effects, complications relating to the creation of an artificial opening in the abdominal wall to facilitate defecation also known as an ostomy, and stress arising from healthcare costs and time away from work, family and hobbies (Blum-Barnett et al., 2019, Kristensen et al., 2022).

Taken together, these pressure points provide an idea of the considerable burden which CRC exerts, and emphasizes the urgency required in developing novel ideas to diagnose and treat CRC more effectively.

The environmental and genetic basis for disease in CRC

Environmental factors have been shown to contribute to CRC. This includes a diet high in red and processed meats (Chan et al., 2011), excessive alcohol consumption (Fedirko et al., 2011), obesity (Ma et al., 2013), diabetes mellitus (Jiang et al., 2011), as well as a sedentary lifestyle (Boyle et al., 2012). Together with non-modifiable risk factors for CRC such as increased age and male gender, these environmental factors likely account for a large proportion of the increased risk at a population level, but do not explain the pathogenesis of CRC at an individual level, for which a molecular basis must be sought. Environmental factors may interact with genetic factors, giving rise to CRC.

CRC carcinogenesis proceeds via three main genetic pathways. Adenocarcinomas can arise as a result of chromosomal instability (CIN), microsatellite instability (MSI) and through the serrated neoplasia pathway (Nguyen, Goel and Chung, 2020).

Chromosomal instability (CIN)

Up to two-thirds of CRC is acquired via the CIN pathway, and therefore represents the most common pathway by which CRC arises (Vogelstein et al., 2013, Grady and Carethers, 2008).

CRC arising from the CIN pathway is characterized by copy number alterations, indels, translocations, amplifications and loss of heterozygosity. These genetic abnormalities can arise due to dysfunctional chromosomal segregation, loss of heterozygosity (LOH) at a tumour suppressor gene, disordered telomere viability, and an ineffective DNA-damage response machinery (Nguyen, Goel and Chung, 2020).

Dysfunctional chromosomal segregation occurs when there is dysregulated signaling at the mitotic checkpoint, which is critical in delaying the onset of anaphase until there is a high fidelity of chromosomal segregation (Pino and Chung, 2010). Dysfunctional chromosomal segregation therefore gives rise to aneuploidy and abnormal chromosomal numbers in daughter cells. Proteins from the *mitotic arrest-deficient* (*MAD*) and *budding uninhibited by benzimidazoles* (*BUB*) family of genes interact with anaphase-promoting complex/cyclosome (APC/C) to induce a conformational change in APC/C which prevents binding and ubiquination of substrates and the premature progression of the cell into anaphase (Herzog et al., 2009). Mutations in the *hBUB1* and *hBUBR1* genes have been shown to induce CIN in cell lines (Cahill et al., 1998), while mutations of *hMAD2* and *hMAD1* have been detected in both breast cancer and leukemia (Li and Benezra, 1996, Jin, Spencer and Jeang, 1998). Upstream of metaphase, abnormal proteins from the genes *hZw10*, *hZwilch*, and *hRod* functioning at the spindle checkpoint during prometaphase resulting in improper chromosomal separation have been reported in CRC (Wang et al., 2004).

LOH is also key to CRC as the loss of an allele at a heterozygous locus may result in the loss of a tumour-suppresor gene and lead to cancer (Lasko, Cavenee and Nordenskjöld, 1991, Fearon et al., 1990, Chang et al., 2005). In CRC, chromosomes which experience the highest frequency of LOH are at chromosomes one, five, eight, seventeen and eighteen (Thiagalingam et al., 2001), though LOH can occur in any chromosome. LOH has been postulated to occur as a result of the aforementioned dysfunctional chromosomal segregation which gives rise to aneuploidy and whole chromosome loss (Doxsey, 1998), as observed in the LOH of 18q (Bardi et al., 1995, Remvikos et al., 1997), but may also arise from double-

stranded DNA breaks coupled with a defective double-strand repair-recombination mechanism (Ferguson et al., 2000), therefore giving rise to chromosomal breakage and fusion events which lead to interchromosomal recombination (Thiagalingam et al., 2001). The latter sequence results in subchromosomal LOH patterns. LOH patterns in CRC have been described in a number of studies. LOH at 14q32 and 22q12.2-q13.33 have been observed in 44.1% (Bando et al., 1999) and 28.4% (Zhou et al., 2002) of CRC respectively while genetic loss at 10q has been associated with an earlier onset of sporadic CRC (Fawole et al., 2002).

Telomeres are hexanucleotide repeats of TTAGGG and are located at the 3' ends of chromosomes. These telomeres serve to prevent the fusion and degradation of chromosomes (Mathieu et al., 2004). Telomeres progressively shorten due to the absence of the catalytic component (hTERT) of the reverse transcriptase telomerase in non-stem cells. As cells age, telomeres become progressively shortened until the telomere reaches its minimum size of approximately 1.5 kb. At this point, cell proliferation is stopped, mortality stage 1 (M1) is reached (Wright, Pereira-Smith and Shay, 1989), and DNA damage checkpoints trigger the senescence pathway. The shortening of telomeres, however, results in the formation of end-to-end fusions either between sister chromatids, or between different chromosomes, leading to dicentric chromosomes (Hastie et al., 1990, Counter et al., 1994, Maser and DePinho, 2002). During cell replication, these fusions present as anaphase bridges, resist being pulled to opposite poles, and result in chromosome breaks that lead to chromosomal deletions, translocations and amplifications. As these chromosomal abnormalities themselves have uncapped telomeric ends, further breakage-fusion-bridge cycles can form, resulting in extensive chromosomal aberrations and CIN (McClintock, 1941).

Mechanisms for DNA damage repair are important to maintain chromosomal integrity, and have been validated based on datasets from The Cancer Genome Atlas (TCGA) (Knijnenburg et al., 2018, Reilly et al., 2019). It is worth noting that DNA damage repair mechanisms refer to both mismatch repair giving rise to MSI, which will be discussed separately, and errors in homologous recombination giving rise to genome instability. Homologous recombination is

required when the DNA suffers a double-stranded break (Ranjha, Howard and Cejka, 2018). Double-stranded breaks are detected during S phase by poly(ADP-ribose) polymerase 1 (PARP1), which attaches ADP-ribose molecules around the double-stranded break (Haince et al., 2008) for recognition by meiotic recombination 11 (MRE11), RAD50, and Nijmegen breakage syndrome (NBS) proteins to form an MRN complex. The MRN complex is an exonuclease which produces single-stranded DNA (ssDNA) around the break, and triggers ATM-modulated downstream repair factors (You and Bailis, 2010, Huen, Sy and Chen, 2010, Dodson et al., 2010). ssDNA is stabilized by replication protein A (RPA) (Maréchal and Zou, 2013), allowing the BCDX2 complex to promote binding of RAD51 to the ssDNA overhand (Candelli et al., 2014). RAD51 is responsible for producing the homologous DNA region with the aid of a sister chromatid (Qi et al., 2015). Defects in molecules responsible for homologous recombination may give rise to CRC, as has been demonstrated with RAD51 (Hiramoto et al., 1999) and other proteins related to the formation of the MRN complex (Krupa et al., 2011).

Microsatellite instability (MSI)

MSI is a readout of deficient DNA mismatch repair, and is present in 15% of CRC (Aaltonen et al., 1998, Sinicrope, 2018). Microsatellite regions refer to nucleotide repeat sequences consisting of cytosine and adenine (Thibodeau, Bren and Schaid, 1993). In long repetitive sequences such as those presented at microsatellites, DNA polymerase is not able to bind effectively to such sequences during DNA synthesis, resulting in either an incorrect number of bases being bound, or an incorrect nucleotide being bound (Vilar and Gruber, 2010, Boland and Goel, 2010). Mutations at these regions give rise to temporary insertion-deletion loops which form DNA hairpins. The mismatch repair (MMR) system functions by correcting such mutations, and plays a critical role in ensure these mutations do not get propagated to daughter cells (Boland and Goel, 2010).

The MMR system consists of two classes of genes, the *Mut S homologue* (*MSH)*, and the *Mut L homologue* (*MLH*). Homologues of the *Mut S* gene include *MSH1* to *MSH6*, while homologues of the *Mut L* gene include *MLH1* to *MLH3*, and *PMS1* and *PMS2*. A heterodimer is formed between two homologues from each family, giving rise to an array of different homologues which have different affinities to specific DNA mismatch mutations (Aquilina et al., 1994, Strand et al., 1995, Palombo et al., 1995). After insertion-deletion loops are recognized by the MMR proteins, excision of the mismatched nucleotide is performed by exonuclease 1 and proliferating-cell nuclear antigen. DNA polymerase δ and DNA ligase complete the repair by resynthesis of the correct base pair and re-ligation (Jiricny, 2006, Hoeijmakers, 2001).

Many mutations in microsatellite regions have been found in genes which are responsible for an array of cellular functions, including genes responsible for growth factor receptors TGF-β receptor II and IGF receptor II, histone modifier HDAC2, and the pro-apoptotic factor BAX2 (Duval and Hamelin, 2002). In CRC, MSI tumours give rise to a unique CRC phenotype, characterized by a less aggressive natural history, younger onset of cancer, increased likelihood of synchronous and metachronous disease, cancer which predominantly affects the right colon, and a histologic appearance of increased lymphocytic infiltration and a mucinous or signet ring appearance (Giardiello et al., 2014). Loss of MMR genes may be germline or non-inherited and associated with promoter hypermethylation resulting in loss of gene expression. Patients with germline MMR deficiency are diagnosed to have Lynch Syndrome, and develop cancer by the third or fouth decade of life (Watson et al., 2008). MSI has been discovered in other cancers, including ovarian, endometrial, gastric cancers, and glioblastoma multiforme (Boland et al., 1998, Pal et al., 2008).

Serrated neoplasia

The serrated neoplasia pathway is responsible for 15 to 25% of CRC (Toyota et al., 1999, Snover, 2011, Jass, 2007), and is named for its characteristic histologic appearance of stellate

crypt folding patterns which appear as serrations on hematoxylin and eosin staining (Nagtegaal et al., 2020). The serrated neoplasia pathway is incompletely understood, but an activating V600E mutation in BRAF, which regulates the mitogen-activated protein kinase pathway, is the putative initial mutation event (O'Brien et al., 2004, Davies et al., 2002, Crockett and Nagtegaal, 2019). *BRAF* mutations result in constitutive activation of the mitogen-activated protein kinase-ERK pathway which leads to uncontrolled cellular proliferation and division (O'Brien et al., 2006, Chan et al., 2002, Spring et al., 2006, Kambara et al., 2004).

Activation of this pathway likely proceeds down two separate routes. In the first route, an MSI pattern is recapitulated and is discussed in the subsequent section. In the second route, tumors with BRAF mutations demonstrate activation of Wnt and TGFB signaling, often in the absence of an *APC* mutation (Nguyen, Goel and Chung, 2020, Borowsky et al., 2018). Wnt pathway activation differs from its CIN counterpart in that truncating mutations do not occur. Instead, Wnt appears to be stimulated via two potential pathways, both involving *RNF43*, which encodes for an E3 ubiquitin ligase (Bond et al., 2016). In the first instance, RNF43, which normally suppresses Wnt function through R-spondin, is mutated (Bond et al., 2016, Yan et al., 2017). Alternatively, RNF43-ZNRF3 complexes, which serve to promote endocytosis of the frizzled-LRP5-LRP6 complex is disrupted, resulting in increased Wnt-frizzled mediated transcription activation (Koo et al., 2012).

Ultimately, *BRAF* activation results in widespread methylation of CpG islands. CpG islands are concentrated clusters of cytosine and guanine nucleotides, and the methylation of such CpG islands gives rise to the CpG island methylator phenotype (CIMP) (Bond et al., 2018, Weisenberger et al., 2006). CpG islands can be found in the promoter region of genes (Bird, 1986), and upstream of key tumour suppressor genes. Mutant BRAF upregulates the transcriptional repressor MAFG (JE et al., 2015), which in turn recruits BACH1, CHD8, and DNMT3B, resulting in a hypermethylation phenotype (Fang et al., 2014). Clinically, CRC which develops by the serrated adenoma pathway is considered to be more aggressive, and

progresses from a premalignant phenotype to cancer within a shorter period of 1-3 years, and are therefore both mechanistically and clinically considered different from CIN tumours (Umar et al., 2004).

Driver mutations play an important role in CRC carcinogenesis

Particularly in the CIN pathway of carcinogenesis, the accumulation of mutations in key tumour suppressor genes and oncogenes plays a critical role in the progression towards cancer (Huang et al., 2018). Each driver mutation confers a slight survival advantage to the subpopulation of cells harbouring the mutation (Raskov et al., 2020), and leads to a proliferative advantage for that clone. Encapsulated in the adenoma-carcinoma sequence, first theorised by Vogelstein et al. (Vogelstein et al., 1988), normal colonic epithelium acquires mutations leading first to an adenoma with low-grade dysplasia, before progressing to high-grade dysplasia and finally carcinoma. In this seminal paper by Vogelstein et al., the technique of allelic-deletion analyses through a combination of restriction enzymes was used to localise the position of putative driver mutations. In this way, the gene responsible for familial adenomatous polyposis syndrome was localised to chromosome 5q, while other driver mutations, unknown at the time, were linked to 17p and 18q. Vogelstein's hypothesis of these driver mutations was validated when the identity of the mutations were discovered. In CRC, four most common driver mutations are *KRAS*, *TP53*, *APC*, and *SMAD4* (Schell et al., 2016).

The contemporary understanding of driver mutations in CRC has expanded far beyond these four driver mutations first described by Vogelstein. A pan-cancer mutation analysis based on specimens from The Cancer Genome Atlas (TCGA) revealed an intriguing landscape of driver mutations across 33 different cancer types (Bailey et al., 2018). This analysis concluded that the majority of driver mutations in cancer are unique to a single cancer type (142 out of 258 driver genes), with 87 driver genes being found in two or more cancer types.

Furthermore, 29 additional driver mutations were ubiquitous among a number of cancer types, including *TP53*, *PIK3CA*, *KRAS*, *PTEN*, and *ARID1A*, each present in at least 15 cancer types. In CRC, unique driver mutations included *AMER1*, *TCF7L2*, *ZFP36L2*, *PCBP1*, *SOX9*, and *TGIF1*.

A different study which analysed CRC samples from the TCGA as well as two other published cohorts (Giannakis et al., 2016) identified genes which were commonly found in CRC, as well as pathways associated with those genes. Pathway identification enabled a deeper understanding of the molecular pathways which become perturbed in CRC. Notably, these genes may not represent driver mutations exclusively, and may include passenger mutations. In total, 62 significantly mutated genes were identified, with 40 in MSI tumours only, 9 in MSS tumours only, and 13 shared between both. A further dichotomy into MSS and MSI cancers provided insights into these two subtypes of CRC. Genes involved in Wnt signalling were mutated in both MSI and MSS subtypes, even though the genes involved were different. In MSS tumours, these genes included *AMER1*, *TGIF1*, *ELF3*, *SMAD4*, and *SMAD2*, while in MSI tumours, these genes included *ZNRF3*, *BCL9L*, *DOCK3*, *CD58*, *AXIN2*, and *WNT16*. The remarkable convergence of different sets of genes involved in the same pathway perhaps reflects the central role which the Wnt pathway plays in CRC, such that its perturbation is a necessary step in oncogenesis regardless of MSI or MSS status. In addition to Wnt pathway genes, a set of immune-related genes were also identified in MSI cancers, such as *XYLT2*, *ZFP36L2*, *KLF3*, and *CD58*, reflecting the propensity of MSI cancers towards immune evasion.

Finally, a recent preprint which sequenced 2023 colorectal cancers significantly expanded the repertoire of driver mutations associated with CRC, increasing the number of known CRC driver mutations by 117 (Cornish et al., 2022), and involved different molecular pathways including the abovementioned Wnt signalling pathway, but also the MAP kinase pathway.

One of the concepts alluded to in this thesis includes the impact which the timing of a driver mutation can have on CRC. Here, the sequential acquisition of colonic organoids with first

APC, then *SMAD4*, *TP53*, and finally *KRAS*, and *PIK3CA* imbued normal human organoids with cancer characteristics (Matano et al., 2015). These organoids morphologically resembled highly proliferative cystic organoids, and were able to grow in the absence of growth factors otherwise required by normal organoids. These results were corroborated by a similar article published by experiments from the laboratory of Hans Clevers (Drost et al., 2015).

Driver mutations have also been identified in normal colonic epithelium

While the presence of driver mutations in CRC is not unexpected, the identification of driver mutations in phenotypically normal tissue raises questions surrounding their function and effect in epithelial cells. The identification of these driver mutations was observed in whole genome sequencing of 571 laser micro-dissected colonic crypts, identifying truncating mutations in *AXIN2*, and *STAG2*, as well as missense hotspot mutations in *PIK3CA*, *ERBB2*, *ERBB3*, and the gene of interest in this study, *FBXW7* (Lee-Six et al., 2019). In this section, I will discuss *FBXW7* in detail, and provide an overview of the other driver mutations which have been identified in normal colonic epithelia.

FBXW7

While a context specific discussion on the role of *FBXW7* will be provided in the introductory section to both the abovementioned themes, here, a general overview of the role and functions of *FBXW7* is provided.

FBXW7 or F-box and WD repeat domain containing 7 (FBXW7) is the substrate recognition component of the Skp1-Cdc53/Cullin-F-box-protein complex (SCF/β-TrCP) (Yeh, Bellon and Nicot, 2018). This SCF complex is an E3 ubiquitin ligase, and is the final component of a target protein with the ubiquitin-proteosome system (UPS). The SCF complex ubiquitinates the target proteins and triggers proteasomal degradation of that target protein. FBXW7 is one of 4 proteins in this SCF complex, the other components being Skp1, Cul-1 and Rbx1.

FBXW7 plays a critical role in the SCF complex, as it is responsible for recognizing proteins which have been phosphorylated at a CDC4 phosphodegron (CPD) motif. This phosphodegron for FBXW7 recognition is (S/T)PXX(S/T), whereby S/T refers to serine or threonine residues, P refers to proline, and X is any amino acid (Welcker and Clurman, 2008). Briefly, target proteins are phosphorylated by a Glycogen synthase kinase 3 beta (GSK3β) process (Welcker et al., 2004). Following this, FBXW7 recruits the target protein to the SCF complex where the target protein is ubiquitinated. The poly-ubiquitinated protein is marked for degradation by the 26S proteasome (Thompson et al., 2021).

FBXW7 is known to recognise at least 90 different substrates (Arabi et al., 2012). Many of the substrates which FBXW7 targets are proto-oncogenes which are key regulators of gene transcription and stem cell differentiation, such that mutations in *FBXW7* impair the proper degradation of these proteins, and result in unabated cancer cell proliferation, and disordered differentiation. Some of these target proteins include c-Myc, Cyclin E and Notch1. *FBXW7* is the most commonly mutated gene in the SCF complex (Yumimoto and Nakayama, 2020). In CRC, *FBXW7* is associated with both tumour initiation and progression (Karczmarski et al., 2019).

A unique feature of the *FBXW7* gene is that mutations in this gene cannot be easily classified into either gain or loss of function changes. This unique feature is related to the structure of the protein. The FBXW7 protein comprises of 8 WD40 repeat regions which each comprise one propellor blade of an eight-bladed propellor. Arginine residues at the propellor tips bind to the CDC4 phosphodegron motif, and are responsible for holding the protein in place for subsequent ubiquitination. Mutations occurring at these arginine residues lead to severe dysfunction of FBXW7, and therefore are the most common mutations seen in mutant *FBXW7*, which affect arginine residues at positions 465, 479 and 505. Missense mutations at other positions of the gene could affect amino acid residues at the propellor blade, but not specifically the propellor tip, and lead to a less severe protein dysfunction. Analysing the distribution of *FBXW7* mutations based on TCGA data, Davis and Tomlinson illustrated the

unique pattern of mutations in *FBXW7* (Davis and Tomlinson, 2012). The authors demonstrated that propellor tip mutations in one allele tended not to be associated with a classical "second hit" to the other allele, and that two protein-truncating mutations were never found. Together, this implied there was selection pressure towards a non-complete loss of *FBXW7* function, and that mutations occurring at different locations along the gene affected the function of the protein to different degrees which reflected the optimal amount of FBXW7 activity in cancer. Put together, these findings suggested a "just enough" model of carcinogenesis in which there exists an optimal level of dysfunction which is preferable to complete gain or loss of function. This was validated in mouse models of $Fbxw7^{R482Q/+}$, whereby the R482Q is the mouse equivalent of the human R479Q mutation, and $Fbxw7fl^{(\Delta ex5)/+}$, which represented a heterozygous *FBXW7* null model (Davis et al., 2014). The authors demonstrated that *Vill-Cre;*$Fbxw7^{R482Q/+}$ in combination with *APC* mutant mice generated a greater burden of intestinal polyposis than *Vill-Cre;*$Fbxw7fl^{(\Delta ex5)/+}$ and *APC* mutant mice, suggesting that different genotypic perturbations of the same gene could result in different phenotypes depending on the degree of FBXW7 dysfunction.

More recently, *FBXW7* has also been found to regulate chromatin accessibility (Thirimanne et al., 2022). In HCT116 cell lines, *FBXW7* mutations affected Histone H3 lysine-27 acetylation (H3K27ac) and Histone H3 lysine-27 trimethylation (H3K27me3), which represent active versus repressed chromatin respectively. Furthermore, the AP-1 motif, which is bound by the FBXW7 substrate Jun, was enriched in 32% of differential H3K27ac sites upregulated by the *FBXW7* mutation. These findings suggest a synergistic effect, where functional FBXW7 loss upregulates both its substrate as well as regions of the chromatin where those substrates bind to, therefore potentiating the effect of the *FBXW7* mutation.

To summarise, FBXW7 is a key regulatory protein which counts among its substrates many proteins which themselves regulate cell proliferation and differentiation, and in turn can potentiate cancer. In spite of this, *FBXW7* is not a classical tumour suppressor in that it does not satisfy the Knudson 2-hit hypothesis for tumour suppressors *in vivo*. The mutational

pattern of *FBXW7* affects the degree of functional dysregulation, which is profound, as evidenced not only in its role in the ubiquitylation and degradation of substrates, but also in modifying chromatin accessibility.

STAG2

STAG2's function arises from its many interactions with the chromosome. STAG2 is a critical component of the cohesin complex, which binds to centrosomes and is responsible for sister chromatid cohesion and chromosome separation during mitosis (Romero-Pérez et al., 2019). STAG2 depletion may also result in increased telomeric recombination, resulting in telomere sister chromatid exchange (T-SCE). In addition, STAG2 is responsible for binding to CTCF and non-CTCF binding sites, therefore maintaining topologically associating domain (TAD) boundaries which maintain DNA-DNA loop topography and accessibility. Mutations in STAG2 can therefore result in changes in the 3-dimensional structure of DNA, and in turn affect the transcription of genes (Viny et al., 2019). As a result, *STAG2* mutants have been found to maintain hematopoietic stem/progenitor cells (HSPC) in a more stem-like state, decreasing lineage commitment, decreasing quiescence, and enhancing stem cell self-renewal. In Ewing sarcoma, tumors harboring *STAG2* mutations possess a more aggressive metastatic phenotype (Adane et al., 2021). More pertinently in the normal colon, the identification of STAG2 clonal marks on immunohistochemistry of formalin fixed paraffin embedded (FFPE) slides has allowed for the direct observation of crypts with cells harboring *STAG2* mutants. *STAG2* mutants were shown to form clone sizes larger than would be expected based on neutral stem cell competition dynamics, suggesting that cells with a *STAG2* mutation have a biased stem cell behaviour which favours crypt fixation and increased crypt fission (Nicholson et al., 2018).

AXIN2

AXIN2 is a critical component of the APC-destruction complex, which comprises the other proteins APC, CK1-α, and GSK-3β (Parker and Neufeld, 2020). As part of this destruction

complex, AXIN2 plays a critical role in regulating β-catenin by phosphorylating β-catenin and marking it for proteasomal degradation. *AXIN2* is therefore a negative regulator of the Wnt/β-catenin pathway and loss of function mutations in *AXIN2* have been found in CRC (Liu et al., 2000). Pathogenic variants of the *AXIN2* gene have also been found to be associated with the FAP-spectrum syndrome AXIN2-associated polyposis in an autosomal dominant fashion, and patients with this condition present with up to hundreds of adenomatous polyposis as well as oligodentia (Lammi et al., 2004).

More recently, gene expression profiles generated from CRC cell lines demonstrated that *AXIN2* is itself a β-catenin target gene, suggesting that AXIN2 may function as a critical negative feedback homeostatic regulator of the Wnt pathway (Yan et al., 2001). The characterisation of *AXIN2* mutations in normal colonic epithelium however remains unexplored.

ERBB2

ERBB2 is a protein which belongs to the EGFR family, and is a tyrosine kinase with intrinsic kinase activity (Stein and Staros, 2000). *ERBB2* mutations are rare in CRC, and may be found in only 3% of patients (Vaghi et al., 2023). In other cancers, *ERBB2* is found to be overexpressed or amplified, resulting in more aggressive cancers due to loss of normal cell-control mechanisms (Hung et al., 1986). *ERBB2* mutations are rare in general, and in CRC may be found in only 3% of patients. In CRC, *ERBB2* mutant cancers are associated with MSI status, as well as increased tumour mutational burden (Uchida and Sugino, 2023). Due to the effectiveness of anti-HER2 agents in breast cancer, there is ongoing interest in testing such agents in the setting of ERBB2-mutant CRC in clinical trials.

ERBB3

Like ERBB2, the ERBB3 protein is also a tyrosine kinase protein which is a member of the EGFR family. ERBB3 expression in CRC has been found to be a negative prognostic marker, and predicts for risk of liver metastasis (Capone et al., 2023). Recently, miRNA-483-3p was

found to result in increased ERBB3 signaling, resulting in the activation of downstream mediators of EMT (Candiello et al., 2023). As with *ERBB2* mutations, *ERBB3* mutations are similarly rare in CRC, and there is a paucity of literature characterizing the effect of *ERBB3* mutations in the setting of normal colonic epithelium.

PIK3CA

Compared to the other driver mutations which have been detected in CRC, *PIK3CA* mutations in CRC are relatively common, and may be found in close to 30% of tumours (Ogino et al., 2014). The gene encodes for the PI3K enzyme, which plays a critical role in the PI3K-AKT pathway, and activating mutations of the *PIK3CA* gene can result in increased cellular proliferation resulting in a range of cancers. These activating mutations cluster at exon 9 and 20 hotspots, and have been found in equal proportions between MSI and MSS CRC (Velho et al., 2005). In light of this critical role which PIK3CA plays in the PI3K-AKT pathway, CRCs which possess a *PIK3CA* mutation are associated with worse prognosis (He et al., 2009, Manceau et al., 2015)

Introduction to the two themes in this thesis

My thesis may be divided into two themes, each with a unique set of experimental procedures, but both using *FBXW7* as the candidate gene to explore both themes. The first theme explores competitive or collaborative clonal interactions in CRC. Our understanding of the function and role of ITH in cancer is an ongoing challenge and hitherto is incompletely described. Specifically, I explore the role of *FBXW7* mutant cancer cells as a clonal population in its interaction with cancer cells wildtype for *FBXW7*. In the second theme, I analyse the role of *FBXW7* mutations which have been discovered in phenotypically normal tissue. Here, I seek to characterise the effect which an *FBXW7* mutation in phenotypically normal tissue may have on the future evolutionary trajectory of those cells within the context of cancer.

Theme 1: Biallelic *FBXW7* knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells

Intratumoral heterogeneity hints at a complex interplay of interactions between driver genes

While mutations in driver mutations certainly confer a strong selective advantage in cells harbouring those mutations, the presence of significant intratumoral heterogeneity (ITH) in cancer suggests that clonal sweeps following the acquisition of these driver mutations do not occur. ITH may be detected by applying multiregional sampling in CRC. A study by Losi et al. (Losi et al., 2005) of ten multiply sampled CRC confirmed the presence of intratumoral heterogeneity (ITH) on single-strand conformation polymorphism (SSCP) analyses in 90% of cases. In two patients, different mutations of the *KRAS* and *TP53* genes were identified in the same tumour. Multi-regional sampling combined with WES provides a further characterisation of ITH. In this way, somatic mutations could be classified as either "ubiquitous," if present in all regions sampled, or "heterogeneous," if present in only some regions sampled. A study by Uchi et al. (Uchi et al., 2016) which sequenced several CRC from a number of patients found 5068 ubiquitous somatic mutations in total, and a further 3107 somatic mutations which were heterogeneous. In addition, 1362 of the heterogeneous somatic mutations were unique to single samples.

Moreover, ITH is a dynamic process which can change along the evolution of the tumour. Multiregional sequencing of cancers suggests that there is an initial period of mutagenesis akin to a "big bang" which results in the generation of mutations that persist late into the evolutionary trajectory of that cancer (Sottoriva et al., 2015). Following this initial period of mutagenesis, there is continued co-evolution of the genome and the epigenome along the transition from normalcy to adenoma and carcinoma (Heide et al., 2022). The changing heterogeneity along the transition from adenoma to carcinoma can be visualised on a evolutionary phylogenetic tree. Here, clonal mutations which are present throughout the

tumour correspond to trunks on the evolutionary tree, while mutations which only occur in a subclone of the tumour are represented by branches. Saito et al. noted that premalignant lesions tended to have shorter trunks with more numerous branches, while advanced CRC lesions are dominated by longer trunks with shorter branches (Saito et al., 2018) (Figure 4).

In addition, single-cell sequencing has greatly facilitated our current understanding of the heterogeneity which lies within each tumour, also known as microheterogeneity (Barber, Davies and Gerlinger, 2015). This microheterogeneity describes the intrinsic heterogeneity of a cancer, and provides an advancement over techniques such as multi-regional sampling described above. Single-cell techniques allow for the entire tumour specimen to be sequenced, and for regions within a tumour to be characterised in greater detail, resulting in the identification not only of differences between subclones, but also of differences within subclones. Microheterogeneity has been observed using somatic copy number aberations (SCNAs) and point mutations to detect current mutation rates in a small number of cells. Mutations observed at the microheterogeneity level were likely acquired many generations before the mutation became pervasive and detectable, whereas microheterogeneity provides insights into the current mutation rate (Alexandrov et al., 2013). In a similar way, single-cell RNA sequencing of glioblastomas revealed cells with varying degrees of stemness co-existing at different regions of the tumour. Microheterogeneity within cancers therefore suggest that the founding cell in cancer may not occupy the most optimal fitness peak, such that tumour subclones may acquire further new driver mutations which enhance the fitness of that cell, resulting in multiple fitness peaks exemplified in branched evolution models (Barber, Davies and Gerlinger, 2015).

These changes may be reflected in the distribution of driver mutations between trunks and branches of the same lesion. In premalignant lesions, 25 out of 51 driver mutations were found in branches, while only 10 out of 45 driver mutations were found in branches in cancer. Accordingly, variant allele frequencies (VAF) were higher in the branches of premalignant lesions relative to cancer. Another study by Cross et al. (Cross et al., 2018a) showed

concordant results. The authors evaluated the presence of somatic mutations in premalignant lesions and CRC, and found that while there were no differences in the median absolute number of SNPs between premalignant lesions and CRC, the proportion of somatic mutations which were ubiquitous was greater in CRC than in premalignant lesions. Taken together, these findings suggest that premalignant lesions tend to exhibit an increased ecological diversity which is reduced along the transition towards cancer, possibly reflecting the stronger selecting pressures found amongst clonal populations present in CRC. Yet, in spite of this reduction in ecological diversity, microheterogeneity continues to be an evolving phenomenon from premalignancy to cancer.

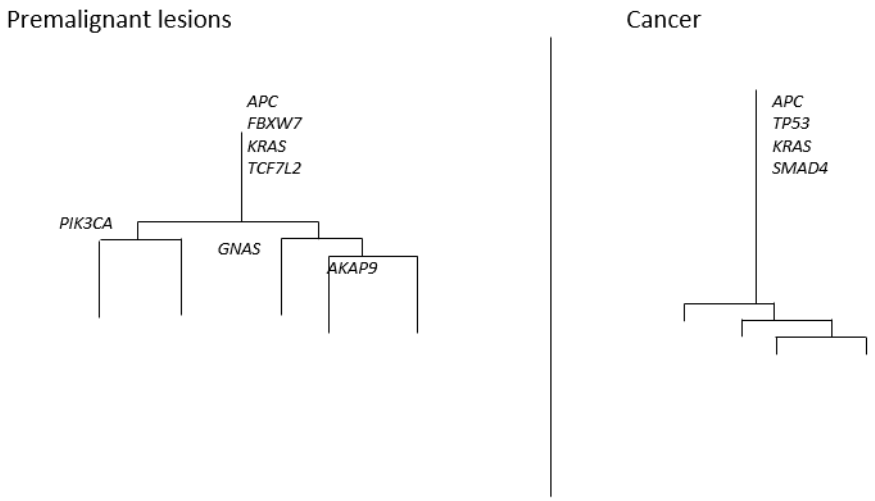


Figure 1: Dendrograms depicting phylogenetic trees in premalignant lesions and in cancer.

In premalignant lesions, trunks tend to be shorter with numerous branches, suggesting the presence of more subclones and greater heterogeneity. In contrast, phylogenetic trees of cancer tend to have longer trunks and shorter branches. Figure adapted from (Cross et al., 2018a) and (Chan and Buczacki, 2021).

Clonal interactions imbue tumours with unique phenotypes

Ample evidence of competitive and cooperative interactions occurring between subpopulations of clones within cancers have been described (Colom et al., 2021, van Neerven et al., 2021a, Snippert et al., 2014, Zhou, Neelakantan and Ford, 2017), and are an ongoing field of investigation. Clonal competition recalls Darwinian evolutionary dynamics in which subclones gain mutations which confer an advantage over other neighbouring subclones. Three seminal articles have revealed the interactions between mutant cells and their wildtype neighbours in the gut. Using a mouse small intestine model with multicolour-labelled oncogene reporters, Yum et al. found that $KRAS^{G12D}$ and $PIK3CA^{H1047R}$ mutant crypts exhibited increased crypt fixation and fission over wildtypes. Mutant crypts interacted with wildtype crypts by reducing the intestinal stem cell fraction in both mutant and neighbouring wildtype cells, priming stem cells towards differentiated cell lineages. Moreover, mutant crypts remodelled the intestinal niche in a paracrine fashion resulting in a reduction of Wnt secretion and an increase in BMP release, leading to apoptosis and differentiation of neighbouring wildtype cells (Yum et al., 2021). In another article by van Neerven et al., *APC* mutant organoids were found to upregulate WNT antagonist genes including *Notum*, *Wif1*, and *Dkk2*. In turn, these WNT antagonists were secreted into the paracrine environment of neighbouring wildtype cells, resulting in increased differentiation of neighbouring wildtype cells, therefore conferring a proliferative advantage to *APC* mutant organoids. This effect was rescued when downstream activation of Wnt by GSK3β inhibitors were administered (van Neerven et al., 2021a). The competitive inhibition of wild-type cells by *APC* mutants was confirmed by Flanagan et al. (Flanagan et al., 2021b). In this study, NOTUM inhibitor treatment resulted in reduced fixation of *APC* mutants. Other examples may be gleaned from the literature. For example, cells which overexpress MYC can result in the non-cell autonomous activation of caspases in neighbouring cells (Di Giacomo et al., 2017). Elsewhere, mutations which target the Hippo pathway by suppressing YAP/TAZ activation

lead to enhanced cellular proliferation and result in epithelial-mesenchymal transition (EMT) (Boopathy and Hong, 2019).

Clonal cooperation is observed when subclones interact in a cooperative way to enhance tumour growth. The paracrine secretion of proteins in one such cooperative interaction. Hobor et al. discovered that cells which would otherwise have been sensitive to cetuximab treatment demonstrated greater resistance to the drug when incubated in media derived from mutant *KRAS* cells (Hobor et al., 2014). This occurred due to the secretion of TGFα and amphiregulin by mutant KRAS cells into the paracrine environment of wildtype cells. Another study analysed the paracrine secretion of progranulin by CRC cells which had become resistant to oxaliplatin chemotherapy. The administration of progranulin to otherwise chemo-sensitive cells led to increased chemo-resistance (Bose et al., 2011). Direct interaction in cancer cell populations have also been observed. Mouse models of breast cancer revealed the interaction between distinct subpopulations of $Hras^{mut}Wnt1^{low}$ and $Hras^{wt}Wnt1^{high}$ cells which together are required for tumour propagation (Cleary et al., 2014). Intriguingly, when Wnt was withdrawn from the biclonal tumour to simulate the administration of targeted therapy, a short-lived regression of the tumour was observed prior to relapse either by recruitment of other Wnt-producing cells or by alternative means of reactivating the Wnt pathway. This observation clearly demonstrated how tumour cells cooperate with different subclones to bring about tumour growth and proliferation.

Similarly, in a model of small cell lung cancer in mice, clonal populations of phenotypically distinct cells bearing neuroendocrine and mesenchymal markers were required to bring about a more aggressive cancer phenotype (Calbo et al., 2011). One population expressed neuroendocrine markers such as synaptophysin (Syp), asymmetric synthesis of *HO* (Ash1), and neural cell adhesion molecule (Ncam), while another subclone expressed mesenchymal markers and were low in Ncam. *In vitro* co-cultures of both subclones not only resulted in greater proliferation and growth than if singly grown, but also potentiated the metastatic potential of neuroendocrine cell types in an *in vivo* mouse model which was otherwise not

observed when the neuroendocrine cells were grown in the absence of their mesenchymal cell type counterparts.

Theme 2: Mutational order and epistasis determine the consequences of *FBXW7* mutations during colorectal cancer evolution

Phenotypically normal epithelia accumulate driver mutations

It is well known that cells acquire genetic alterations throughout life. Advances in next generation sequencing (NGS) have uncovered the burden of somatic mutations not only in cancer, but more pertinently in phenotypically normal tissues.

The pattern of somatic driver mutations remains highly variable across the spectrum of tissues. Firstly, tissues acquire mutations at different rates. Massively parallel sequencing performed on normal brain frontal cortex and colonic epithelium during childhood, adolescence and adulthood showed that while the frequency of mutations was similar in children, this begun to diverge during adolescence, such that the mutational rate acquired a tissue-specific context by the time of adulthood (childhood: colon vs brain, $1.8 \pm 0.5 \times 10^{-7}$ vs $1.1 \pm 0.3 \times 10^{-7}$, $p > 0.05$; adolescence: $5.5 \pm 1.6 \times 10^{-7}$ vs $2.2 \pm 1.1 \times 10^{-7}$, $p < 0.05$; adult: $1.1 \pm 0.2 \times 10^{-6}$ vs $6.3 \pm 2.3 \times 10^{-7}$, $p < 0.01$) (Hoang et al., 2016). Similar results were observed in a separate study which evaluated the skin, liver, colon and small intestine (Werner and Sottoriva, 2018). Furthermore, while a handful of genes remain recurrently mutated across the range of tissues, such as *ARID1A*, *TP53*, *NOTCH1* and *FAT1*, many genes are unique to that specific tissue type, such as *AXIN2* in the colon, *SMARCA4* in the duodenum, and *KMT2D* in the liver (Li et al., 2021).

Another approach to characterise driver mutations in normal tissues would be to consider mutations occurring both in normal and cancer of that tissue type, and mutations which are only or more prevalent in normal but not in cancer. For example, mutations occurring in phenotypically normal skin show an almost 100% overlap with genes present in cancer. This

is contrasted with mutated genes in the normal spleen and brain, which are rarely present in cancer (García-Nieto, Morrison and Fraser, 2019). Specific examples include *MUC6*, which is present only in normal stomach and not in gastric cancer (Li et al., 2021), or *KRAS* in the endometrium, which is more prevalent in the normal endometrium than in endometrial cancer (Moore et al., 2020). Characterising driver mutations in this approach raises questions regarding a hitherto unknown molecular phenomenon which must account for the decline in prevalence of the mutated gene in the transition from phenotypically normal to cancer, and therefore could yield promising implications for early cancer initiation.

Taken together, these findings draw importance to studying driver mutations from a tissue-specific approach. Given that a greater period of an individual's life is spent in normalcy rather than in cancer, driver mutations present in normal tissue could be relevant in cancer initiation.

Intestine as a model to study stem cell dynamics

The unique anatomy of the intestinal crypt lends itself to being a suitable model to study early cancer evolution. Four to six stem cells reside at the base of each crypt, where increased Wnt activity and bone morphogenetic protein (BMP) suppression maintain stemness at this intestinal stem cell niche (Mah, Yan and Kuo, 2016). Along the crypt axis towards the luminal surface, stem cells divide and give rise to daughter cells which constitute the diversity of cell types present in the colon. This highly efficient process allows for cells in the crypt to undergo complete renewal every three to ten days. In homeostasis, stem cells compete equally to repopulate the entire crypt in a process of neutral competition known as crypt fixation (Lopez-Garcia et al., 2010, Snippert et al., 2010). Crypt fixation alone is however insufficient to cause CRC. Crypts may outcompete neighbouring crypts through a process of increased crypt fission. Here, stem cell dynamics are recapitulated using organoid models (discussed in a later section).

Under homeostatic conditions, lineage tracing of Lgr5+ crypt based columnar (CBC) cells established these cells as the *bona fide* stem cell in small and large intestine (Barker et al., 2007), capable of regenerating all cells in the intestinal epithelium. In non-homeostatic conditions, such as during colitis, injury, or cancer, a non-Lgr5+ population appears capable of precipitating rapid growth and cellular turnover. Notably, in mouse models, inflammation resulted in the dedifferentiation and proliferation of non-stem cell, non-Lgr5+ cells (Schwitalla et al., 2013, Davis et al., 2015). Furthermore, the ablation of the Lgr5+ population did not prevent recurrence of tumour (de Sousa e Melo et al., 2017). These findings suggest an upending of the traditional stem cell hierarchy during non-homeostatic states, and suggest that cells in the intestinal crypt are able to switch from one differentiation lineage to another depending on the circumstance. For example, Morral et al. identified a subset of cells with elevated expression of POLR1A which is responsible for stemness independent of the Lgr5 status (Morral et al., 2020). Elsewhere, KRT20+ differentiated cells were observed on lineage tracing to revert to an LGR5+ population (Shimokawa et al., 2017). More recently, an equilibrium between classical Lgr5+ CBC cells and an Lgr5- regenerative stem cell (RSC) population has been described. Disequilibrating forces which may originate from either epithelial mutations, such as *KRAS* or *BRAF* mutations, or microenvironmental signalling disruption bias this equilibrium towards an RSC hierarchy which leads to treatment resistance (Vasquez et al., 2022).

There has also been renewed interest in fetal states of intestinal epithelial cells. Guiu et al. demonstrated that fetal intestinal cells were markedly different from adult intestinal cells. Adult cells can be readily demarcated into differentiated and undifferentiated types, in which undifferentiated stem cells found at the base of crypts clearly give rise to the differentiated cell phenotypes. In contrast, fetal intestinal cells uniformly possess the ability to contribute towards intestinal growth and development, and retain the potential to give rise to adult stem cells (Guiu et al., 2019). This plasticity associated with fetal intestinal cells appears to be recalled during instances in which regeneration of the intestine following injury is required. Following

helminthic infection, but also during intestinal irradiation and the ablation of *Lgr5*+ cells, an interferon-gamma (IFN-γ)-dependent transcription program associated with Sca-1 expression was observed (Nusse et al., 2018). Sca-1 expression appeared to be highly characteristic of fetal intestinal cells. Likewise, Yui et al. described a similar association between regenerating intestinal epithelial cells and a fetal cellular state, establishing that collagen type one in the ECM interacts with epithelial cells via FAK and Src, leading to the activation of a YAP/TAZ mediated pathway to induce this fetal state (Yui et al., 2018). Taken together, these findings suggest that adult intestinal cells revert to a fetal-like cell state during injury to bring about repair.

Together, these findings demonstrate the phenomenal plasticity associated with stem cell states in the intestine, such that cellular phenotypes cannot be fully explained by the genotype of those cells. In theme 2, I explore the reordering of stem cell hierarchies and how it may contribute to cancer.

Molecular and functional characteristics of *FBXW7*

While a context specific discussion on the role of *FBXW7* will be provided in the introductory section to both the abovementioned themes, here, a general overview of the role and functions of *FBXW7* is provided.

FBXW7 or F-box and WD repeat domain containing 7 (FBXW7) is the substrate recognition component of the Skp1-Cdc53/Cullin-F-box-protein complex (SCF/β-TrCP) (Yeh, Bellon and Nicot, 2018). This SCF complex is an E3 ubiquitin ligase, and is the final component of a target protein with the ubiquitin-proteosome system (UPS). The SCF complex ubiquitinates the target proteins and triggers proteasomal degradation of that target protein. FBXW7 is one of 4 proteins in this SCF complex, the other components being Skp1, Cul-1 and Rbx1. FBXW7 plays a critical role in the SCF complex, as it is responsible for recognizing proteins which have been phosphorylated at a CDC4 phosphodegron (CPD) motif. This

phosphodegron for FBXW7 recognition is (S/T)PXX(S/T), whereby S/T refers to serine or threonine residues, P refers to proline, and X is any amino acid (Welcker and Clurman, 2008). Briefly, target proteins are phosphorylated by a Glycogen synthase kinase 3 beta (GSK3β) process (Welcker et al., 2004). Following this, FBXW7 recruits the target protein to the SCF complex where the target protein is ubiquitinated. The poly-ubiquitinated protein is marked for degradation by the 26S proteasome (Thompson et al., 2021).

FBXW7 is known to recognise at least 90 different substrates (Arabi et al., 2012). Many of the substrates which FBXW7 targets are proto-oncogenes which are key regulators of gene transcription and stem cell differentiation, such that mutations in *FBXW7* impair the proper degradation of these proteins, and result in unabated cancer cell proliferation, and disordered differentiation. Some of these target proteins include c-Myc, Cyclin E and Notch1. *FBXW7* is the most commonly mutated gene in the SCF complex (Yumimoto and Nakayama, 2020). In CRC, *FBXW7* is associated with both tumour initiation and progression (Karczmarski et al., 2019).

A unique feature of the *FBXW7* gene is that mutations in this gene cannot be easily classified into either gain or loss of function changes. This unique feature is related to the structure of the protein. The FBXW7 protein comprises of 8 WD40 repeat regions which each comprise one propellor blade of an eight-bladed propellor. Arginine residues at the propellor tips bind to the CDC4 phosphodegron motif, and are responsible for holding the protein in place for subsequent ubiquitination. Mutations occurring at these arginine residues lead to severe dysfunction of FBXW7, and therefore are the most common mutations seen in mutant *FBXW7*, which affect arginine residues at positions 465, 479 and 505. Missense mutations at other positions of the gene could affect amino acid residues at the propellor blade, but not specifically the propellor tip, and lead to a less severe protein dysfunction. Analysing the distribution of *FBXW7* mutations based on TCGA data, Davis and Tomlinson illustrated the unique pattern of mutations in *FBXW7* (Davis and Tomlinson, 2012). The authors demonstrated that propellor tip mutations in one allele tended not to be associated with a

classical “second hit” to the other allele, and that two protein-truncating mutations were never found. Together, this implied there was selection pressure towards a non-complete loss of *FBXW7* function, and that mutations occurring at different locations along the gene affected the function of the protein to different degrees which reflected the optimal amount of FBXW7 activity in cancer. Put together, these findings suggested a “just enough” model of carcinogenesis in which there exists an optimal level of dysfunction which is preferable to complete gain or loss of function. This was validated in mouse models of $Fbxw7^{R482Q/+}$, whereby the R482Q is the mouse equivalent of the human R479Q mutation, and $Fbxw7fl^{(\Delta ex5)/+}$, which represented a heterozygous *FBXW7* null model (Davis et al., 2014). The authors demonstrated that *Vill-Cre;*$Fbxw7^{R482Q/+}$ in combination with *APC* mutant mice generated a greater burden of intestinal polyposis than *Vill-Cre;*$Fbxw7fl^{(\Delta ex5)/+}$ and *APC* mutant mice, suggesting that different genotypic perturbations of the same gene could result in different phenotypes depending on the degree of FBXW7 dysfunction.

More recently, *FBXW7* has also been found to regulate chromatin accessibility (Thirimanne et al., 2022). In HCT116 cell lines, *FBXW7* mutations affected Histone H3 lysine-27 acetylation (H3K27ac) and Histone H3 lysine-27 trimethylation (H3K27me3), which represent active versus repressed chromatin respectively. Furthermore, the AP-1 motif, which is bound by the FBXW7 substrate Jun, was enriched in 32% of differential H3K27ac sites upregulated by the *FBXW7* mutation. These findings suggest a synergistic effect, where functional FBXW7 loss upregulates both its substrate as well as regions of the chromatin where those substrates bind to, therefore potentiating the effect of the *FBXW7* mutation.

To summarise, FBXW7 is a key regulatory protein which counts among its substrates many proteins which themselves regulate cell proliferation and differentiation, and in turn can potentiate cancer. In spite of this, *FBXW7* is not a classical tumour suppressor in that it does not satisfy the Knudson two-hit hypothesis for tumour suppressors *in vivo*. The mutational pattern of *FBXW7* affects the degree of functional dysregulation, which is profound, as

evidenced not only in its role in the ubiquitylation and degradation of substrates, but also in modifying chromatin accessibility.

Organoids as models for studying early tumour evolution

Organoids are three-dimensional *in vitro* cell models which recapitulate structures and cell-types seen *in vivo*, and are initiated by either pluripotent stem cells, or organ-restricted adult stem cells (Clevers, 2016). Organoids formed from pluripotent stem cells take advantage of developmental processes to form a myriad of tissues. The research encapsulated in this thesis however relies on adult stem cells, and this will be the focus here. Adult stem cells require a unique cocktail of growth factors in order to maintain the stem cell niche environment necessary for organoid formation and growth (Sato et al., 2011). One of the key molecules crucial to maintenance of this stem cell niche is Wnt, a hydrophobic molecule expressed by Lgr5+ cells and supporting niche cells which limits its spread in the stem cell microenvironment and therefore restricts its effect to a localised area, in essence creating a stem cell niche (Clevers, Loh and Nusse, 2014). The cocktail of growth factors and small molecules necessary for development and maintenance of gastrointestinal organoids was refined by Sato et al. (Sato et al., 2011), and led to the emergence of organoids as a viable model for studying cancer evolution.

Two earlier studies combined the use of clustered regularly interspaced short palindromic repeats (CRISPR) gene editing technology with gastrointestinal organoids to model the acquisition of driver mutations (Drost et al., 2015, Matano et al., 2015). These papers demonstrated that the sequential accumulation of driver mutations *APC*, *TP53*, *SMAD4*, and *KRAS*, resulted in organoids proliferating free of growth factors *in vitro*.

The advantages of human organoid models over the more commonly used cell line models include the ability to introduce genetic mutations *de novo*, particularly when deriving organoids from phenotypically normal adult stem cells. These adult stem cells are regarded to

be wildtype, and rarely harbour any cancer driver mutations. In addition, human organoids give rise to different cell types which may be seen in the tissue of origin, and are therefore less homogeneous than cell lines, which are mostly of epithelial cell origin. Disadvantages of the organoid model system include the increased cost and time associated with the manipulation of organoids, as well as technical difficulty even when applying commonly used molecular biology techniques. Cell lines are used to study the subclonal interaction of *FBXW7* with other subclones in cell lines in theme 1, while organoid models are used to study the role of *FBXW7* in phenotypically normal tissue in theme 2.

Another model which warrants mentioning but is not used in the experiments of this thesis include mouse models. Mouse models bear the advantage of providing a microenvironmental niche which includes stromal and immune cells, and therefore can be used to study the interaction of epithelial cells in the tumour microenvironment. The main disadvantage of the mouse model is that it requires either the xenotransplantation of human cells into immunocompromised mice, or the use of mouse cells. There are however some crucial divergences between humans and mice which could influence the reproducibility of results obtained from mouse models in humans. One study which evaluated gene expression similarities and differences in mouse and humans found that the expression of tumour-related genes were the most divergent. In particular, the PI3K signalling cascade was highly divergent owing to differences in the expression of *mTOR* and *AKT2* (Monaco et al., 2015). More startlingly, there is little similarity between the key tumour suppressor *TP53*. Gene regulatory networks of mouse and human *TP53* have been shown to be species-specific, such that 39% of mouse and 48% of human p53 DNA binding sites are unique to the species, resulting in more than 1000 genes being differently expressed in mice and human (Fischer, 2019, Fischer, 2021).

Hypotheses of this thesis

In this thesis, I explore two themes using mutations in the *FBXW7* gene as a basis to explore ITH. In the first theme, I explore the impact a subclonal population may have on neighbouring cells within an established cancer. In the second theme, I evaluate driver mutations present in phenotypically normal tissue, and seek to explore whether these driver mutations contribute to the generation of ITH.

My hypothesis for each theme is as follows.

Theme 1:

1. *FBXW7* mutations in a subclonal population can have cooperative or competitive effects on neighbouring populations

Theme 2:

1. *FBXW7* in phenotypically normal tissue plays a protective role in the generation of ITH

Conclusion

CRC exerts a global burden of disease. Given that driver mutations exert the greatest selection advantage in cells of the colorectum, research thus far has understandably focused on interactions between these driver mutations. There is however great impetus for novel approaches in the diagnosis and treatment of CRC. CRC has been shown to have significant ITH, which could be harnessed to uncover new treatment approaches which could change how CRC is diagnosed and managed today. This thesis therefore studies two facets of ITH, both involving the gene *FBXW7*. In the first theme, I study the interaction between a clonal population of mutant *FBXW7* with neighbouring wildtype cells. In the second theme, I focus on *FBXW7* mutations which are present in phenotypically normal epithelium, seeking to understand the relationship between *FBXW7* mutations in normal tissue and how this could give rise to future carcinogenesis and ITH. While the first theme uses cell lines as the model,

the second theme combines CRISPR-Cas9 gene editing with adult stem-cell derived human organoids. This novel approach lends itself to studying the second theme as organoids derived from such phenotypically normal organoids are considered to be wildtype for driver mutations, and therefore allows for gene editing to uncover the effect of individual and sequential gene mutations. This thesis therefore seeks to uncover how *FBXW7* contributes to carcinogenesis at different evolutionary time points of a cancer.

Chapter 2: Methods for Theme 1 experiments

The following methodologies cover techniques utilised in the first theme of this thesis. All experiments were performed by myself unless otherwise specified.

Cell culture

The RKO and LS513 cell lines, purchased from ATCC(USA), were used in this project. Cells were cultured in DMEM/F12 (Gibco) supplemented with 10% FBS (Gibco) and 100U/ml of penicillin-streptomycin (Gibco). When cells had reached 80% confluency, the cells were first washed with Dulbecco's Phosphate Buffered Saline (DPBS, Gibco), before incubating with TrypLe (Gibco) at 37°C for five min. Cells were collected by centrifugation at 1200 rpm for 5min, before being subcultured into new flasks. Cells were cultured in a humidified 5% CO_2 chamber at 37°C. Mycoplasma testing was carried out regularly using PlasmoTest Mycoplasma Detection Kit (InvivoGen).

Biological replicates referred to in this Theme were replicates different isogenic clones selected following CRISPR-Cas9 gene editing of the wildtype cell lines. Each individual clone was validated using Sanger sequencing and western blotting as described in the following sections.

Generation of clonal CRISPR-Cas9 gene knockouts

sgRNAs were designed using Synthego's CRISPR Design Tool (Table 1). This tool prioritises sgRNA with minimal off-target effects and early exonic locations. For *FBXW7*, the sgRNA targets Y509. This was designed to coincide as closely to the *FBXW7* hotspot R505C as possible, considering limitations imposed by the need for an NGG protospacer adjacent motif (PAM) complex. For *AKAP8*, the recommended sgRNA based on the tool was used. For

AAVS1, a commonly used safe habour locus was used. AAVS1 experiments were performed as this sgRNA targets an intronic location of the *PPP1R12C* gene, and therefore should not have an protein-altering effect. AAVS1 therefore served as a control to ensure studied effects were not a function of CRISPR gene editing.

Gene target	sgRNA sequence
FBXW7	GUCCGCUGUGUUCAAUAUGA
AAVS1 (*PPP1R12C*)	GGGGCCACUAGGGACAGGAU
AKAP8	CUCGAACGGCUGGAAGCGGA

Table 1: sgRNA sequences for CRISPR-Cas9 gene editing used in Theme 1

An electroporation-based approach was used to maximise gene editing efficacy. Cells were first dissociated into a single-cell suspension using TrypLe (Gibco) by incubating at 37°C for 5min. A ribonucleoprotein (RNP) complex was generated by mixing Cas9 enzyme (Sigma-Aldrich), and sgRNA to generate a molar ratio of 1:1. Cells were mixed with the RNP complex, and then electroporated with a Neon Transfection System (ThermoFisher). The settings for both cell lines were the same. This was voltage 1300V, width 20ms and two pulses. Following electroporation, cells were incubated for one week and allowed to expand. Primers which generated amplicons of approximately 500bp in length were designed on Benchling (Table 2). Amplicons were Sanger sequenced by an external vendor (Genewiz).

Gene target	Forward primer	Reverse primer
FBXW7	TCTGCAGAGTTGTTAGCGGT	TGGACTGTACTGGATCAGCAA
AAVS1 (*PPP1R12C*)	GGTCCGAGAGCTCAGCTAGT	GGCTCCATCGTAAGCAAACC
AKAP8	AGTGTTCCAGGGAGGGTGTT	GGGTCTCGGCATTCACTGTA

Table 2: Sequencing primers to validate gene knockout in Theme 1

Serial dilution was performed by diluting cells to a concentration of one cell/20µl and 20µl of this cell suspension was transferred to each well of a 96-well plate. Cells were allowed to expand to achieve confluency of a T25 flask. Amplification of the target region was again performed as described above and Sanger sequencing was used to check that isogenic biallelic knockouts were obtained. Further validation of knockout was performed using western blot.

Protein lysates and western blot

Proteins were harvested using RIPA (ThermoFisher) supplemented with both protease and phosphatase inhibitor cocktails (Sigma-Aldrich). Protein concentrations were quantified using bicinchoninic acid (BCA) assay (ThermoFisher).

For western blot, 30µg of protein was used for each lane. Proteins were separated using a NuPage 4 – 12% Bis-Tris gel (Invitrogen) and transferred onto a 0.45µm nitrocellulose membrane. Membranes were blocked with 5% low-fat milk for one hour before overnight incubation with primary antibodies at 4°C. Antibodies used can be found in table 3. GAPDH was used as the control. Membranes were incubated with secondary antibody at room temperature for one hour. The membranes were imaged with a ChemiDoc XRS+ system (Bio-Rad).

Antibody	Host	Dilution	Category number	Manufacturer
Anti-FBXW7	Rabbit	1:2500	BS-8394R	Bioss, USA
Anti-ATP5A	Rabbit	1:2500	ab176569	Abcam, USA
Anti-HSP60	Rabbit	1:2500	ab46798	Abcam, USA
Anti-AKAP95	Rabbit	1:2500	ab134923	Abcam, USA
Anti-AKAP95	Rabbit	1:2500	ab72196	Abcam, USA
Anti-GAPDH	Rabbit	1:2500	ab9485	Abcam, USA
Anti-Rabbit	Goat	1:5000	ab6721	Abcam, USA

Table 3: Antibodies and dilutions used for western blotting in Theme 1

Quantitative reverse transcription-PCR (RT-PCR)

Total RNA was extracted using an RNeasy Mini Kit (Qiagen) as per manufacturer's protocol. cDNA was synthesised using SuperScript IV VILO Master Mix (Invitrogen). Real-time qPCR was then performed using TaqMan Fast Advanced Master Mix (Applied Biosystems). Taqman probes used can be found in Table 4.

Gene	Assay ID	Manufacturer
MKI67	Hs04260396_g1	Applied Biosystems, USA
PCNA	Hs00427214_g1	Applied Biosystems, USA

Table 4: Taqman probes used for RT-PCR

Live-cell microscopy

The Incucyte (Sartorius) imaging system was used for live-cell microscopy. Cells were seeded at a density of 5 x 10^4 cells/ml and observed at 10X magnification every two hours for 72 hours.

Flow cytometry

Flow cytometry was used to determine the proportion of actively proliferating cells, as well as to ascertain cell apoptosis. All flow cytometry experiments were performed on a CytoFLEX (Beckman Coulter) flow cytometer. Analysis was performed with the CytExpert 2.4 software (Beckman Coulter).

For proliferating cells, a Click-iT™ EdU Alexa Fluor™ 647 Flow Cytometry Assay Kit (Invitrogen) was used. Incorporation of EdU with cells was performed at an EdU concentration of 10 μM for a 2 hour duration. The Click-iT reaction was them performed after cells had been fixed and permeabilised. Propidium iodide was added 10min prior to flow cytometry as a live/dead marker.

For cellular apoptosis, an Annexin V Apoptosis Kit with 7-AAD, FITC (Stemcell Technologies) was used. Cells were harvested and counted, and 1 x 10^6 cells were stained with Annexin V Binding Buffer. Annexin V conjugated to FITC and 7-AAD were added to cells for a duration of 15min, and left at room temperature protected from light. This was then proceeded by flow cytometry.

Extreme limiting dilution colony forming assay

Cells were seeded at the following densities five, 25, 50, 100, 200, 300, 400 and 500 cells/ml and incubated for 72 hours at 37°C. Colonies were visualised under a tabletop microscope, counted manually by eye, and analysed using the webtool interface http://bioinf.wehi.edu.au/software/elda/ (Hu and Smyth, 2009).

Metabolic assay

The Agilent Seahorse Real-Time ATP Rate Assay kit was used in accordance with the manufacturer's instructions, and the Seahorse Agilent XFe96 Analyzer was used. Briefly, a range of cell densities of both wildtype and *FBXW7*$^{-/-}$ cells were seeded on the Seahorse XF microplate. A total of 8 different cell densities at increments of 0.5 x 10^4 cells/well from 0.5 x 10^4 cells/well to 4 x 10^4 cells/well were seeded 24 hours before the assay was run to ensure good adherence of cells to the microplate. The sensor cartridge was hydrated at 37°C in a non-CO_2 incubator 24 hours before. On the day of the experiment, supplementation of Seahorse XF DMEM, pH 7.4 with 10mM of XF glucose, 1mM of XF pyruvate, and 2mM of XF glutamine was performed, and this media was allowed to warm to 37°C. Media from the microplate with adherent cells were removed, and the freshly prepared supplemented Seahorse XF DMEM was added. The cells were further incubated for another 60min at 37°C in a non-CO_2 incubator. To determine the oxygen consumption rate (OCR) and extracellular acidification rate (ECAR), 1.5µM of oligomycin and 0.5µM of Rotenone/Antimycin A was added to the cells 20min and 40min after the analyser was initiated respectively. Metabolic measurements were performed at intervals of 6.5min. At each cell density seeded, mean values of the OCR and ECAR between wildtype and *FBXW7*$^{-/-}$ cells were calculated.

Apoptosis assays

Three different apoptosis assays were performed. These were the quantification of cleaved Caspase-3, Cathepsin B, and Calpain activity. All measurements were recorded on a POLARstar Omega microplate reader (BMG Labtech).

For quantification of cleaved Caspase-3, instructions from the Caspase 3 (Cleaved) Human ELISA Kit (Invitrogen) was followed. Briefly, 5 x 10^6 cells were harvested and resuspended in cell lysis buffer (Invitrogen) supplemented with protease inhibitor cocktail (Sigma-Aldrich)

and PMSF protease inhibitor (Thermo Scientific). This suspension was then centrifuged, and the supernatant was collected and diluted 1:10, before being transferred to an antibody coated plate. A detector antibody followed by an anti-rabbit IgG HRP solution was then added and allowed to incubate for 30min. Stabilised chromogen was then added and stop solution was transferred to the wells after 30min. Immediately after, absorbance at 450nm was measured on a plate reader. All steps were performed at room temperature.

For quantification of Cathepsin B, instructions from the Human Cathepsin B ELISA Kit (Abcam) was followed. Briefly, 5 x 10^6 cells were harvested and resuspended in cell lysis buffer (Invitrogen) supplemented with protease inhibitor cocktail (Sigma-Aldrich) and PMSF protease inhibitor (Thermo Scientific). This suspension was then centrifuged, and the supernatant was collected and diluted 1:10, before being transferred to an antibody coated plate. The biotinylated anti-human Cathepsin B antibody was then added to each well and incubated for 90min. After washing, an avidin-biotin-peroxidase complex working solution was added to wells for 40min. After washing, TMB colour developing agent was then added for 30min, and the reaction arrested with the TMB stop solution. Immediately after, absorbance at 450nm was measured on a plate reader. All steps were performed at room temperature.

For measurement of calpain activity, instructions from the Calpain Activity Assay Kit (Abcam) was followed. Briefly, 1 x 10^6 cells were harvested and resuspended in extraction buffer on ice for 20min. Protein concentration was then measured and 100µg of protein was used for each assay reaction. Calpain substrate was added to each well and incubated for 60min at 37°C. Immediately after, fluorescence following excitation at 400nm was measured on a plate reader.

RNA sequencing and analysis

Extraction of total RNA was performed using the RNeasy Mini Kit (Qiagen) as per manufacturer's protocol. The quantity and quality of the RNA was ascertained using the Agilent 2100 Bioanalyzer. Sequencing was performed by the Oxford Genomics Centre, Oxford using an Illumina NextSeq500 instrument using the standard paired-end protocol with a read length of 150bp. Fastq reads were processed to clip low quality leading and trailing edges and remove any adapter content using Cutadapt (v.3.5). Quality checked fastq reads were then aligned to the human genome (GRCh38) and gene annotation (Ensembl release 105) using STAR aligner (v.2.7.3a) two-pass mode to generate gene-level quantification. Raw counts were then processed in R (v.4.1.2) for all statistical testing and plotting purposes. Normalisation and differential expression were performed using limma (v.3.50.0), and pathway level significance testing was performed using fgsea (v.1.20.0) on MSigDb (Molecular Signature Database, v.7.5.1) hallmark pathways. Subsequent bioinformatic analysis was performed by Dr Amit Mandal, Post-Doctoral Fellow in Bioinformatics, Buczacki Lab, who returned a differentially expressed gene list between cells which were wildtype for *FBXW7* and mutant for *FBXW7*. Subsequent analysis was performed by myself. This included the generation of heatmaps using gplots (v.3.1.1) and volcano plots using ggplot2 (v3.3.5). Analysis was also performed using Gene Set Enrichment Analysis (GSEA) for 'Hallmark' gene sets (50 gene sets), obtained from the molecular signatures database (MSigDB v4.0) (Liberzon et al., 2015).

Transwell cell culture

Transwell inserts were used for co-culture of cell lines. All experiments utilised the 24mm diameter Transwell inserts, made of polyester and with a pore size of 0.4μm. These were placed in 6-well plates. A cell density of 3 x 10^6 cells/insert surface were cultured. Briefly, cells were first seeded onto the underside of the insert, and placed in a 37°C incubator for one

hour to allow cells to adhere. Following this, the inserts were flipped and cells were seeded onto the topside of the insert. A total of 1ml of media was placed into the insert, while 2ml of media was placed in the receiver. All Transwell cultures were incubated in a humidified 5% CO_2 chamber at 37°C.

Immunofluorescence and imaging

An 8-well μ-Slide (ibidi) was used for all immunofluorescence (IF) experiments. Cells were seeded at a density of 3 x 10^4 cells/well. Cells were initially fixed and permeabilised, before blocking was performed with 2% BSA in PBS for one hour at room temperature. Cells were washed with PBS before primary antibody was added and incubated for 2 hours at room temperature. Cells were washed again with PBS and secondary antibody was added and allowed to incubate for 45min at room temperature. All imaging was performed on an the Andor Dragonfly High Speed Confocal Microscope Systems (Oxford Instruments). Antibodies and concentrations used for IF are listed in Table 5. Quantification of fluorescence intensity was done using ImageJ while a CellProfiler4 pipeline for speckle counting using CellProfiler4 was used to quantify DNA damage foci of γH2AX and 53BP1 (Stirling et al., 2021).

Antibody	Host	Dilution	Category number	Manufacturer
Anti-HSP60	Rabbit	1:1000	ab46798	Abcam
Anti-ATP5A1	Mouse	1:1000	459240	Invitrogen
Anti- γH2AX	Rabbit	1:2500	ab81299	Abcam
Anti-53BP1	Rabbit	1:2500	ab172580	Abcam
Phalloidin 647	-	1:1000	ab176759	Abcam
Hoescht 33342	-	1µg/ml	62249	Thermo Scientific
Anti-Rabbit	Goat	1:1000	ab150077	Abcam
Anti-mouse	Goat	1:1000	ab150113	Abcam

Table 5: Antibodies and dilutions used for immunofluorescence in Theme 1

Single-cell gel electrophoresis assay

CometAssay (biotechne) was used for single-cell electrophoresis, and all steps were in compliance with the manufacturer's instructions. Briefly, 50µl of cells at a concentration of 1 x 10^5 cells/ml were placed on a CometSlide (biotechne) and allowed to spread evenly across the slide's marked sample area. The slide was then incubated in the dark at 4°C for 10min. Slides were immersed in lysis solution for 60min at 4°C, and subsequently transferred to alkaline unwinding solution for 20min in the dark at room temperature. Slides were then transferred to an electrophoresis slide tray and submerged in alkaline electrophoresis solution which had been pre-cooled to 4°C. Power settings were 21 volts, and this was allowed to run for 30min. Following electrophoresis, slides were washed by gentle immersion in distilled water twice, and finally in 70% ethanol for 5min. Samples were allowed to dry at 37°C for 15min. DNA was visualised using 100µl of 1X SYBR Gold infiltrated over the dried agarose. Staining was completed in 30min and the slide was washed in distilled water again. Fluorescence microscopy was immediately performed. Quantification of comet assay parameters was performed using the comet assay pipeline on CellProfiler4. (Stirling et al., 2021). Each replicate was used to generate a single slide of cells, and approximately 100 cells

per slides were analysed. Results from biological replicates of each condition were pooled together for the final quantification analysis.

Proteomics analysis

Two sources of proteins were analysed by mass spectrometry. These were obtained firstly from the cell lysates, and secondly from the media in which the cells were grown in. Proteins derived from cell lysates were harvested using RIPA (ThermoFisher) supplemented with protease and phosphatase inhibitor cocktails (Sigma-Aldrich). BCA assay (ThermoFisher) was used for the quantification of protein concentration. Proteins from the media fraction were obtained from media without FBS supplementation in which cells had been growing for four days. Media was aspirated from cell culture flasks and transferred into a centrifuge tube. This was subjected to centrifugation at 1200 rpm to remove cell debris. Supernatant was then transferred into a clean tube.

Proteomic mass spectrometry was performed by our collaborators at the Target Discovery Institute, University of Oxford. Filter aided sample preparation (FASP) was performed. Briefly 20ug of lysate and 2.5ml of media were loaded onto a Vivacon 500 (Sartorius) filter and spun at 14, 300 rcf for 10min. Samples were denatured with 8M urea for 30min at room temperature followed by reduction with 10mM Tris(2-carboxyethyl)phosphine(TCEP) and 50mM Iodacetamide for 30min at room temperature. Two cycles of washes with 200ul 50mM TEAB were applied to reduce the 8M urea to less than 1M prior to trypsin addition. 2ug of trypsin in 200ul 50mM TEAB was added and incubated overnight at 37°C. The next day, peptides were eluted from the Vivacon 500 filter with 200ul 0.1% trifluoroacetic acid (TFA)/water and 200ul 50% acetonitrile (ACN)/water. Eluate was then dried in a SpeedVac Vacuum Concentrator (ThermoFisher).

For mass spectrometry, peptides were resuspended in 5% formic acid and 5% DMSO and then trapped on an Acclaim™ PepMap™ 100 C18 HPLC Columns (PepMapC18; 300µm x

5mm, 5µm particle size, ThermoFisher) using solvent A (0.1% Formic Acid in water) at a pressure of 60 bar and separated on an Ultimate 3000 UHPLC system (ThermoFisher) coupled to a QExactive mass spectrometer (ThermoFisher). The peptides were separated on an Easy Spray PepMap RSLC column (75µm i.d. x 2µm x 50mm, 100 Å, ThermoFisher) and then electro-sprayed directly into an QExactive mass spectrometer (Thermo Fischer Scientific) through an EASY-Spray nano-electrospray ion source (ThermoFisher) using a linear gradient (length: 60 minutes, 5% to 35% solvent B (0.1% formic acid in acetonitrile and 5% dimethyl sulfoxide), flow rate: 250 nL/min). The raw data was acquired in the mass spectrometer in a data-independent mode (DIA). Full scan MS spectra were acquired in the Orbitrap (inclusion list with scan range 495 to 995 m/z, 20m/z increments, with an overlap of +/- 2 Daltons, resolution 35000, AGC target 3e6, maximum injection time 55ms). After the MS scans peaks were selected for HCD fragmentation at 28% of normalised collision energy(nce)/stepped nce. HCD spectra were also acquired in the Orbitrap (resolution 17500, AGC target 1e6, isolation window 20m/z).

For data analysis, the search engine used was DIA-NN 1.8 in a library-free search mode. For protein identification, peaks were searched against the database UPR_Homo sapiens_9606_UP000005640 with the following parameters – deep learning-based spectra, RTs and IMs prediction, maximum number of protease (trypsin) missed cleavages set to 1, post-translational modifications: Carbamidomethyl (C), Oxidation (M), false discovery rate (FDR) set to 1%, and match between runs was used.

Metabolomics analysis

The entire metabolomic mass spectrometry was performed by me, from extraction of metabolites to initial bioinformatics, and final analysis of results. I was supervised by experts from the Target Discovery Institute. Metabolomic analysis was performed by GCxGC-MS as

described (Yu et al., 2017). Metabolites were analysed from the media in which cells were grown in. The protocol to harvest media was the same as used for proteomic analysis.

Briefly, extraction of metabolites was carried out at room temperature by adding 1 mL of tert-butyl methyl ether (MTBE) and 800 µL methanol to 100 µL of media sample. Myristic acid-14,14,14-D3 was used as an internal standard. Homogenisation was performed using a bead beater (Precellys 24, Bertin Technologies). The supernatant of MTBE and methanol were combined and dried using a Speed Vac Centrifugation (ThermoFisher).

The dried samples were then resuspended in 50µl of methoxyamine hydrochloride in pyridine (20 µg/µL) and shaken at a shaking speed of 1200 rpm for 90 min at 30°C. 70 µL N-Methyl-N-trimethylsilyltrifluoroacetamide (MSTFA) with 1% chlorotrimethylsilane (TMCS) and 30 µL pyridine were added to the samples, followed by incubation for one hour at 60 °C and shaken at 1200 rpm. The samples were cooled to room temperature and injected directly for GCxGC-MS analysis.

Analysis of derivatised samples was performed using a GCxGC-MS system comprising of a gas chromatograph coupled to a quadrupole mass spectrometer (Shimadzu GCMS QP2010 Ultra) and a Shimadzu AOC-20i/s auto sampler. The first dimensional separation was carried out on a SHM5MS capillary column (30 m×0.25 mm i.d.×0.25 µm film thickness, Shimadzu) while the second dimensional separation was on a BPX-50 capillary column (5 m×0.15 mmi.d.×0.15 µm film thickness, SGE). The carrier gas used was Helium gas. 4 seconds was set as the modulation period. The samples were injected at 280°C. The oven temperature was programmed from 60 °C to 320 °C at 10 °C/min unless stated otherwise and held at 320°C for 8 min. The interface temperature to the mass spectrometer was set at 300°C and ion source was heated at 230°C. The mass spectrometer was operated at scan speeds at 20 000 amu covering a range of m/z 45–600. Electron Ionization spectra were recorded at 70 eV.

Raw GCxGC MS data were processed using GCMSsolution software (v2.72/4.20 Shimadzu), and Chromsquare software (v2.1.6, Shimadzu) in combination with the NIST 11, NIST 11/s,

OA_TMS, FA_ME and YUTDI in-house libraries. Annotation of metabolites was carried out both by comparing them to external standards (IM spectra and retention times adjusted to the internal standard myristic acid-14,14,14-d3) and also by spectrum matching based searches using the above databases for metabolites which did not have external standards. The similarity score threshold was set to 80 out of 100, and the confidence of identification further validated by manual inspection of matches between experimentally observed and reference EI spectra.

Transient transfection of plasmid for AKAP8 overexpression

pcDNA3.1(+)-AKAP95-FRB was a gift from Jin Zhang (Addgene plasmid # 138254 ; http://n2t.net/addgene:138254 ; RRID:Addgene_138254) (Clister et al., 2019). pcDNA3.1(+) empty vector was a gift from David Kerr. Plasmids were extracted using PureYield™ Plasmid Miniprep System (Promega). Lipofectamine 3000 (Invitrogen) was used for the transient transfection of cells. Validation of successful transfection was performed by western blot 48 hours after transfection.

Validation of results from datasets

Gene expression from the TCGA-COAD dataset was used to validate my *FBXW7*$^{-/-}$ model. RNAseq results from CRC cases which were wildtype and mutant for *FBXW7* were compared, generating a gene set of upregulated genes. This gene set was denoted as the "TCGA_FBXW7_Mutant" gene set and was used subsequently. Using transcriptional data obtained from RNAseq results from a comparison of *FBXW7*$^{-/-}$ and wildtype cells from my model, I performed GSEA using the "TCGA_FBXW7_Mutant" gene set to generate enrichment score values as well as the FDR q-value.

Statistical analyses

All statistical analysis was performed using GraphPad Prism 9. Data are presented as mean ± standard deviation (sd). p-values < 0.05 were considered significant. Significance was further defined as follows: *p<0.05,**p<0.01, ***p<0.001, ****p<0.0001. Two tailed unpaired students' t-test was used for comparisons between groups with parametric data. All experiments were performed with biological triplicates taken from distinct samples.

Chapter 3: Methods for Theme 2 experiments

The following methodologies cover techniques used in the second theme of this thesis. All experiments were performed by myself unless otherwise specified.

Human material for organoid cultures

Ethics approval for the retrieval of human colon specimens was accorded by the University of Oxford – Translation Gastroenterology Unit (16/YH/0247). Written consent was obtained from all patients.

Organoid culture

Colonic tissue was obtained from adult patients who were undergoing surgery for colorectal cancer. A 1 x 1 cm piece of colon from a region at least 5cm distant from the tumour, and appearing phenotypically normal was resected. Organoids were derived based on my previously published protocol (Chan, Collins and Buczacki, 2023). Briefly, colon specimens are first prepared by dissecting away omental or mesenteric fat, if present. Specimens were washed twice in DPBS with 10% FBS to remove fecal debris and other contaminants. EDTA harvesting solution comprising of Hanks' Balanced Salt Solution (HBSS) supplemented with 12mM EDTA and 10mM NaOH was prepared. Specimens were then placed into EDTA harvesting solution and incubated in a 37°C water bath. The specimens were agitated vigorously by manual shaking every 2 min for a total duration of 10min. Specimens were then centrifuged at 400g for 6 min to allow crypts to pellet. The supernatant was aspirated and resuspended in 5ml of DPBS/FBS solution. A 10µl aliquot of the crypt suspension was placed onto a petri dish and the number of crypts present in the aliquot was counted using a light microscope. This was then used to determine the total number of crypts which had been

harvested. Depending on the number of wells anticipated for organoid culture, a volume of the crypt suspension equivalent to 350 crypts per well was transferred to a new centrifuge tube. Centrifugation was repeated to pallet the crypts. This was then resuspended in reduced growth factor BME, Type 2 Pathclear (Cultrex), and the BME-crypt suspension was then transferred to 24-well plates which were placed in a 37°C incubator for 30min to allow for polymerisation of the BME. Growth media was added when BME had polymerised.

Media was changed once every 2 – 3 days, and subcultured after 7 days. To subculture organoids, media is aspirated from each well, and the domes are disrupted with Organoid Harvesting Solution (Cultrex) on ice for 30 min. Organoids are then aspirated and transferred to a centrifuge tube, where organoids are palleted by centrifugation at 400g for 6 min. Organoids are then mechanically disrupted with a P200 pipette, and centrifuged again at 400g for 6 min. Organoids are then resuspended in BME as described previously, and growth media is replaced when BME has polymerised.

The constituents of organoid culture growth media can be found in Table 6. Wnt3a conditioned media was produced from ATCC CRL-2647 cell line, while R-spondin conditioned media was produced from Cultrex HA-R-Spondin1-Fc 293T cell line.

Reagent	Concentration	Manufacturer
Advanced DMEM/F12	-	Gibco
Penicillin/Streptomycin	100U/ml	Gibco
HEPES buffer solution	10mM	Gibco
GlutaMAX	2mM	Gibco
Wnt3a conditioned medium	50% (v/v)	-
R-spondin conditioned medium	25% (v/v)	-
B-27 plus supplement	1X	Gibco
SB202190	10μM	Tocris
SB431542	0.5μM	Tocris
Prostaglandin E2	1μM	Tocris
Noggin	50ng/ml	Peprotech
EGF	50ng/ml	Peprotech
Nicotinamide	10mM	Sigma-Aldrich
N-acetylcysteine	1.25mM	Sigma-Aldrich

Table 6: Reagents used to produce human organoid growth media

For selection of *APC* mutants, Wnt3a and R-spondin conditioned media were withdrawn from the growth media for at least 3 weeks. *APC* mutant cells are not able to form the *APC* destruction complex, allowing for β-catenin to avoid phosphorylation and to enter the nucleus. On the other hand, wildtype cells have a functioning *APC* destruction complex. In the absence of Wnt3a, the destruction complex becomes unbound from the LRP5/6 and Frizzled transmembrane proteins, resulting in the phosphorylation and ubiquitin-dependent degradation of β-catenin. ZNRF3/RNF43 are negative regulators of the Wnt pathway through their effect of degrading the LRP5/6 and Frizzled transmembrane proteins (Ter Steege and Bakker, 2021). R-spondins potentiate the effect of the Wnt pathway by clearing ZNRF3/RNF43 from the membrane. In the absence of R-spondin, ZNRF3/RNF43 are not cleared from the cellular membrane, resulting in a further dampening of the Wnt pathway.

For selection of *TP53* mutants, growth media was supplemented with 10μM of Nutlin-3 for at least 1 month (Sigma-Aldrich). MDM2 is an inhibitor of p53. Nutlin-3 is an inhibitor of

MDM2. In *TP53* wildtype cells, the addition of Nutlin-3 results in an increase in p53 activity, leading to an increase in apoptosis and cell death. This increased apoptosis would however not be observed in mutant *TP53* cells.

I attempted to add reagents to select for *FBXW7* mutants but as this was not successful, I diverted my effort towards improving the efficacy of CRISPR/Cas9 gene editing instead.

Finally, biological replicates referred to in this Theme were replicates derived from the same biological specimen but differing only in their passage number. I was unfortunately not able to generate replicates from different patient samples because of the lengthy time as well as the high quantity of resources required first to generate the organoids, and then to perform CRISPR-Cas9 gene editing on each of the samples.

CRISPR/Cas9 gene editing of human organoids

sgRNAs were designed using Synthego's CRISPR Design Tool (Table 7). For *FBXW7*, a multiguide approach utilising three different guide RNAs in equal concentration was used. The multiguide targets Exon 2 of the *FBXW7* gene. This increased the targeting efficacy given the lack of a selecting agent which can be used to enrich for *FBXW7* mutants. The AAVS1 site is a commonly used safe harbour locus which served as a control to ensure studied effects were not a function of CRISPR gene editing. Table 7 lists the guide RNA sequences which were used for the different gene editing targets.

Gene target	Guide RNA sequences
FBXW7	GCAAGGAAUGGUGAAGUUGU GAUGAAUCGUGUGGUAGAGG AGCAAAAGACGACGAACUGG
APC	UCUGUAUAAAUGGCUCAUCG
TP53	CCGGUUCAUGCCGCCCAUGC
AAVS1(*PPP1R12C)*	GGGGCCACUAGGGACAGGAU

Table 7: sgRNA sequences for CRISPR-Cas9 gene editing used in Theme 2

Our protocol for gene editing human organoids has been described previously (Chan, Collins and Buczacki, 2023). An electroporation-based approach was used to maximise gene editing efficacy. Cells were first dissociated into a single-cell suspension using TrypLe (Gibco) by incubating in a 37°C water bath for 10min. triturating at regular intervals. Organoids were then pelleted with centrifugation at 400g for 6 min. Pelleted organoids were resuspended in advanced DMEM, triturated and palleted by centrifugation once more. A ribonucleoprotein (RNP) complex comprising Cas9 (Sigma-Aldrich) and guide RNA in a molar ratio of 1:4 was resuspended in Buffer R (Invitrogen) and left to stand at room temperature for 10min. Pelleted organoids were resuspended in Buffer T (Invitrogen). The RNP complex and resuspended organoids were then mixed. Electroporation was performed using a Neon Transfection System (ThermoFisher) with the following settings – voltage 1350V, width 20ms, and pulses 2. After electroporation, organoids are transferred into advanced DMEM and left to recover for 10min. Centrifugation was performed to pellet organoids, which were then resuspended in BME and transferred to cell culture plates as described above. Following electroporation, organoids were incubated for one week and allowed to expand. Primers which generated amplicons of approximately 500bp in length were designed on Benchling (Table 8). Amplicons were Sanger sequenced by an external vendor (Genewiz).

Gene target	Forward primer	Reverse primer
FBXW7	TTTTCCCCTGCAGAATGTGA	ACTGGAGTTCGTGACACTGTT
APC	ATGCTGCAGTTCAGAGGGTC	TTTTTCTGCCTCTTTCTCTTGGT
TP53	CTTGCCACAGGTCTCCCCAAG	AGCCACAGGTTAAGAGGTCC
AAVS1 (*PPP1R12C*)	GGTCCGAGAGCTCAGCTAGT	GGCTCCATCGTAAGCAAACC

Table 8: Sequencing primers to validate gene knockout in Theme 2

Protein lysates and western blot

Proteins were harvested using RIPA (ThermoFisher) supplemented with both protease and phosphatase inhibitor cocktails (Sigma-Aldrich). Protein concentrations were quantified using bicinchoninic acid (BCA) assay (ThermoFisher).

For western blot, 30µg of protein was used for each lane. Proteins were separated using a NuPage 4 – 12% Bis-Tris gel (Invitrogen) and transferred onto a 0.45µm nitrocellulose membrane. For non-phosphorylated proteins, membranes were blocked with 5% low-fat milk for 1 hour before overnight incubation with primary antibodies at 4°C. For phosphorylated proteins, membranes were blocked with 5% BSA for 1 hour before overnight incubation with primary antibodies at 4°C. Antibodies used can be found in table 9. GAPDH was used as the control. Membranes were incubated with secondary antibody at room temperature for 1 hour. The membranes were imaged with a ChemiDoc XRS+ system (Bio-Rad).

Antibody	Host	Dilution	Category number	Manufacturer
Anti-FBXW7	Rabbit	1:2500	BS-8394R	Bioss
Anti-pcJUN (T239)	Rabbit	1:2500	PA5-40193	Invitrogen
Anti-cJUN	Rabbit	1:2500	ab40766	Abcam
Anti-pCCNE1 (T395)	Rabbit	1:2500	ab52195	Abcam
Anti-CCNE1	Rabbit	1:2500	ab33911	Abcam
Anti-pcMYC (T58)	Rabbit	1:2500	ab185655	Abcam
Anti-pcMYC (S62)	Rabbit	1:2500	ab185656	Abcam
Anti-cMYC	Rabbit	1:2500	ab32072	Abcam
Anti-NICD	Rabbit	1:2500	4147	Cell signalling technologies
Anti-NOTCH1	Rabbit	1:2500	ab52627	Abcam
Anti-GAPDH	Rabbit	1:2500	ab9485	Abcam
Anti-Rabbit	Goat	1:5000	ab6721	Abcam

Table 9: Antibodies and dilutions used for western blotting in Theme 2

Immunofluorescence and imaging

IF for organoids was previously described in our protocol (Chan, Collins and Buczacki, 2023). Briefly, media from the wells of organoids was removed and fixation with 2% paraformaldehyde was performed. This was incubated at room temperature for 30min. Organoids were then washed with ice-cold DPBS gently, and allowed to settle by gravity. The supernatant was then removed, and blocked with Organoid Blocking Solution comprising DPBS supplemented with 3% BSA, 1% Triton X-100, 1% saponin and 1% secondary antibody animal serum. Blocking was performed for 3 hours at room temperature. Organoids were again allowed to settle by gravity, and the supernatant was removed. Primary antibody diluted in organoid blocking solution was then added to the organoids for 24 hours at 4°C. Organoids were then washed with DPBS supplemented with 3% BSA, 1% Triton X-100, and 1% saponin. Secondary antibody diluted in the organoid washing solution was then added to the organoids for 2 hours at 4°C. After 2 hours, organoids were again washed with organoid

washing solution. Phalloidin stain was added to the organoids for 60min, before DAPI stain was added to the organoids for 5min. Imaging was performed with the Andor Dragonfly High-Speed Confocal Microscope Systems (Oxford Instruments). All antibodies and stains used for IF is represented in Table 10.

Antibody	Host	Dilution	Category number	Manufacturer
Anti-FBXW7	Mouse	1:100	H00055294-M02	Novus Biologicals
Phalloidin iFluor 647	-	1:400	ab176759	Abcam
DAPI	-	300nM	D1306	Invitrogen
Anti-mouse IgG H&L (Dylight 550)	Goat	1:500	ab96872	Abcam

Table 10: Antibodies and dilutions used for immunofluorescence in Theme 2

Co-culture of organoids

Transduction of organoids was performed to achieve stable expression of GFP or mCherry fluorescence. Lentiviral particles were first generated by transfection of HEK293T cells with pMDLg/pRRE, pRSV-Rev, pSF-CMV-VSVG (Sigma-Aldrich) and either pCDH-EF1-copGFP-T2A-Puro for GFP fluorescence or pCDH-EF1-copGFP-T2A-Puro for mCherry fluorescence. pMDLg/pRRE, and pRSV-Rev were gifts from Didier Trono (Addgene plasmid # 12251; http://n2t.net/addgene:12251 ; RRID:Addgene_12251) (Dull et al., 1998), pCDH-EF1-copGFP-T2A-Puro and pCDH-CMV-mCherry-T2A-Puro were gifts from Kazuhiro Oka (Addgene plasmid # 72263 ; http://n2t.net/addgene:72263 ; RRID:Addgene_72263).

To transduce organoids with the lentivirus particle, the organoids were first dissociated into single-cells by TrypLe for 10min in a 37°C water bath, with trituration at regular intervals. Single-cells were centrifuged at 400g for 6 min to pellet cells, which were then passed through a 50µm filter. The number of cells was then counted using a hemocytometer, and the

required number of cells was aspirated, centrifuged and then pelleted. These cells were then resuspended in organoid growth media, and viral particles were added to achieve a multiplicity of infection (MOI) of 20. Cells were infiltrated over wells which had been overlaid with a pre-polymerised BME layer and left in a humidified 37°C incubator for 24 hours. After 24 hours, dead cells were found floating above the BME, while live cells would have become embedded in the BME. The media and dead cells were aspirated and fresh organoid growth media was added. Fluorescence was checked using an EVOS M5000 cell imaging system (ThermoFisher) after 72 hours.

For the co-culture of organoids, a GFP-labelled population was co-cultured with an mCherry-labelled population in equal amounts to ascertain differences in proliferation. Cell culture wells were imaged on day 0, 1, 3, 5 and 7 using an EVOS M5000 cell imaging system (ThermoFisher). ImageJ was used to determine fluorescent intensity on each of these days.

Bulk RNA sequencing and analysis

Extraction of total RNA was performed using the RNeasy Mini Kit (Qiagen) as per manufacturer's protocol. The quantity and quality of the RNA was ascertained using the Agilent 2100 Bioanalyzer. Sequencing was performed by the Oxford Genomics Centre, Oxford using an Illumina NextSeq500 instrument using the standard paired-end protocol with a read length of 150bp. Fastq reads were processed to clip low quality leading and trailing edges and to remove any adapter content using Cutadapt (v.3.5). Quality checked FASTQ reads were aligned to the GRCh38 human genome and gene annotation (Ensembl release 105) using STAR aligner (v.2.7.3a) two-pass mode to generate gene-level quantification. Raw counts were processed in R (v.4.1.3) for all statistical testing and plotting purposes. Normalisation and differential expression were performed using limma (v.3.50.3). Subsequent bioinformatic analysis was performed by Dr Amit Mandal, Post-Doctoral Fellow in Bioinformatics, Buczacki Lab, who returned a differentially expressed gene list between

different organoid genotypes. Subsequent analysis was performed by me. This included the generation of heatmaps using gplots (v.3.1.1) and volcano plots using ggplot2 (v3.3.5). Analysis was also performed using Gene Set Enrichment Analysis (GSEA) for 'Hallmark' gene sets (50 gene sets), obtained from the molecular signatures database (MSigDB v4.0) (Liberzon et al., 2015).

Single-cell RNA sequencing and analysis

Organoids were dissociated into single-cell suspension by the disruption of organoid domes with TrypLe for 10min at 37°C. Cells were then counted using a hemocytometer, and resuspended at a concentration of 1000 cells/µl. ScRNA-seq was conducted using a 5' scRNA-seq gene expression workflow (10x Genomics). Cells were loaded onto a 10x Chromium Controller for GEM generation followed by single cell library construction using 10x Chromium Next GEM Single Cell 5' Library and gel bead kit v1.1 following manufacturer's instructions. Size profiles of amplified cDNA and final libraries for sequencing were verified by electrophoresis (Agilent 2100 Bioanalyzer system, High Sensitivity DNA Kit), and the concentration of final libraries was measured with Qubit (Thermo Fisher Scientific). Libraries were sequenced on an Illumina NextSeq 2000 (26 cycles read 1, 8 cycles i7 index, 98 cycles read 2), achieving a mean of 39,412 reads per cell. Gene by cell barcode quantification matrix were obtained using Cell Ranger (v.7.0.1) and Seurat (v.4.2.0) was used to perform normalisation and UMAP based cluster assignment of pass cells. Single sample gene set enrichment analyses (ssgsea) method from GSVA (v.1.42.0) was used on a custom list of gene signature sets defined by editing the MSigDb (v.7.5.1) pathways to leave out SPERMATOGENSIS, MYOGENESIS and PANCREAS_BETA_CELLS, and replacing the default WNT_BETA_CATENIN_SIGNALING set by WNT_SIGNALING gene set (MSigDb systematic name M5493), as per Househam et al. (Househam et al., 2022). Additional gene sets included were – genes up- or down-regulated in bulk RNA-seq

comparison of 'FBXW7-APC' versus 'APC-FBXW7' organoids (adjusted p-value < 0.05 and logFC >= |1.5|; n= 248 up- and n= 166 down-regulated genes), mouse foetal gastric epithelium gene set from Vallone et al. 2016 (n= 122 human homologues), intestinal stem cell signature from Merlos-Suárez et al. 2011 (n= 74), YAP pathway signature from Gregorieff et al. 2015 (n= 213 human homologues), crypt-base columnar (CBC) and regenerative stem cell (RSC) signature sets from Vazquez et al. 2022 (n= 340 and 206 human homologues respectively)(Supplementary table 3). Enrichment scores obtained per gene set was z-score transformed, and cells with scores greater than 3rd quartile or less than 1st quartile of the score distribution were termed as enriched or depleted respectively, for that gene set. To assess statistical enrichment or depletion of a given gene set in *FBXW7* mutant versus wildtype categories, chi-square test was applied on contingency table for cell counts from enriched or depleted classes. Pearson residuals were visualised using mosaic plot from R package vcd (v.1.4-11). For RNA velocity analysis, separate count matrices for spliced and unspliced transcripts were created using Kallisto-Bustools (v.0.27.3) with the La Manno et al. 2018 strategy, while using the same reference annotation from Cell Ranger (GRCh38, v.2020-A). Spliced/ unspliced count data were combined with Seurat based UMAP clustering using scVelo (v.0.2.5) under python v.3.10.6. RNA velocity graph computed was overlaid on UMAP cluster embeddings to infer trajectory direction. The analysis of results was performed by myself, with bioinformatics support from Dr Amit Mandal.

Assay for transposase-accessible chromatin with sequencing and analysis

Library preparation for ATACseq was performed using the ATAC-Seq Kit (Active Motif) as per the manufacturer's instructions. Multiplexing was performed using unique i7 and i5 indexed primers. The quantity and quality of DNA was assessed using the Agilent High

Sensitivity DNA Kit, ensuring that transposed DNA fragments were between 200 and 1000bp, with a periodicity of ~200bp. Sequencing was performed by the Oxford Genomics Centre, Oxford using an Illumina NextSeq2000 instrument. Quality checked fastq reads were aligned to the human genome (GRCh38) using bwa aligner (v.0.7.17). Alignment output (BAM files) were filtered, and TA (Tag Alignment) format data created using samtools (v.1.14), picard (v.2.6.0) and bedtools (v.2.30.0). Further, MACS3 (v.3.0.0b2) was used to perform peak calling which were then annotated using ChIPseeker (v.1.30.3). MACS3 identified peaks for the 'AF vs W' and 'FA vs W' comparisons were thresholded to select those with score (i.e., -10*log(q-value)) >median of the respective distribution, overlapping regions merged and the 50bp span from the centre of the peak region were extracted as input for *de novo* transcription factor motif enrichment analysis using rGADEM (v.2.46.0). Identified motifs were plotted using TFBSTools (v.1.36.0) and annotated for the closest matching core human transcription factors using JASPAR2020 (v.0.99.0). Gene/ motif-set level visualisation of ATAC-seq signal enrichment was performed using deeptools (v.3.5.2). Comparative bigwig files were created for 'AF vs W', 'FA vs W' and 'FA vs AF' with BPM (Bins per million) normalisation, non-covered regions skipped, ENCODE blacklist regions filtered out and effective genome size adjusted to reflect the filtering. The bigwig files were then profiled for ATAC-seq signal enrichment with reference-point at the TSS (Transcription Start Site) and upstream/ downstream 5kb region. Signal enrichments were profiled for each of these sets –all genes in the genome (gencode v.38; 'Whole Genome'), the CBC gene signature set ('CBC') and the 50bp genomic regions where de novo motif search identified TEAD1/2 motifs (542 regions; 'TEAD1/2 motif') or, SNAI1 motifs (491 regions; 'SNAI1 motif'). The analysis of results was performed by myself, with bioinformatics support from Dr Amit Mandal.

Validation of results from datasets

Two independent and published datasets were used to validate our findings. Somatic mutation and gene expression for the TCGA-COAD cohort were obtained from NCI-GDC data portal, to create a dataset of colon adenocarcinoma cases from 320 unique patients of whom 48 were mutated and the remainder wildtype for FBXW7. Differential gene-expression between $FBXW7^{mut.}$ and $FBXW7^{WT}$ samples was performed and the GSEA score for the "FA" up- and down-regulated gene sets was evaluated using fgsea (v.1.20.0). Per-sample enrichment scores for the gene sets was calculated using 'ssgsea' method in GSVA (v.1.42.0). The second validation made use of transcriptomic results derived from the EPICC dataset (Heide et al., 2022). Here, individual gland-level transcriptomic results were obtained. Generation of heatmaps was performed using gplots (v.3.1.1) and PCA plots using ggplot2 (v3.3.5).

Statistical analyses

All statistical analysis was performed using GraphPad Prism 9. Data are presented as mean ± standard deviation (sd). p-values < 0.05 were considered significant. Significance was further defined as follows: *p<0.05,**p<0.01, ***p<0.001, ****p<0.0001. Two tailed unpaired students' t-test was used for comparisons between groups with parametric data. All experiments were performed with biological triplicates taken from distinct samples.

Chapter 4: Biallelic *FBXW7* knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells

This work has been published with the following reference,

Chan DKH, Mandal A, Hester S, Yu Z, Higgins GS, Kessler BM, Fischer R, Buczacki SJA. Biallelic FBXW7 knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells. Cell Death Discov. 2023 Jun 29;9(1):200. doi: 10.1038/s41420-023-01494-y. PMID: 37386001; PMCID: PMC10310709.

Introduction

Regional sequencing of CRC has demonstrated a highly variegated landscape of ITH (Cross et al., 2018b, Sottoriva et al., 2015), where subclonal populations which harbour different mutations coexist in proximity with each other. Until recently, there has been limited understanding regarding the function of ITH in carcinogenesis. At present, there is significant interest in unpicking possible cooperative and collaborative features in the clonal dynamics between these subclonal populations (Chan and Buczacki, 2021). *Apc* mutant cells have been found to secrete NOTUM, a Wnt antagonist, into the tumour microenvironment, resulting in increased fitness of *Apc* mutant intestinal stem cells relative to wildtype counterparts (van Neerven et al., 2021b, Flanagan et al., 2021a). While there has been extensive research on the molecular interactions in driver mutations such as *APC* (Flanagan et al., 2021a, van Neerven et al., 2021b), *KRAS* (Zhu et al., 2021), *TP53* (Voorneveld et al., 2015), and *SMAD4* (Siraj et al., 2019), there are fewer studies which have elaborated the molecular mechanisms in other driver mutations. Multiregional genome and exome sequencing of both colon adenomas and CRC has revealed *FBXW7* to be a truncal and branch driver mutation (Cross et al., 2018b). Based on The Cancer Genome Atlas (TCGA) data for CRC adenocarcinoma (COAD-READ), the incidence of *FBXW7* mutations in CRC was found to be 17.5% (cBioPortal, 2023). Given

the high prevalence of *FBXW7* mutations in CRC, this study was designed to interrogate the possible cooperative or competitive roles of *FBXW7* mutants to neighbouring wildtype cells.

FBXW7 has been shown to recognise at least 90 different substrates (Arabi et al., 2012). Many of these substrates are protooncogenes which regulate gene transcription and stem cell differentiation. Some of these targets include c-Myc, Cyclin E and Notch1. Mutations of *FBXW7* abrogate the degradation and termination of pathways involving these proteins, and in the setting of cancer, can lead to persistent activation of cancer cell proliferation. FBXW7 is the most commonly mutated protein in the SCF complex amongst human cancer cells (Yumimoto and Nakayama, 2020). In CRC, *FBXW7* has been associated with tumour initiation and progression (Karczmarski et al., 2019). CRC is also the second most common cancer to harbour *FBXW7* mutations.

In the gut, *FBXW7* plays a critical role in the determination of intestinal cell lineage. In murine intestine, depletion of FBXW7 in the gut resulted in the upregulation of Notch signalling, which in turn upregulates *Hes* target genes involved in suppressing genes of a secretory lineage (Kopan and Ilagan, 2009). In this way, loss of FBXW7 results in an intestinal epithelium with markedly reduced goblet cells and enteroendocrine cells (Babaei-Jadidi et al., 2011).

Crucially, *FBXW7* is also known to interact with several other common driver mutations. In an $APC^{min/+}$ mouse model, mice developed greatly enhanced tumour formation when there was a co-occurring *FBXW7* mutation. Control mice heterozygous for *APC* succumbed to cancer at an average of 21.6 ± 1.65 weeks, compared with mice which had both the heterozygous *APC* mutation as well as an *FBXW7* mutation, which succumbed to cancer at 6.3 ± 1.45 weeks (Sancho et al., 2010). Separately, FBXW7 is also responsible for the degradation of p53 following DNA damage by ionising or ultraviolet radiation (Cui et al., 2020, Tripathi et al., 2019, Galindo-Moreno et al., 2019). p53 possesses a phosphodegron motif which allows for binding to FBXW7. In cell line models, mutant FBXW7 incapable of degrading p53 resulted in reduced proliferation and increased apoptosis following DNA

damage. Collectively, these results highlight the importance which *FBXW7* plays not only in its role as an E3 ubiquitin ligase, but also in its interaction with many other canonical CRC driver mutations. Given the prevalence of *FBXW7* mutationsin CRC, existing research has been limited to understanding the effects of *FBXW7* mutations in a cell autonomous fashion. This study builds on previous work by interrogating the possible cooperative or competitive roles of *FBXW7* mutants to neighbouring wildtype cells. In this way, I seek to uncover the possible interactions of mutant *FBXW7* cells in a milieu of ITH present in CRC.In this study, I defined one mechanism of interaction between mutant *FBXW7* and wildtype subclones. I identified that *FBXW7* mutants upregulate metabolism via oxidative phosphorylation. DNA damage was also upregulated in cells which were mutant for *FBXW7*, but strikingly, was also observed in neighbouring wild-type cells. I propose that A-kinase Anchoring Protein 8 (AKAP8), a member of the protein kinase A-anchoring protein family, which operates exclusively in the nucleus, is secreted in the microenvironment of *FBXW7* mutant cells, and could contribute to the increased DNA damage observed in *FBXW7* wildtype cells.

It is crucial to acknowledge in this introduction a number of characteristics of the *FBXW7* mutation which are relevant in the interpretation of results for this theme.

Firstly, I selected for a homologous knockout of the *FBXW7* gene as the model for the *FBXW7* mutant. My CRISPR targeting strategy involved an R505 truncating mutation, corresponding to the location of a propellor tip mutation which provides sufficient derangement of the SCF complex in a single allele (Davis and Tomlinson, 2012). *In vivo*, unlike the model selected in my study, two propellor tip truncating mutations were not observed. Based on the “just enough” hypothesis, two protein truncating propellor tip mutations is likely to result in such a reduction in SCF complex activity to render a less than ideal cellular phenotype. Nonetheless, such a model has been used to demonstrate the effect of *FBXW7* mutations in mice. Here, a bilateral truncating mutation possessed an equivalent phenotype of adenomatous polyposis as an R482Q/+ model, albeit to a lesser effect (Davis et al., 2014).

Secondly, the assignment of molecular weights for FBXW7 in western blots in depicted as 66kDa, which corresponds strictly to the γ-isoform of FBXW7. Again, the CRIPSR strategy which I employed targeted the R505 hotspot location, which should lead to a truncating mutation . In these experiments, the settings did not allow for sufficient resolution of the α- and γ- isoforms, which differ by a relatively short 15kDa in size across all three isoforms (Yeh, Bellon and Nicot, 2018), such that the three isoforms are most likely to be resolved into a single band on western blotting. To remain fiducial to my initial interpretation of the western blot, as well as the published manuscript, I have decided to maintain the molecular weight which I had previously annotated on the western blot.

Results

FBXW7$^{-/-}$ cells are at a proliferative disadvantage compared to wildtype cells

To evaluate the role of *FBXW7* subclones on neighbouring populations, I generated isogenic *FBXW7*$^{-/-}$ truncating mutations targeting the hotspot R505C on exon 10 in two cell lines, RKO and LS513, using CRISPR-Cas9. Successful gene editing was confirmed by Sanger sequencing (Figure 5a and 5c) and western blot (Figures 5b and 5d). The rationale for choosing RKO and LS513 for the cell lines to be used in this experiment was such as to reflect the molecular diversity of CRC, RKO being a microsatellite stable (MSS) cell line, while LS513 exhibits microsatellite instability (MSI) (Fichtner et al., 2020). Given known interactions between *FBXW7* and *TP53* as described in the introduction, it was important to choose cell lines which possessed wildtype *TP53*.

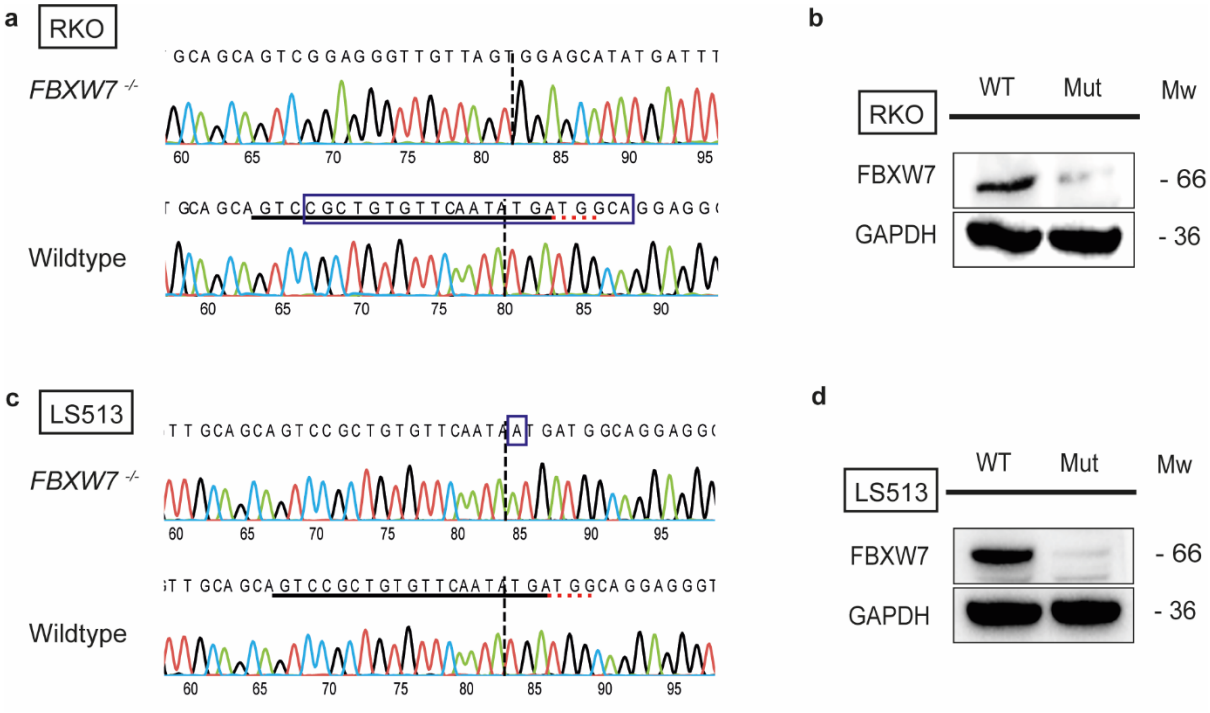


Figure 2: Validation of RKO and LS513 cell line models.

a. Sanger sequencing trace of RKO cell line showing an isogenic *FBXW7*$^{-/-}$ clone with 22 nucleotide deletion (blue box) from wildtype. b. Western blot validation showing loss of FBXW7 protein in RKO cells. c. Sanger sequencing trace of LS513 cell line showing an isogenic *FBXW7*$^{-/-}$ clone with 1 nucleotide insertion (blue box) from wildtype. d. Western blot validation showing loss of FBXW7 protein in LS513 cells. All experiments were performed with N = 3 biological replicates.

As *FBXW7* plays an important role in the regulation of genes involved in the cell cycle and cellular proliferation, the initial expectation was that *FBXW7*$^{-/-}$ cells would have increased proliferation. Instead, reduced proliferation in *FBXW7* $^{-/-}$ cells compared to wildtype cells was observed (Figure 6a). This finding was confirmed on an extreme limiting dilution clonogenic assay, where *FBXW7* $^{-/-}$ cells were also observed to form less colonies compared with their wildtype counterparts (Figure 6b). This was further corroborated with a reduction in the EdU incorporation rate on flow cytometry (Figures 6c and d) and decreased relative expression of the proliferation markers *MKI67* and *PCNA* on reverse transcription-polymerase chain reaction (RT-PCR) (Figure 6e) in *FBXW7* $^{-/-}$ cells.

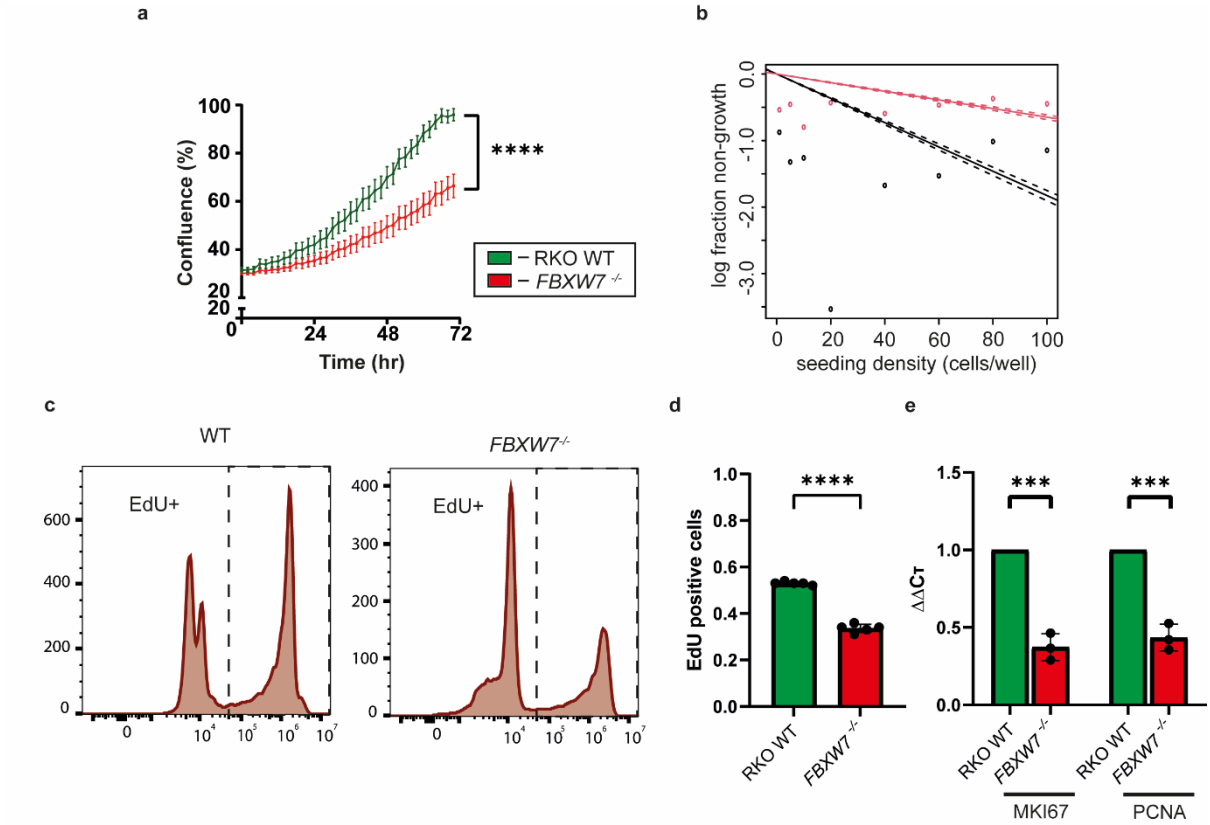


Figure 3: Comparison of proliferation and growth between FBXW7$^{-/-}$ and wildtype cells.

a. Live-cell proliferation analysis using IncuCyte imaging system showed increased proliferation of wildtype cells (green) relative to *FBXW7*$^{-/-}$ cells (red). (p< 0.0001). b. Results from clonogenic assay showing the increased ability of wildtype cells (black) to form clones relative to *FBXW7*$^{-/-}$ cells (red) (p< 0.0001). c. Histogram showing gating thresholds to determine EdU positivity. d. EdU incorporation assay showed increased proliferation occurring in wildtype cells relative to *FBXW7* $^{-/-}$ cells (p< 0.0001). e. RT-PCR showed increased MKI67 (p= 0.0002) and PCNA (p= 0.0003) in wildtype relative to *FBXW7* $^{-/-}$ cells. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

To ensure that the differences in growth were due to decreased proliferation and not increased apoptosis, markers of apoptosis in cells was investigated. The proportion of cells positive for Annexin V and 7-amino-actinomycin (7-AAD) on flow cytometry was not found to be significantly different between wildtype and *FBXW7* $^{-/-}$ cells (Figure 7a). In addition, neither an enzyme-linked immunosorbent assay (ELISA) for cleaved caspase 3 (Figure 7b) nor cathepsin B (Figure 7c) was significantly different, as was a quantitative calpain activity assay (Figure 7d). Taken together, these findings suggest that *FBXW7* $^{-/-}$ cells indeed proliferated slower than wildtype cells, and this was not a function of increased apoptosis.

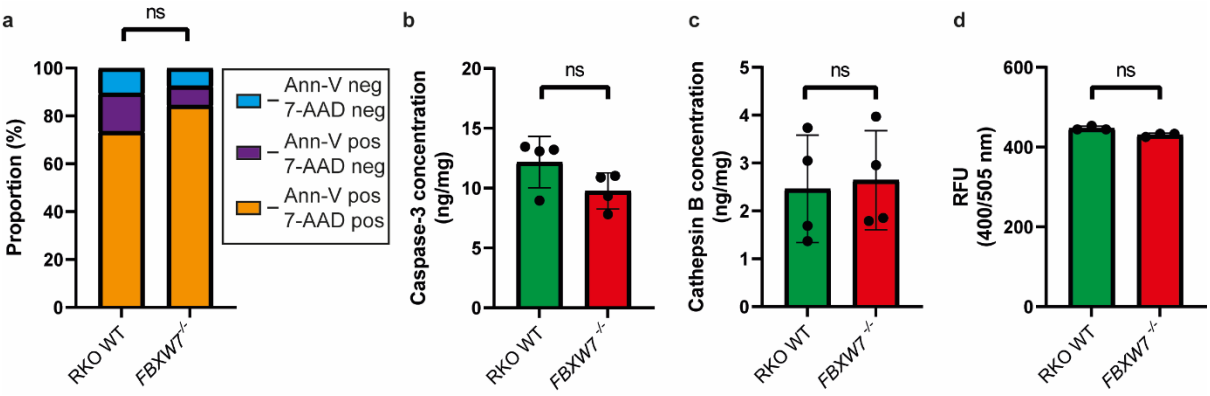


Figure 4: Comparisons of apoptosis markers between FBXW7$^{-/-}$ and wildtype cells.

a. No statistically significant differences in both early and late apoptotic markers (blue: Annexin V and 7-AAD negative; violet: Annexin V positive, 7-AAD negative; orange: Annexin V and 7-AAD positive). b. No difference in cleaved caspase 3 levels to suggest increased apoptosis accounting for the observed difference in cellular proliferation. c. No difference in Cathepsin B between wildtype and *FBXW7*$^{-/-}$ cells. d. No difference in Calpain activity between wildtype and *FBXW7*$^{-/-}$ cells. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

FBXW7$^{-/-}$ demonstrate upregulation of oxidative phosphorylation

To further characterise the molecular effects of knocking out *FBXW7* on a tumour cell, bulk RNA sequencing (RNAseq) on both wildtype and *FBXW7* $^{-/-}$ RKO cells was performed. Applying a cut-off of $\log_2$FC > 1.0 or < -1.0 and $p < 0.05$, 362 genes were significantly upregulated, while 532 genes were downregulated (Figure 8a). Transcriptomic expression of *FBXW7* was decreased in mutant cells compared with wildtype cells (Figure 8b). Substrate targets of FBXW7 were interrogated. Among *NOTCH1*, *MYC*, *CCNE1*, *MCL1*, *MTOR*, and *JUN*, only *MYC* transcription was downregulated(Figure 8c). I hypothesise that this could be a function of negative feedback homeostatic regulation of c-Myc levels resulting from decreased ubiquitination of c-Myc by the SCF complex. c-Myc is finely regulated at the transcriptional and translational levels (Penn et al., 1990, Secombe, Pierce and Eisenman, 2004). Moreover, the half-life of *MYC* mRNA is approximately 15min, and the mRNA is regulated by microRNAs (miRs) including miR-145, which can inactivate the *MYC* mRNA in response to p53 (Sachdeva et al., 2009).

Differential expression of a group of matrix metalloproteinases (*MMP12*, *MMP3*, *MMP1*, *MMP19*, and *MMP10*) were also observed to be significantly upregulated in *FBXW7* $^{-/-}$ cells (Figure 9). MMPs play a critical role in remodelling the extracellular matrix (ECM), and have been linked to oncogenic processes such as the proteolytic degradation of the ECM, as well as angiogenesis which ultimately lead to metastases (Gialeli, Theocharis and Karamanos, 2011). This could contribute to the observation that CRCs with mutant *FBXW7* have worse prognosis, greater risk of recurrence, and worse disease-free and overall survival (Korphaisarn et al., 2017, Kawashita et al., 2017, Kawaguchi et al., 2021). This remains the first experimental evidence of a link between *MMPs* and *FBXW7.*

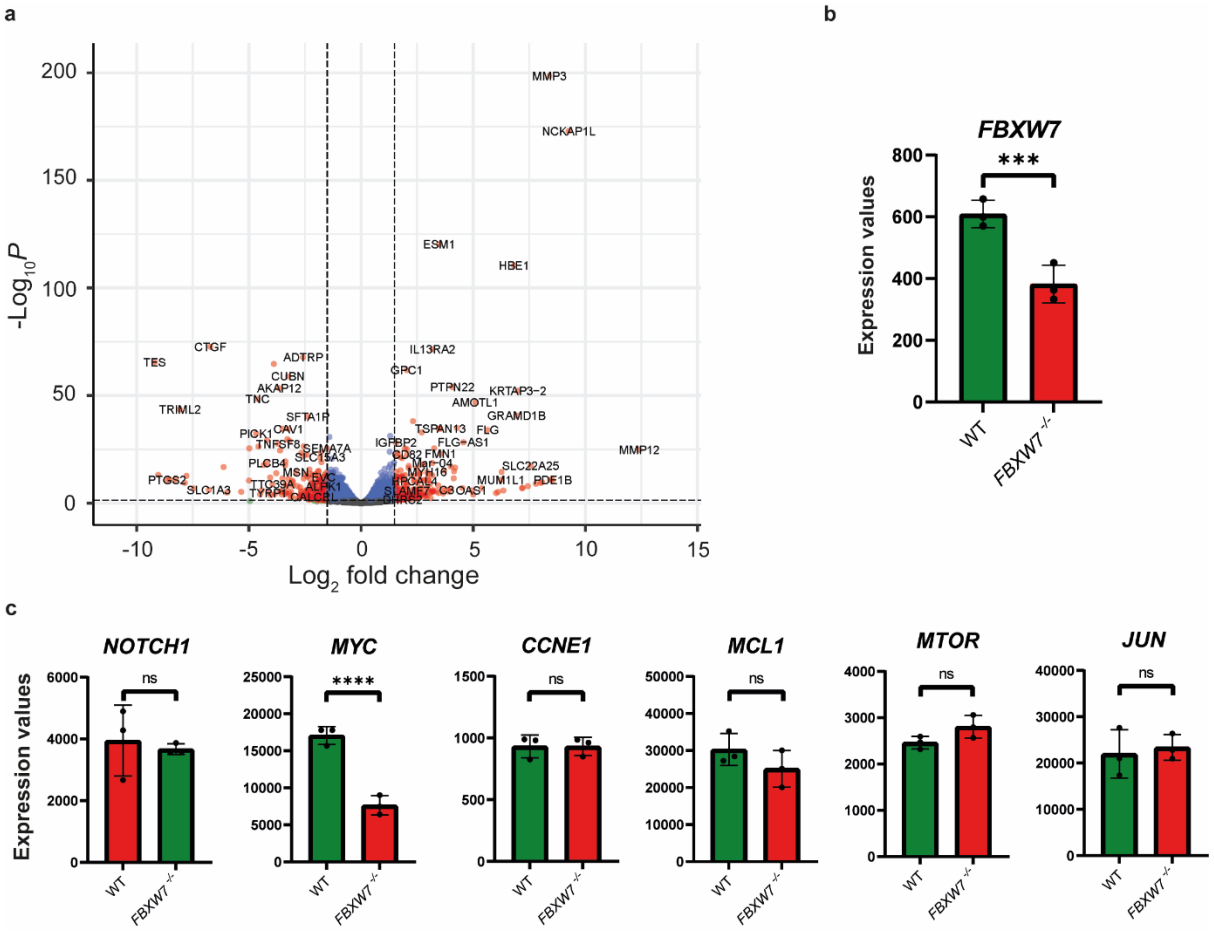


Figure 5: RNAseq results between wildtype and FBXW7$^{-/-}$ cells.

a. Heatmap showing top 50 differentially expressed genes between wildtype and *FBXW7*$^{-/-}$ cells. b. Transcript level for *FBXW7* was significantly decreased in *FBXW7*$^{-/-}$ cells compared to wildtype on bulk RNAseq (p= 0.0065). c. Transcript levels for genes which act as direct substrates for *FBXW7*. All differences were non-significant except for *MYC* (p= 0.0008) which was unexpectedly decreased in an *FBXW7* knockout, possibly due to negative feedback homeostatic regulation. Expression values depict the raw read counts for each gene. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

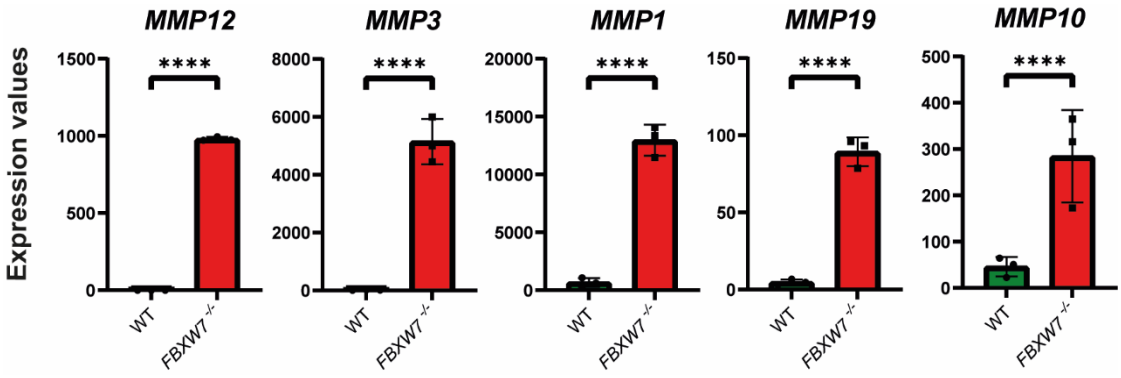


Figure 6: RNAseq results between wildtype and FBXW7$^{-/-}$ cells for MMP genes.

Transcript differences in metalloproteinases which were observed to be upregulated in *FBXW7*$^{-/-}$ cells. All were significant with p< 0.0001. Expression values represents the raw read count for each gene. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. Biological triplicates, N = 3, were submitted for RNAseq. MMP: metalloproteinase

RNAseq data was further analysed using Gene Set Enrichment Analysis (GSEA) for 'Hallmark' gene sets (50 gene sets), obtained from the molecular signatures database (MSigDB v4.0) (Liberzon et al., 2015). This revealed that the oxidative phosphorylation gene set had the greatest magnitude of increase based on the normalised enrichment score (NES) (Figure 10a). Out of the 200 genes included as part of the oxidative phosphorylation gene set, 78 genes were upregulated, while 8 genes were downregulated in *FBXW7* $^{-/-}$ cells relative to wildtype. Among these 86 genes, 14 genes had a $\log_2$ fold change < 0.7 or > 0.7 (Figure 10b). Curiously, this finding suggested a decrease in the "Warburg effect", in which aerobic glycolysis is upregulated in association with more aggressive cancers. In concordance with this association, the hypoxia hallmark gene set was downregulated in *FBXW7* $^{-/-}$ cells (Figure 10a). Our study validates earlier bioinformatic analyses based on the TCGA dataset which suggested that *FBXW7* mutations in a spectrum of various cancers cause metabolic reprogramming by upregulating oxidative phosphorylation. I also demonstrate through RNAseq that this metabolic reprogramming is the most upregulated pathway as determined by GSEA (Davis et al., 2018).

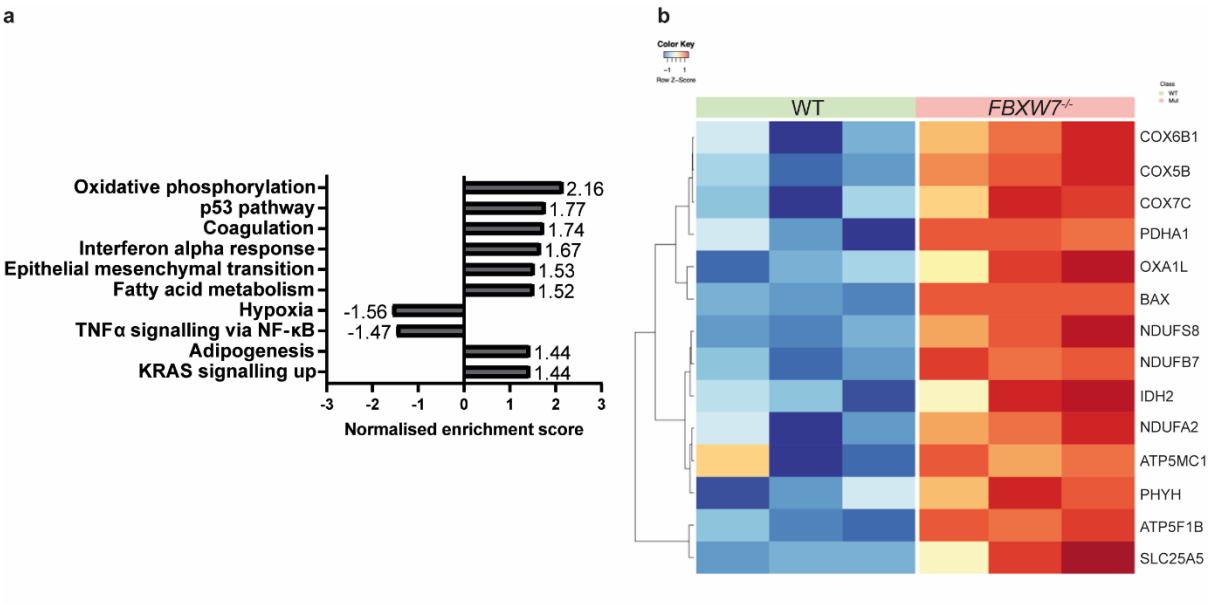


Figure 7: Analysis by GSEA reveals upregulation of oxidative phosphorylation in FBXW7 $^{-/-}$ cells.

a. Analysis of Hallmark pathways using GSEA demonstrates that the oxidative phosphorylation pathway was the most upregulated in *FBXW7*$^{-/-}$ cells. This graph shows the top ten differentially expressed gene sets by normalised enrichment scores. b. 86 genes within the oxidative phosphorylation gene set were differentially expressed. Among these 86, the 14 which showed log2FC >0.7 or <0.7 have been represented in this heatmap.GSEA: Gene Set Enrichment Analysis

Our findings on RNAseq were validated by measuring the metabolic activity of cells in terms of the oxygen consumption rate (OCR) and extracellular acid production rate (ECAR), as surrogates for oxidative phosphorylation and glycolysis respectively. I performed this metabolic measurement using an Agilent Seahorse Real-Time ATP Rate Assay kit on a Seahorse Agilent XFe96 Analyzer. Reagents are added to the cell media sequentially which function by arresting various facets of cellular metabolism. Oligomycin, a mitochondrial ATP synthesis inhibitor, is first added to the media. This results in mitochondrial ATP synthesis arrest, and results in a decrease in the OCR, which represents the ATP production value of the cell. Subsequently, Rotenone and Antimycin A is added, resulting in complete mitochondrial and respiratory inhibition. This suggests that any further acidification of the media detected must originate from glycolysis and not from the carbon dioxide produced through oxidative phosphorylation, therefore giving the calculation of the glycolysis-induced ATP production value, or the ECAR. Details on concentrations, and timing of measurements are provided in the Chapter 2. A significant increase in the OCR was noted, but changes in the ECAR was found to be non-significant (Figure 11a). HSP60 is a chaperone protein which is required for the formation of the mitochondrial respiratory chain complex (Zhou et al., 2018), while ATP5A is the catalytic subunit in complex V which generates adenosine triphosphate (ATP) from adenosine diphosphate (ADP). This was also further validated with western blot (Figure 11b) and immunofluorescent staining (Figures 11c – 11f) for both HSP60 and ATP5A.

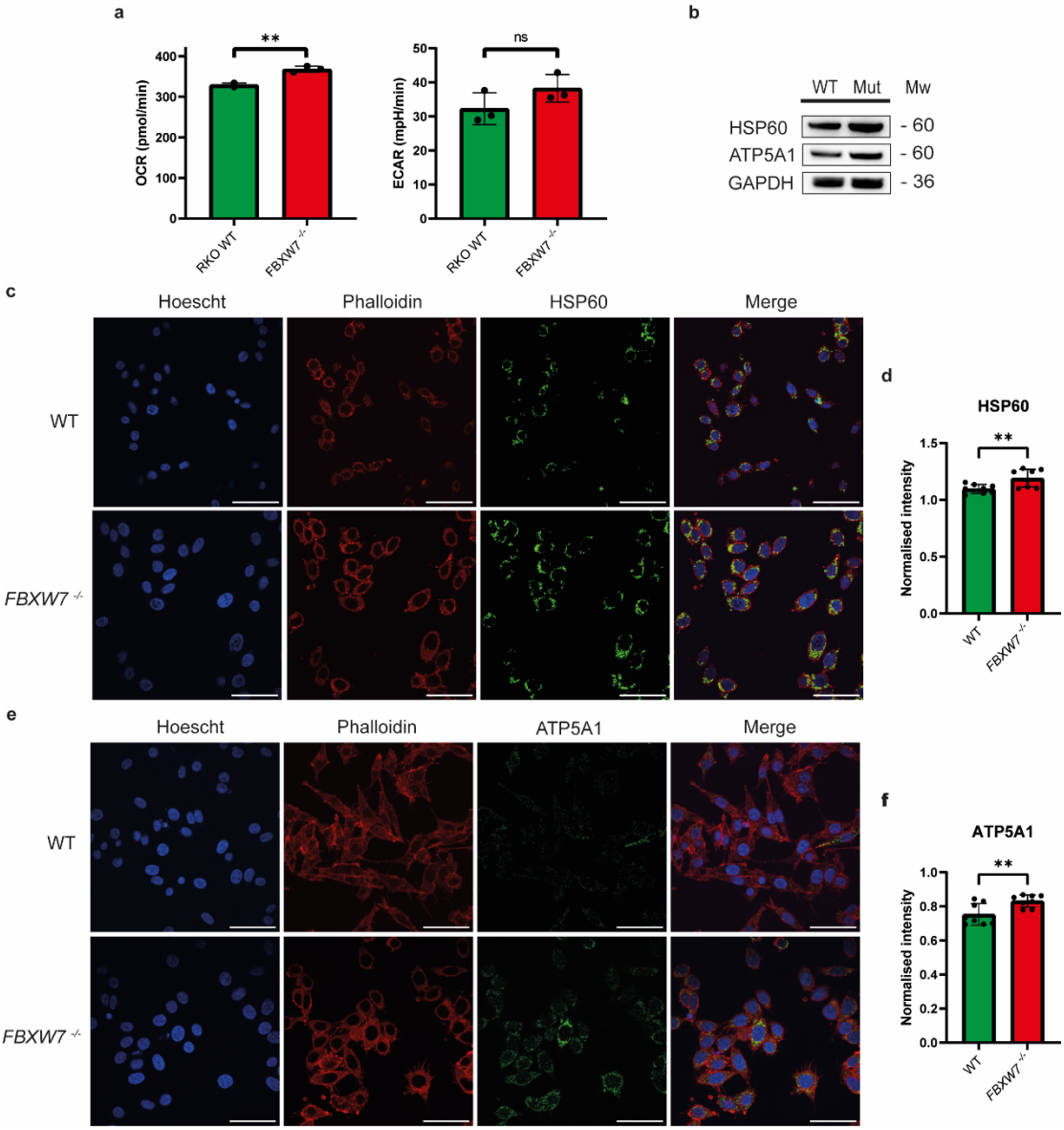


Figure 8: Validation of upregulated oxidative phosphorylation in *FBXW7* $^{-/-}$ cells.

a. Metabolic analysis of oxygen consumption rate was increased in *FBXW7*$^{-/-}$ cells relative to wildtype cells (p= 0.0015). There was no significant difference in the extracellular acid production rate. b. Western blot showing upregulation of HSP60 and ATP5A1 in *FBXW7*$^{-/-}$ cells compared to wildtype cells. c. Immunofluorescence of wildtype and *FBXW7*$^{-/-}$ cells with Hoechst 33342 nuclear stain (blue), phalloidin cytoskeleton stain (red) and HSP60 (green). Scale bar represents 50μm. d. Immunofluorescence of wildtype and *FBXW7*$^{-/-}$ cells with Hoechst 33342 nuclear stain (blue), phalloidin cytoskeleton stain (red) and ATP5A1 (green). Scale bar represents 50μm e. Immunofluorescence of HSP60 was normalised against Hoechst and was increased in *FBXW7*$^{-/-}$ cells (p= 0.0086). f. Immunofluorescence of ATP5A1 was normalised against Hoechst and was increased in *FBXW7*$^{-/-}$ cells (p= 0.0085). The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

FBXW7$^{-/-}$ induce DNA damage in neighbouring wildtype cells

The process of oxidative phosphorylation can result in the production of reactive oxygen species (ROS), which can result in DNA damage (Shadel and Horvath, 2015, Sabharwal and Schumacker, 2014). As expected, increased DNA damage was observed in *FBXW7* $^{-/-}$ cells in comparison with wildtype cells on IF staining for DNA damage markers such as anti-γH2AX and anti-53BP1 in both RKO (Figures 12b, 12d and 12e) and LS513 (Figures 12c, 12f and 12g) cells.

I had earlier observed decreased proliferation in *FBXW7* $^{-/-}$ cells compared with wildtype cells. In CRC, there is an association between worse overall survival in cancers with *FBXW7* mutations (Shang et al., 2021). This led me to consider a cooperative or competitive effect of *FBXW7* $^{-/-}$ cells on its neighbouring wildtype cancer cells, particularly given CRC's significant ITH where subclones with different driver mutations could interact with one another (Chan and Buczacki, 2021). To determine the effect of the *FBXW7* $^{-/-}$ cells on neighbouring wildtype cells, mutant and wildtype cells were cultured on either surface on the insert of a Transwell (Corning, New York, USA) plate (Figure 12a). A Transwell insert port size was selected such that media could freely permeate between both compartments of the Transwell setup, but without allowing for the migration of cells. To ensure all cells were treated uniformly, wildtype cells were also co-cultured with wildtype cells, and mutant cells were similarly co-cultured with mutant cells. These cells were similarly grown on both the apical and basal sides of the Transwell plate. At the time of harvesting, cells grown on either side of the Transwell plate were pooled together for onward experiments and analysis. For example, wildtype cells on the apical and basal layers of the cells were harvested and then pooled together. All conditions were cultured contemporaneously. For nomenclature purposes, wildtype cells cocultured with wildtype cells were henceforth collectively denoted as WT cells. Mutant cells cocultured with mutant cells were henceforth collectively denoted as Mut cells. Wildtype cells cocultured with *FBXW7* $^{-/-}$ cells were harvested separately after coculture and denoted CC-WT and CC-Mut cells respectively. Coculture experiments

involved both RKO and LS513 cell lines separately. IF for DNA damage markers was performed for anti-γH2AX and anti-53BP1 in all coculture conditions. Surprisingly, coculturing wildtype and mutant cells resulted in an upregulation of DNA damage markers in CC-WT cells when compared with WT cells. There was however no difference in the burden of DNA damage between CC-Mut and Mut cells (Figures 12b – 12g). These observations were further validated in a comet assay with confirmed increased tail length as well as tail moment in CC-WT cells compared with WT cells. (Figures 13a – 13c).

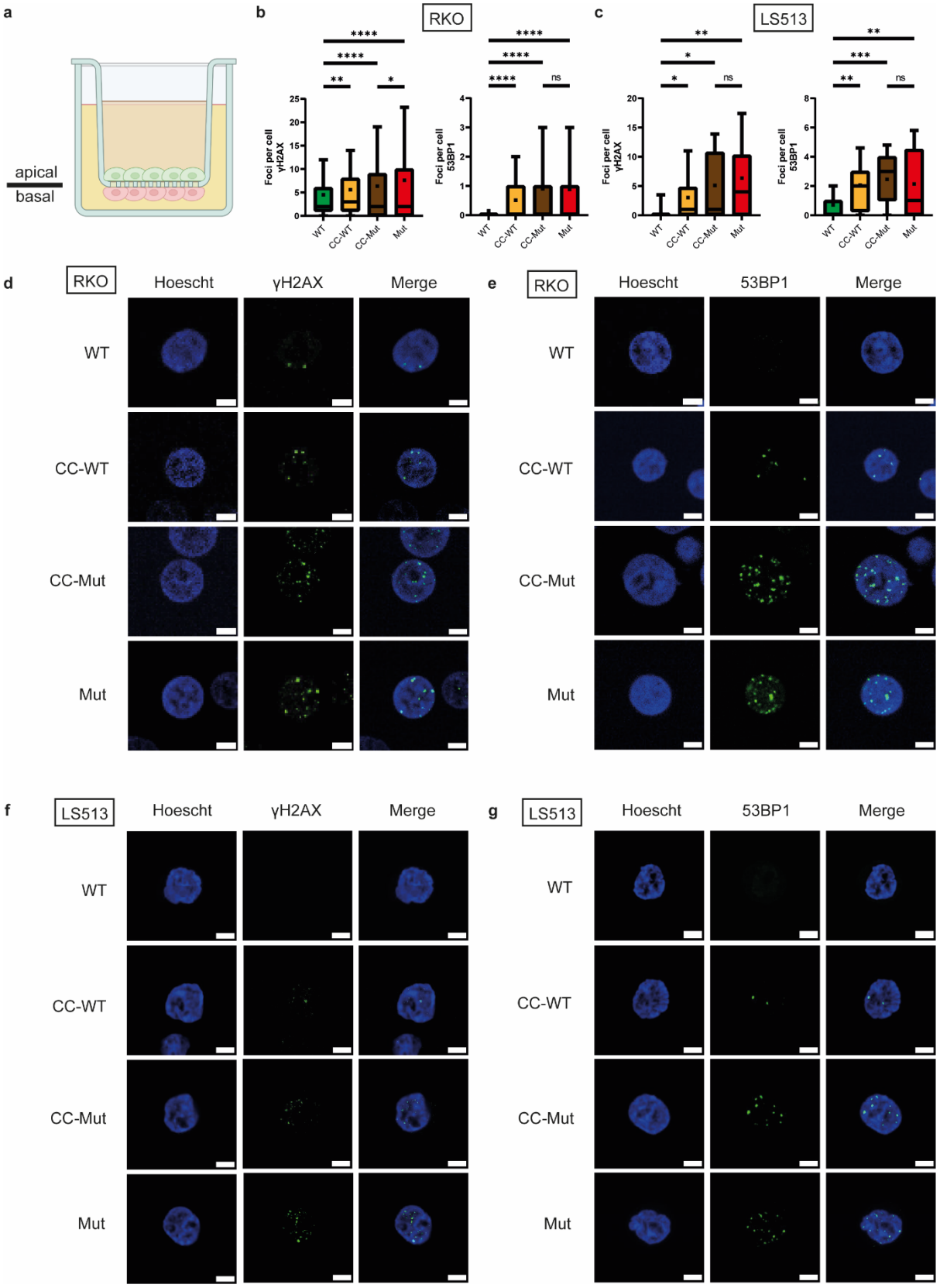

a
apical
basal
b
RKO
Foci per cell γH2AX
Foci per cell 53BP1
WT
CC-WT
CC-Mut
Mut
c
LS513
d
RKO
Hoescht
γH2AX
Merge
e
RKO
Hoescht
53BP1
Merge
f
LS513
Hoescht
γH2AX
Merge
g
LS513
Hoescht
53BP1
Merge

Figure 9: DNA damage was increased in wildtype cells co-cultured with FBXW7$^{-/-}$ cells.

a. Schematic showing how cells were cultured using the Transwell insert. Cells were cultured on both sides of the insert, and were therefore in close proximity with each another without the transfer of cells across the membrane. b. In RKO cells, γH2AX foci per nuclei was significantly increased between wildtype and all other groups. Notably, γH2AX foci was increased between WT and CC-WT cells. (WT vs CC-WT, p= 0.0032; WT vs CC-Mut, p< 0.0001; WT vs CC-Mut^{-}, p< 0.0001; CC-Mut vs Mut, p= 0.0154.) This same effect was observed when staining for 53BPI foci. (WT vs CC-WT, p< 0.0001; WT vs CC-Mut, p< 0.0001; WT vs Mut, p< 0.0001; CC-Mut vs Mut, ns.) c. In LS513 cells, γH2AX foci per nuclei was significantly increased between wildtype and all other groups. As in RKO cells, γH2AX was increased between WT and CC-WT cells (WT vs CC-WT, p= 0.0404; WT vs CC-Mut, p= 0.0189; WT vs Mut, p= 0.0029; CC-Mut vs Mut, ns). A similar effect was observed when staining for 53BP1 foci (WT vs CC-WT, p= 0.0022; WT vs CC-Mut, p= 0.0004; WT vs Mut, p= 0.0052; CC-Mut vs Mut, ns). Immunofluorescent images of nuclei with γH2AX foci in RKO cells (d) and LS513 cells (f). Immunofluorescent images of nuclei with 53BP1 foci in RKO cells (e) and LS513 cells (g). All scale bars represent 5μm. Boxplots are presented as 10 – 90 percentiles with the median value represented by a line across the box, and an asterisk to denote the mean. Statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates. WT: wildtype cells, CC-WT: wildtype cells co-cultured with mutant cells, CC-Mut: mutant cells co-cultured with wildtype cells, Mut: mutant cells

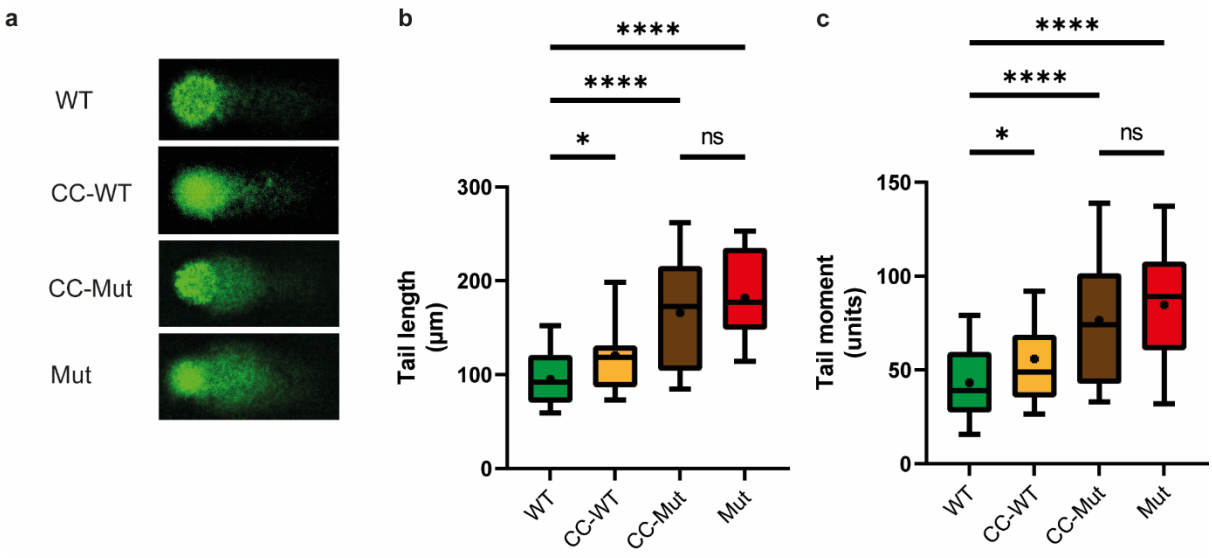


Figure 10: Comet assay showed DNA damage was increased in wildtype cells co-cultured with FBXW7$^{-/-}$ cells.

a. Single-cell gel electrophoresis images for WT, CC-WT, CC-Mut and Mut cells. b. Tail length was significantly increased between wildtype and all other groups. Notably, tail length was increased between WT and CC-WT cells (WT vs CC-WT, p= 0.0185; WT vs CC-Mut, p< 0.0001; WT vs Mut, p< 0.0001; CC-Mut vs Mut, ns). c. A similar relationship was observed for tail moment (WT vs CC-WT, p= 0.0367; WT vs CC-Mut, p< 0.0001; WT vs Mut, p< 0.0001; CC-Mut vs Mut, ns). Boxplots are presented as 10 – 90 percentiles with the median value represented by a line across the box, and an asterisk to denote the mean. Statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates. WT: wildtype cells, CC-WT: wildtype cells co-cultured with mutant cells, CC-Mut: mutant cells co-cultured with wildtype cells, Mut: mutant cells

Given that CRISPR-Cas9 induces double-strand breaks in the genome, I considered that the effects of increased DNA damage observed may be attributed to the gene editing process and related to *FBXW7*. To ascertain if the effects observed were related to *FBXW7*, I targeted the AAVS1 safe harbour locus and performed single-cell cloning to obtain a clonal population (Figure 14a). Measurement of immunofluorescence for anti-γH2AX and anti-53BP1 between wildtype and AAVS1 mutants did not show any difference, suggesting that the effect observed between CC-WT and CC-Mut cells was due to the introduction of a mutation in *FBXW7* (Figure 14b – 14e). Taken together, these findings suggest that *FBXW7* $^{-/-}$ cells induce DNA damage in neighbouring wildtype cells with no observable reverse effect noted.

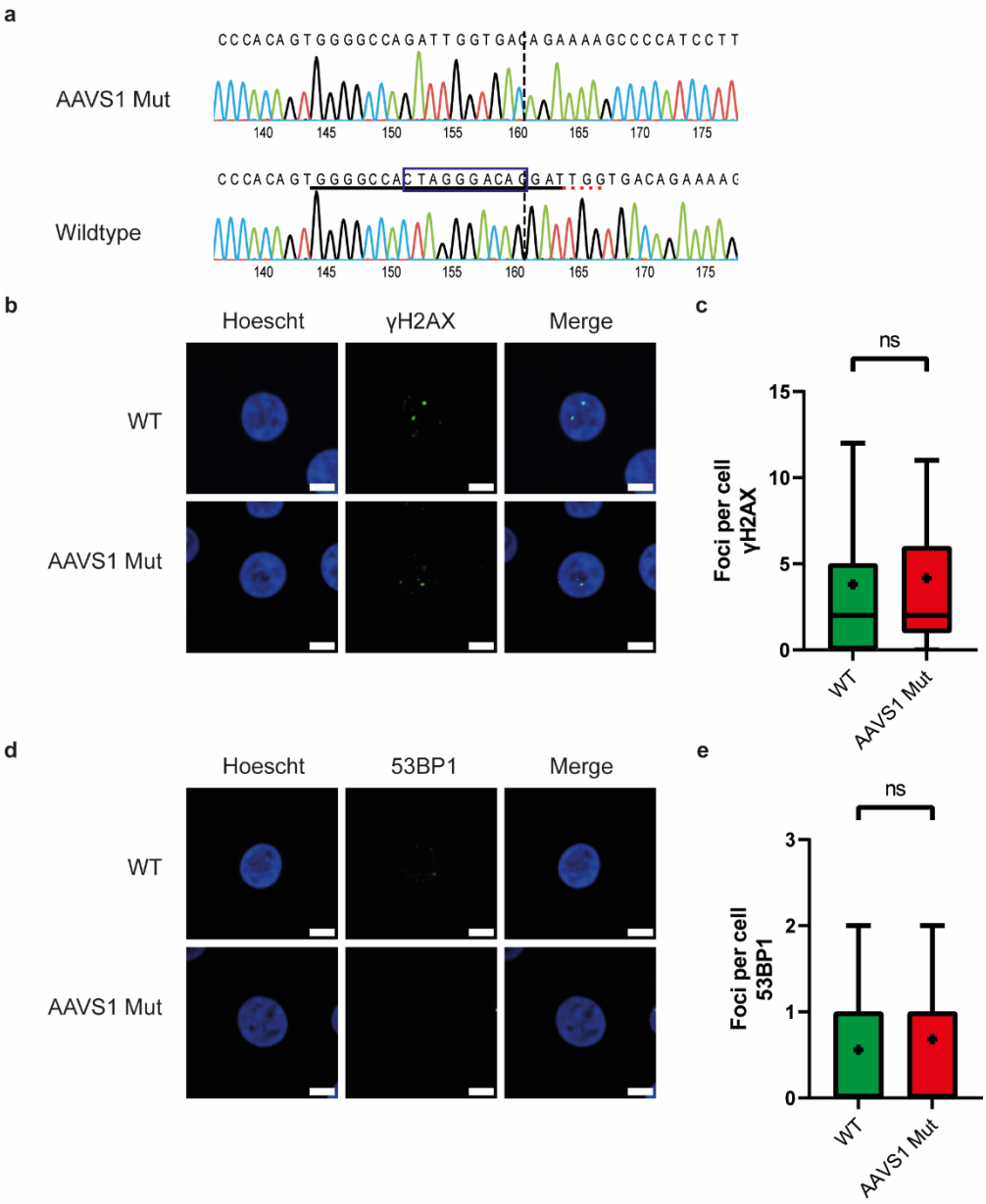


Figure 11: AAVS1 knockouts did not show increased DNA damage.

a. Sanger sequencing trace showing an isogenic AAVS1 knockout with 10 nucleotide deletion (blue box) from wildtype. b. Immunofluorescent images of nuclei with γH2AX foci showing no difference in the number of foci per nuclei (c). d. Immunofluorescent images of nuclei with 53BP1 foci showing no difference in the number of foci per nuclei (e). All scale bars represent 5μm. Boxplots are presented as 10 – 90 percentiles with the median value represented by a line across the box, and an asterisk to denote the mean. Statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

AKAP8 from *FBXW7*$^{-/-}$ mediates DNA damage on wildtype cells

I turned towards trying to uncover a number of observed phenomena. Firstly, I was interested in discovering the mechanism by which DNA damage was upregulated in Mut compared with WT cells. Next, I wanted to understand the mechanism which resulted in increased DNA damage in CC-WT cells compared with WT cells, which I hypothesised could be due to either a metabolite or a protein secreted into the media by *FBXW7*$^{-/-}$ cells. To investigate this, mass spectrometry was performed both on cell lysates as well as the media. In addition, as many of the proteins involved in DNA damage and repair require phosphorylation, mass spectrometry was used to analyse the global proteome as well as the phospho-proteome. Figure 15 summarises the various mass spectrometry experiments performed.

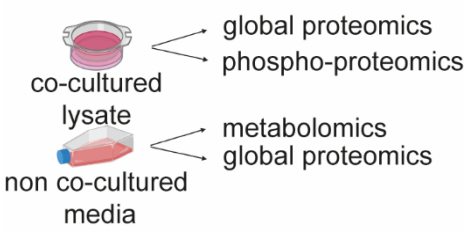


Figure 12: Schematic showing mass spectrometry experiments performed

Comparing between Mut and WT cells and applying a cut-off of $\log_2$FC > 2 or < -2 and p< 0.05, SERPINB9, HDAC2 and LUM were observed to be upregulated in Mut cells while ANXA1, MRS2, CMTR2, LMAN1, BIN3, RUNX1, TBP and MCFD2 were downregulated (Figure 16a). In CRC, ANXA1 (Annexin A1) was associated with a worse prognosis and an increased immune cell infiltrate (Figure 16b) (Liang and Li, 2021). MCFD2 and LMAN1 form a protein complex which shuttles cargo between the endoplasmic reticulum and the Golgi apparatus for transport out of the cell. Deficiencies in either protein have been shown to

result in α1-antitrypsin deficiency (Zhang et al., 2022) and factor V and factor VIII deficiency (Zhang, Liu and Zhang, 2022). Combined downregulation of MCFD2 and LMAN1 in *FBXW7* $^{-/-}$ cells (Figures 16c and 16d) suggests possible impairment of protein transport out of the cell, and could be a novel effect of abrogating the function of *FBXW7*. Conversely, SERPINB9 and HDAC2 were observed to be upregulated in *FBXW7* $^{-/-}$ cells and both have been described as a good prognostic marker for CRC (Figures 16e and 16f) (Vycital et al., 2019, Hu et al., 2020). Comparing phosphorylated proteins present in both Mut and WT cell lysates, and applying a cut-off of $\log_2$FC > 1 or < -1 and $p < 0.05$, seven proteins were found to be significantly downregulated in Mut compared with WT cells (Figure 16g); none were significantly upregulated. Among phosphorylated proteins which were exclusively expressed in either WT or Mut cell lysates, I observed 51 proteins found in cell lysates of WT cells only, and 25 found in cell lysates of Mut cells only.

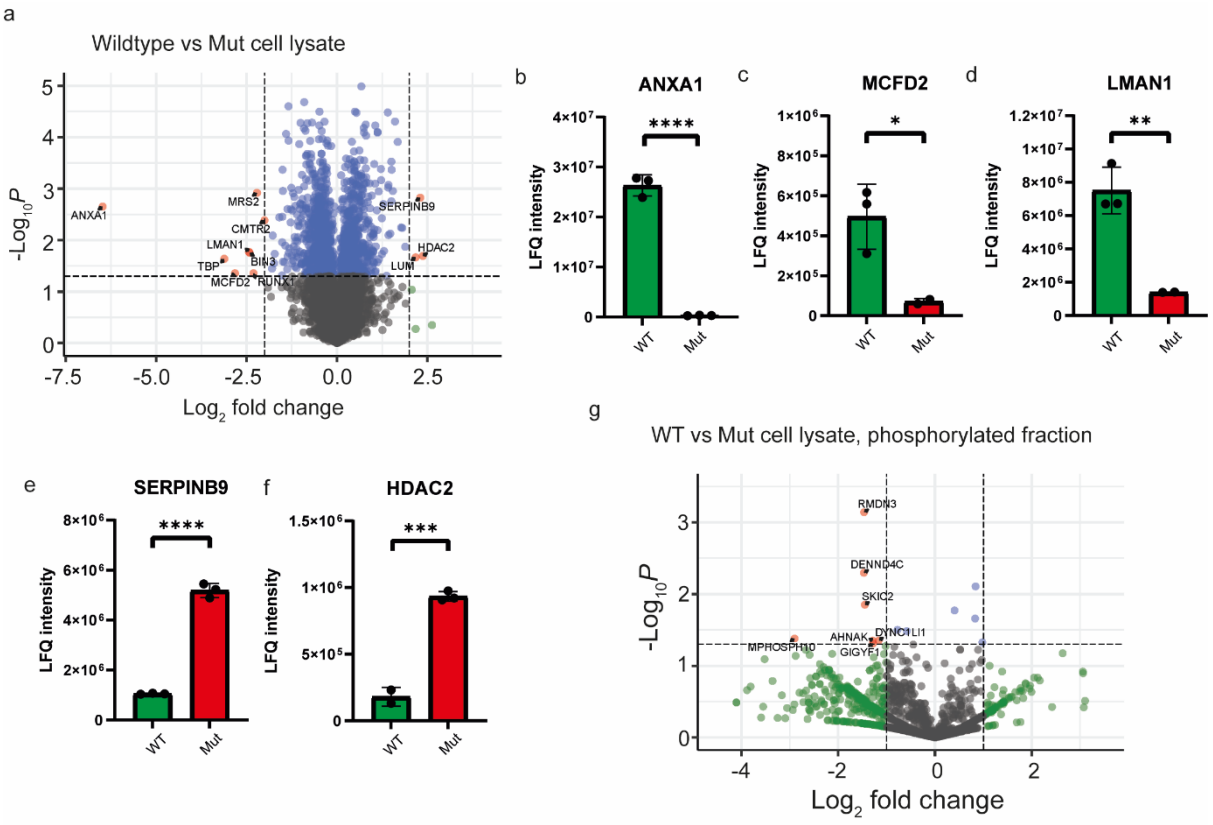


Figure 13: Proteomic analysis of WT and Mut cells.

a. Volcano plot showing proteomic differences between WT and Mut cells. Highlighted proteins had log_2FC > or < 2 and p< 0.05. Individual LFQ intensity differences for (b) ANXA1 (*p*< 0.0001), (c) MCFD2 (*p*= 0.0396), (d) LMAN1 (*p*= 0.0098), (e) SERPINB9 (*p*< 0.0001), and (f) HDAC2 (*p*= 0.0005) are represented here. g. Volcano plot showing proteomic differences between WT and Mut cells. Highlighted proteins had log_2FC > or < 1 and p< 0.05. Proteomic analyses were performed with N = 3 biological replicates. WT: wildtype cells, Mut: Mutant cells

Comparing between CC-WT and WT cells and applying a cut-off of $log_2FC > 0.7$ or < -0.7 and $p< 0.05$, STXBP2, TTC2, GEMIN2, LMF1, ADAMTSL1 and RBL2 were upregulated in CC-WT cells while SLC2A1, NDRG1, METTL15, UPRT and P4HA2 were downregulated (Figures 17a – 17d). *SLC2A1* is the gene for the glucose transporter 1 (GLUT1) protein, an important mediator of glycolysis. GLUT1 facilitates the diffusion of glucose molecules across cell membranes, and its expression therefore closely mirrors glycolytic activity within cells (Wu et al., 2022). This observation suggests metabolic reprogramming away from glycolysis and towards oxidative phosphorylation, and therefore appeared concordant with our earlier findings of the metabolic reprogramming in *FBXW7*$^{-/-}$ cells (Figure 17b). In turn, this could account for the increase in DNA damage observed in CC-WT cells. The downregulation of NDRG1 (Figure 17c) and upregulation of LMF1 (Figure 17d) in CC-WT cells were also observed. These two genes have both been found to be independently associated with a poorer prognosis in CRC (Chen et al., 2018, Liu et al., 2019) suggesting that wildtype cells neighbouring *FBXW7* $^{-/-}$ cells could have a more aggressive phenotype compared with wildtype cells which were not located adjacent to *FBXW7* $^{-/-}$ cells. Comparing phospho-proteins which were isolated in both CC-WT and WT cells, and applying a cut-off of $log_2FC > 0.7$ or < -0.7 and $p< 0.05$, only SRRM2 was downregulated in CC-WT cells, while only SEPTIN9 was upregulated (Figure 17e).

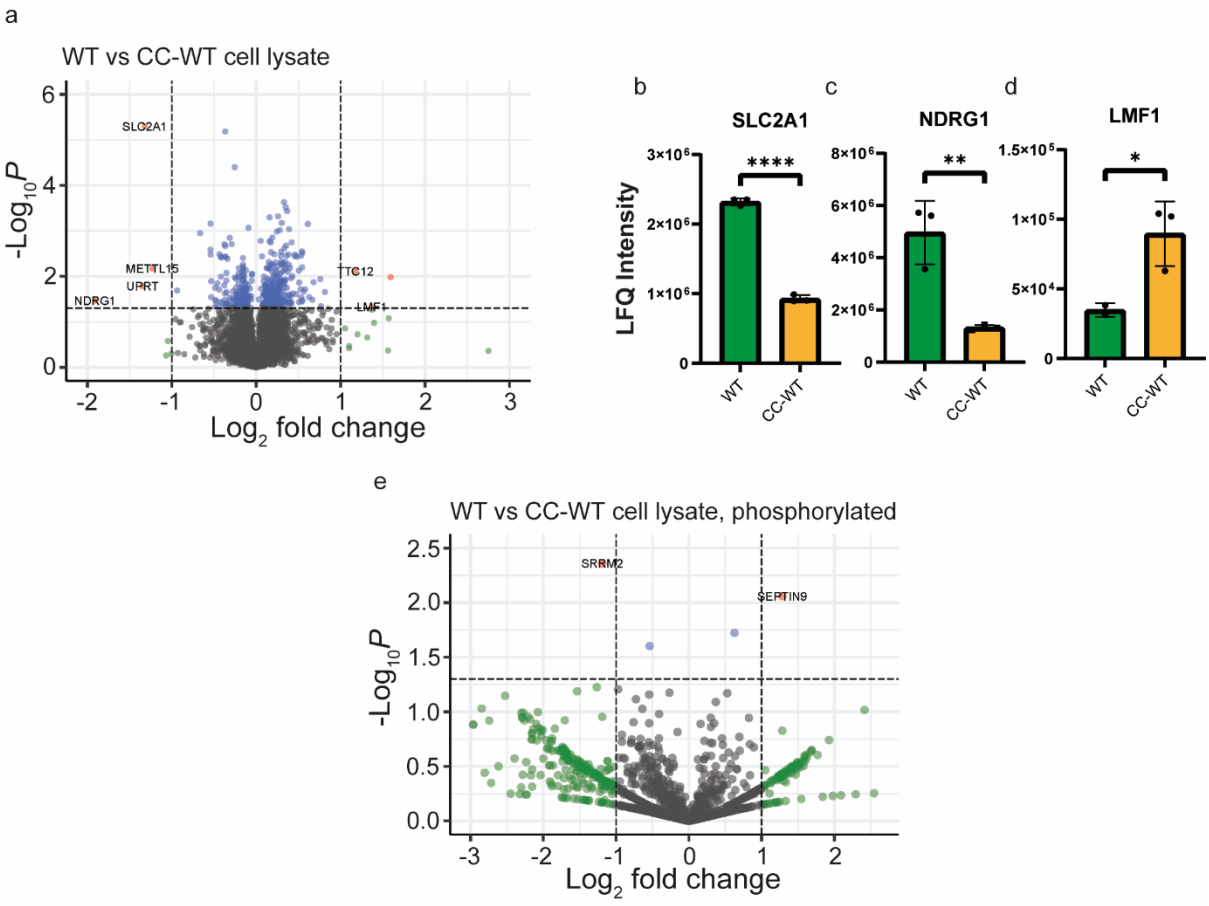


Figure 14: Proteomic analysis of wildtype and CC-WT cells.

a. Volcano plot showing global proteomic differences between wildtype and CC-WT cells. Highlighted proteins had log2FC > or < 1 and p< 0.05. Individual LFQ intensity differences for (b) SLC2A1 (p< 0.0001), (c) NDRG1 (p= 0.0343), and (d) LMF1 (p= 0.0476). e. Volcano plot showing phospho-proteomic differences between wildtype and CC-WT cells. Highlighted proteins had log2FC > or < 1 and p< 0.05. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. Proteomic analyses were performed with N = 3 biological replicates. CC-WT: wildtype cells co-cultured with mutant cells

Given that our Transwell coculture system allowed for media to freely permeate between different cell coculture types, I considered that the causative agent which resulted in increased DNA damage in otherwise wildtype cells could be caused by metabolites or proteins which were secreted into the media by *FBXW7* $^{-/-}$ cells. This compound could be upregulated in media from *FBXW7* $^{-/-}$ cells. Media from both wildtype and *FBXW7* $^{-/-}$ cells were harvested separately and subjected to metabolic and proteomic analysis. Comparing proteins which were found in media from both wildtype and *FBXW7* $^{-/-}$ cells, and applying a cut-off of $\log_2FC > 0.7$ or < -0.7 and $p < 0.05$, EVPL, TMUB1, RPS15 and CREBRF were found to be upregulated in the media of *FBXW7* $^{-/-}$ cells (Figure 18a). In addition, AKAP8 was found to be only secreted by *FBXW7* $^{-/-}$ cells and not by wildtype cells (Figure 18b), and therefore not reflected on the volcano plot. Metabolic analysis of media from both wildtype and *FBXW7* $^{-/-}$ cells did not yield significant differences, except for lactate (Figure 18c), which likely reflected the increased metabolic demand in *FBXW7* $^{-/-}$ cells.

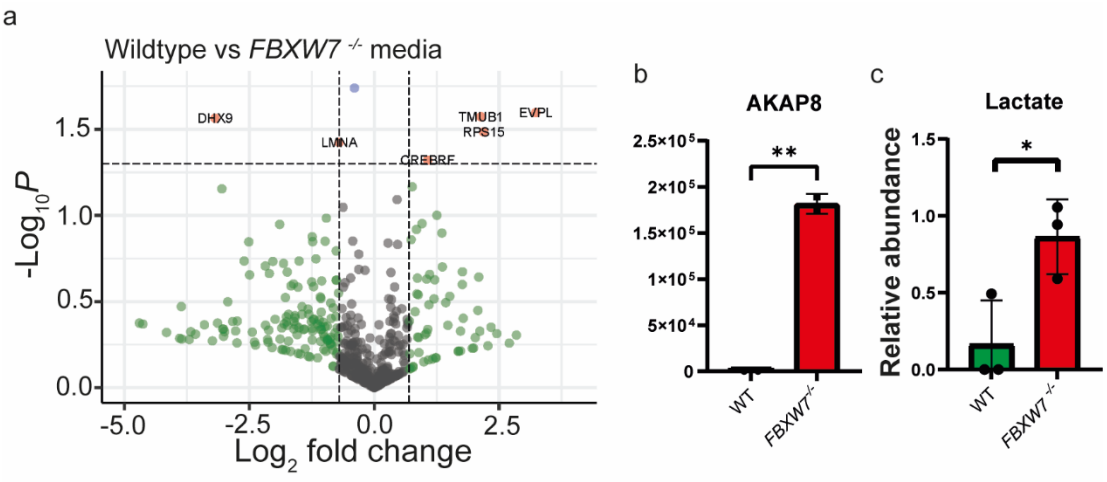


Figure 15: Proteomic and metabolic analysis of media from wildtype vs FBXW7$^{-/-}$ cells.

a. Volcano plot showing proteomic differences between the media from wildtype and *FBXW7*$^{-/-}$ cells for proteins which were present in both media fractions. Highlighted proteins had log_2FC > or < 0.7 and p< 0.05. b. AKAP8 was found in the *FBXW7*$^{-/-}$ cell media but not in the wildtype media. The LFQ intensity difference is represented here (*p*= 0.0018). c. In the metabolic analysis of media from wildtype and *FBXW7*$^{-/-}$ cells, only lactate was found to be significantly increased in media from *FBXW7*$^{-/-}$ cells. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. Proteomic and metabolic analyses were performed with N = 3 biological replicates.

Given that AKAP8 was found only in the media of *FBXW7* $^{-/-}$ cells and not in wildtype cells, I focused on AKAP8 as the putative agent causing increased DNA damage in adjacent wildtype cells. This was performed via two modalities. In the first approach, overexpression of AKAP8 with a pcDNA3.1(+)-AKAP95-FRB plasmid was performed. The pcDNA3.1 with an empty vector was used as a control. Western blot in both RKO and LS513 cells confirmed upregulation of AKAP8 in the cell population after overexpression (Figures 19a and 19b). Immunofluorescence of γH2AX and 53BP1 was observed to be upregulated in both RKO (Figures 19c, 19e, 19g) and LS513 (Figures 19d, 19f, 19h) with overexpression of AKAP8, and not in those which had been transfected with the empty vector.

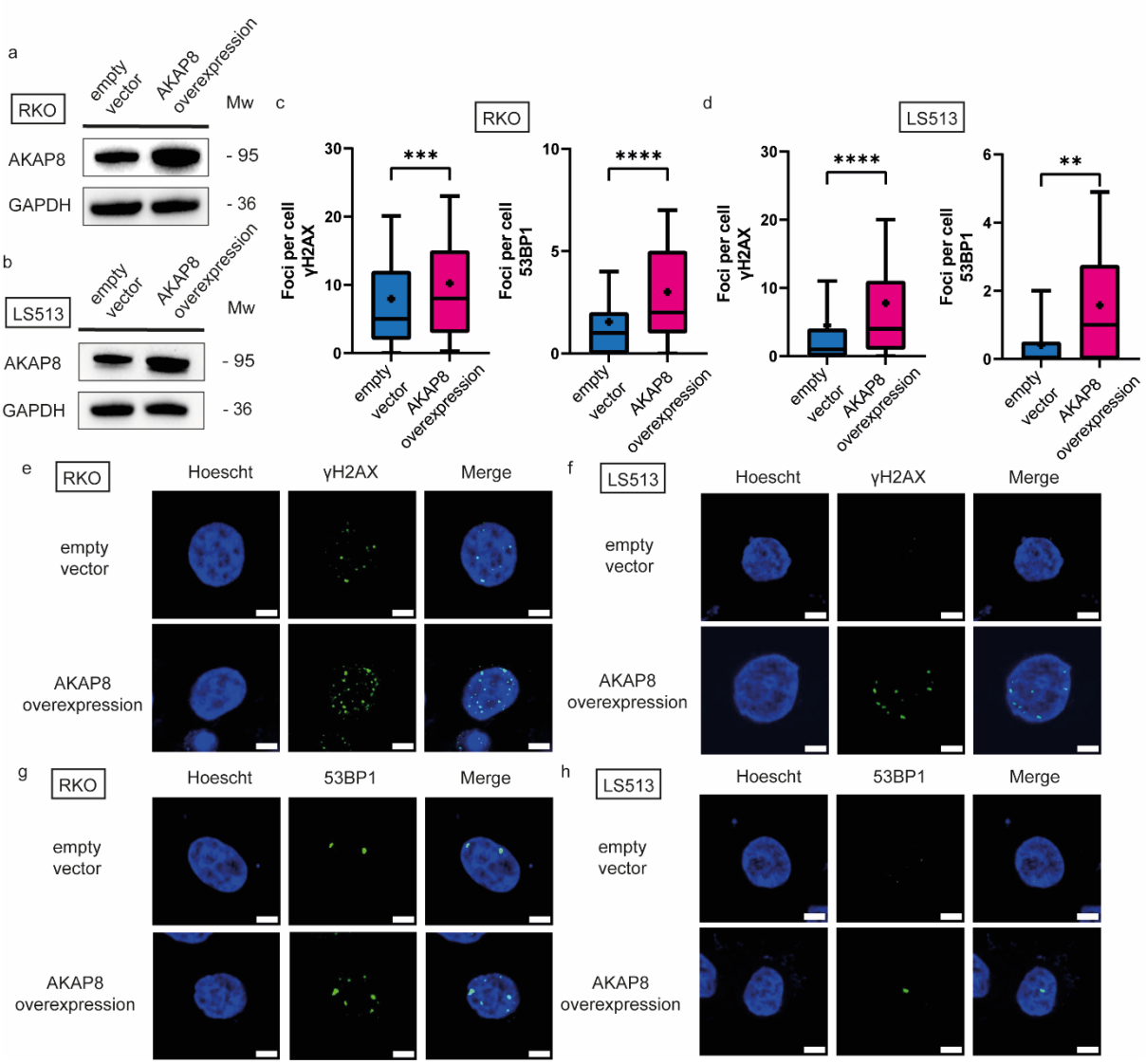


Figure 16: AKAP8 overexpression induces DNA damage in wildtype cells.

AKAP8 was overexpressed in wildtype cells and observed for DNA damage. Western blot demonstrating upregulation of AKAP8 in RKO cells (a) and LS513 cells (b). c. In RKO cells, γH2AX foci was upregulated between cells transfected with an empty vector and cells with the AKAP8 overexpression vector (p= 0.0008). 53BP1 was also upregulated in cells with the AKAP8 overexpression vector (p< 0.0001). d. Similar results were observed in LS513 cells for γH2AX (p< 0.0001) and 53BP1 (p= 0.0026). Immunofluorescent images of nuclei with γH2AX foci in RKO cells (e) and LS513 cells (f). Immunofluorescent images of nuclei with 53BP1 foci in RKO cells (g) and LS513 cells (h). All scale bars represent 5μm. Boxplots are presented as 10 – 90 percentiles with the median value represented by a line across the box, and an asterisk to denote the mean. Statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

In a second approach, an *AKAP8* mutation was introduced into *FBXW7*$^{-/-}$ cells to determine whether AKAP8 was indeed a crucial component of the secretome which resulted in DNA damage of neighbouring wildtype cells. As with previously, double mutant isogenic *FBXW7*$^{-/-}$ / *AKAP8*$^{-/-}$ cells were generated in RKO and LS513 lines using CRISPR-Cas9 and the knockout validated using both Sanger sequencing and western blot (RKO: Figures 20a and 20b; LS513: Figures 20c and 20d). Wildtype cells were cocultured with either a single mutant *FBXW7*$^{-/-}$ or double mutant *FBXW7*$^{-/-}$ / *AKAP8*$^{-/-}$ cells. Immunofluorescence of DNA damage markers was decreased in wildtype cells co-cultured with double mutant *FBXW7*$^{-/-}$ / *AKAP8*$^{-/-}$ cells compared with single mutant *FBXW7*$^{-/-}$ cells in both RKO (Figures 20e, 20g, and 20i) and LS513 (Figures 20f, 20h, and 20j) cells. Taken together, these suggest a crucial role for AKAP8 secreted by *FBXW7*$^{-/-}$ cells in inducing DNA damage in neighbouring wildtype cells.

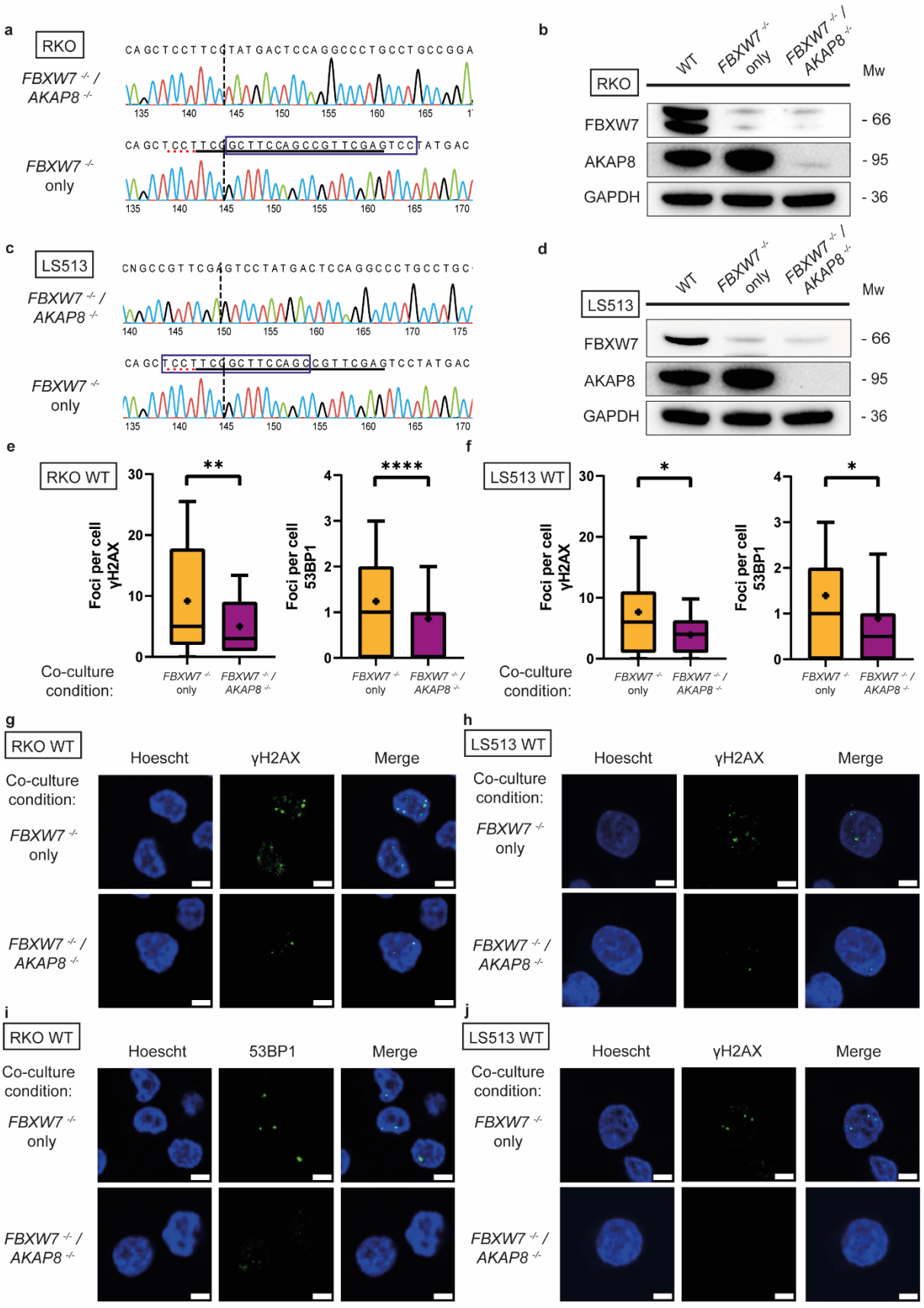

a
RKO
FBXW7 -/- / AKAP8 -/-
FBXW7 -/- only
b
RKO
WT
FBXW7 -/- only
FBXW7 -/- / AKAP8 -/-
Mw
FBXW7
AKAP8
GAPDH
- 66
- 95
- 36
c
LS513
FBXW7 -/- / AKAP8 -/-
FBXW7 -/- only
d
LS513
e
RKO WT
Foci per cell γH2AX
Foci per cell 53BP1
Co-culture condition:
FBXW7 -/- only
FBXW7 -/- / AKAP8 -/-
f
LS513 WT
g
RKO WT
Hoescht
γH2AX
Merge
h
LS513 WT
i
RKO WT
53BP1
j
LS513 WT

Figure 17: Loss of AKAP8 protein abrogates the DNA damage effect seen in neighbouring cells. a. Sanger sequencing trace of RKO cell line showing an isogenic *AKAP8*$^{-/-}$ clone with 20 nucleotide deletion (blue box) from wildtype. b. Western blot validation showing loss of FBXW7 or AKAP8 protein in RKO cells. c. Sanger sequencing trace of LS513 cell line showing an isogenic *AKAP8*$^{-/-}$ clone with 16 nucleotide deletion (blue box) from wildtype. d. Western blot validation showing loss of FBXW7 or AKAP8 protein in LS513 cells. RKO wildtype cells were co-cultured with either single mutant *FBXW7*$^{-/-}$ cells or double mutant *FBXW7*$^{-/-}$ */ AKAP8*$^{-/-}$ cells. In RKO cells, γH2AX (p= 0.0011; Figures 20e and 20g) and 53BP1 (p< 0.0001; Figures 20e and 20i) foci were decreased in double mutant cells compared with single mutant cells. Similarly, LS513 wildtype cells were co-cultured with either single mutant *FBXW7*$^{-/-}$ cells or double mutant *FBXW7*$^{-/-}$ */ AKAP8*$^{-/-}$ cells. In LS513 cells, γH2AX (p= 0.0285; Figures 20f and 20h) and 53BP1 (p= 0.0479; Figures 20f and 20j) foci were decreased in double mutant cells compared with single mutant cells. Boxplots are presented as 10 – 90 percentiles with the median value represented by a line across the box, and an asterisk to denote the mean. Statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates.

Effects of *FBXW7* mutation are recapitulated in TCGA-COAD data

Finally, TCGA-COAD data was analysed to determine if my results could be recapitulated in patient cancer data. Comparing transcriptomic differences between patients with wildtype and mutant *FBXW7*, and using a q-value cut-off of <0.05, 302 genes showed increased expression in CRC with mutant *FBXW7*, while 394 genes showed increased expression in CRC with wildtype *FBXW7* cells. GSEA using 'Hallmark' gene sets (50 gene sets) obtained from the molecular signatures database (MSigDB v4.0) showed that the oxidative phosphorylation gene set was significantly upregulated in mutant tumours. The oxidative phosphorylation gene set had an NES of 1.99, and a false discovery rate (FDR) q-value of < 0.0001 (Figure 21a).

The aforementioned 302 genes which were significantly upregulated in mutant *FBXW7* cells were grouped as a "TCGA_FBXW7_mutant" gene set. GSEA was performed on our bulk RNAseq analysis comparing RKO wildtype and *FBXW7* $^{-/-}$ cells against this "TCGA_FBXW7_mutant" gene set (Figure 21b). A good correlation between our *FBXW7* model and TCGA-derived patient data of *FBXW7* mutant cells was observed, suggesting that results from our experiments using our *FBXW7* cell line models may be extrapolated to real-world patient data.

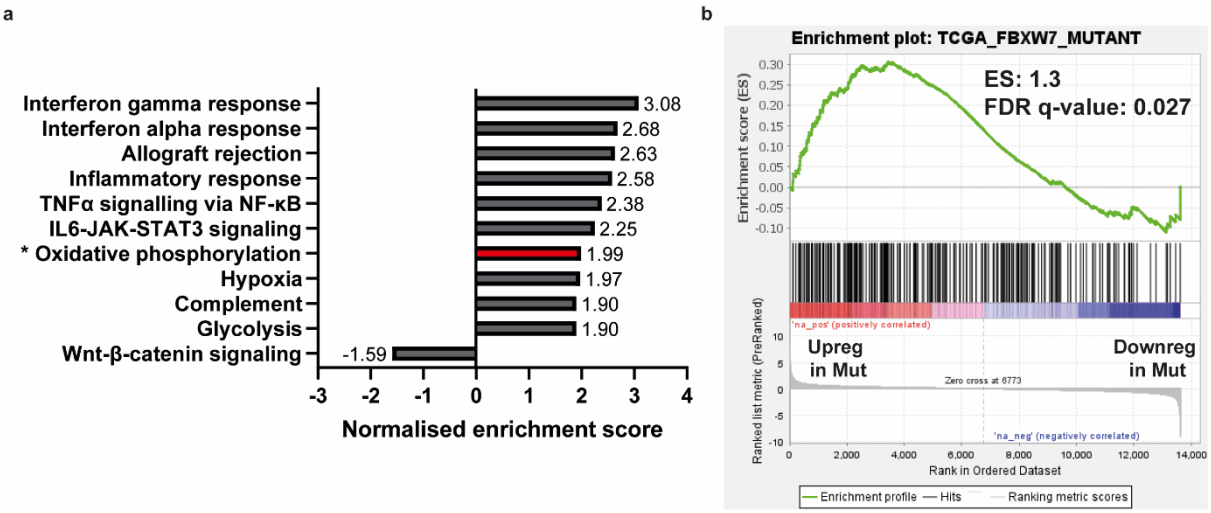


Figure 18: Correlation of the FBXW7 cell line models with TCGA-COAD-derived patient samples. a. Analysis of Hallmark pathways using GSEA demonstrates that the oxidative phosphorylation pathway was also upregulated in TCGA-COAD-derived FBXW7 mutant cancers. The top 10 differentially expressed gene sets are represented here, with the addition of the most differentially expressed gene set which had a negative correlation in FBXW7 mutants. For Oxidative phosphorylation (highlighted with asterisk, and bar in red), the NES was 1.99, with an FDR q-value< 0.0001. b. Enrichment plot depicting a significant correlation between the gene set associated with TCGA-COAD-derived FBXW7 mutant patient samples, and our own RKO cell line derived bulk RNAseq data. TCGA-COAD: The Cancer Genome Atlas – colon adenocarcinoma, GSEA: gene set enrichment analysis, NES: Normalised enrichment score

Discussion

AKAP8, also known as AKAP95, is one of over 50 different types of A-kinase anchoring proteins. The main function of AKAP proteins is to mediate cAMP-dependent protein kinase A (PKA) activity. PKA activity is governed by two regulatory subunits, while its effect is mediated by two catalytic subunits. Binding of AKAPs to PKAs confines the activity of PKA to a narrow subset of potential substrates (Carnegie, Means and Scott, 2009). AKAP8 is primarily localised to the nucleus. Association of AKAP8 to PKA as well as PKA-mediated cAMP signalling is responsible for maintaining chromatin in a condensed state during mitosis (Collas, Le Guellec and Taskén, 1999). Independent of PKA, AKAP8 can also regulate DNA replication (Eide et al., 2003), and promotes tumorigenesis and increased proliferation by binding to cyclin D and cyclin E during G1/S phase (Kong et al., 2016). In addition, RNAseq performed on MDA-MB-231 cells in which AKAP8 was knocked down showed on gene ontology analysis that DNA repair genes were most significantly upregulated (Li et al., 2020), suggesting that AKAP8 is involved in the transcription of DNA repair genes and that overexpression in our case, could lead to the downregulation of DNA repair genes and therefore the accumulation of DNA damage. However, although I performed mass spectrometry on both phosphorylated and non-phosphorylated proteins in an attempt to detect differences in the amount of DNA damage and repair proteins, none of such proteins were differentially produced in our model.

Apart from the specific mechanism in which AKAP8 causes DNA damage in cells, other unexplored questions which arise from this research include the mechanism in which AKAP8 is secreted out of the *FBXW7*$^{-/-}$ cell, and how AKAP8 is taken up by the neighbouring wildtype cell. Reexploring the bulk RNAseq results did not reveal any significant difference in the transcription of *AKAP8*, suggesting that the increased AKAP8 found in the media was not a function of increased transcription of AKAP8, but instead could be related to either decreased degradation of AKAP8, or increased extracellular transport of AKAP8. These factors remained to be explored. In addition, only two cell lines were used in this experiment.

There were a number of limitations from my work. Firstly, although an attempt was made to use one cell line which was microsatellite stable and another which was microsatellite unstable, this by no means reflects the molecular diversity of CRC, which could be improved by testing the findings of my research on more cell lines, and ultimately to use patient samples to test for an increased secretion of AKAP8 into the tumour microenvironments of *FBXW7* mutant cells. Furthermore, I generated truncating mutations at R505 instead of the more commonly found R505C missense mutation found in CRC. One recent study using HCT116 cells however suggested that most enriched gene sets were common in cells which either had a homozygous null $FBXW7^{-/-}$ mutation, or a heterozygous $FBXW7^{R505C/+}$ mutation (Thirimanne et al., 2022). In addition, I did not explore the effect of AKAP8-induced DNA damage, which could have explained more robustly the effects of microheterogeneity within subclonal populations of CRC. It is entirely possible that AKAP8-induced DNA damage has minimal effect on the phenotype of a cell, which would therefore limit the potential impact of the *FBXW7* mutant cell on its neighbouring cells.

Regardless, my experiments reveal that wildtype cells neighbouring $FBXW7^{-/-}$ cells accumulate DNA damage via AKAP8 (Figure 22). This unravels a novel mechanism of interaction between subclones in cancer, by which the phenotype of a mutant may be inherited by virtue of being in the microenvironment of a mutant subpopulation. In this case, $FBXW7^{-/-}$ cells which possess an increased burden of DNA damage result in this increased burden being found in neighbouring wildtype *FBXW7* cells. This finding could have significant implications for subclonal interactions. For example, a dominant subclone could potentially imbue neighbouring subclones with phenotypes which would otherwise not have been expected based on the mutational profile of those cells. This could explain why some cancer cells do not respond as expected to targeted treatment, even though those cells possess a mutation which should render that cell susceptible.

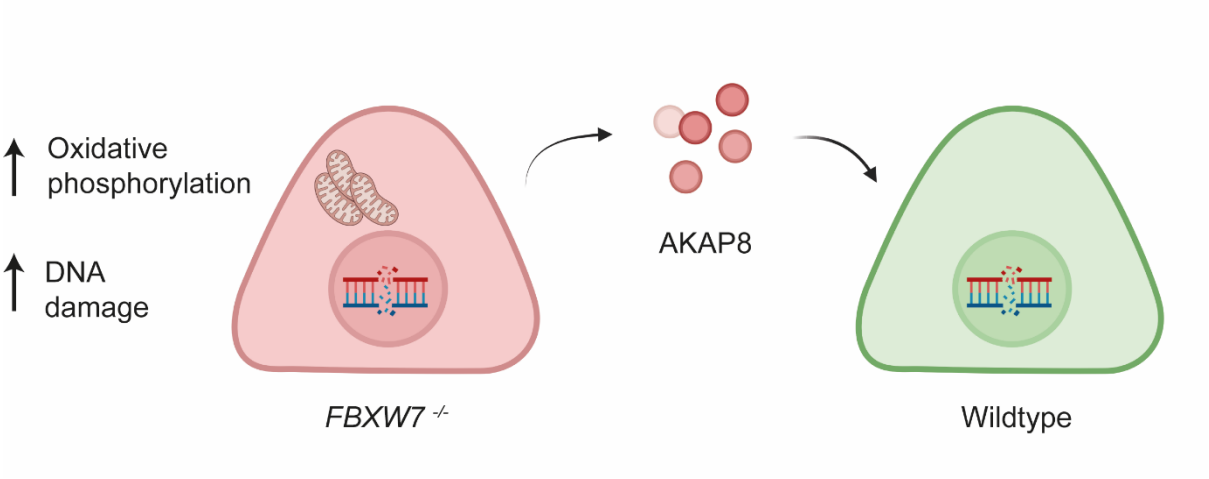


Figure 19: Schematic showing the interaction between FBXW7$^{-/-}$ and wildtype cells.

FBXW7$^{-/-}$ cells demonstrate increased oxidative phosphorylation metabolism as well as an increased accumulation of DNA damage. *FBXW7*$^{-/-}$ cells secrete AKAP8 into its microenvironment, and this is taken up by cells which are wildtype for *FBXW7*. This results in increased DNA damage in wildtype cells.

This finding has another more fundamental implication for cancer evolution. Given that *FBXW7*$^{-/-}$ cells cause DNA damage in neighbouring wildtype cells, this interaction could potentially be a source of ITH in cancer, as neighbouring cells accumulate DNA damage which result in the accumulation of mutations. My findings therefore suggest that further unpicking of mechanisms which exist at the subclonal level could provide major insights into tumour initiation and therapeutic vulnerabilities.

Chapter 5: Mutational order and epistasis determine the consequences of FBXW7 mutations during colorectal cancer evolution

Introduction

In normal colonic epithelium, mutations attributable to eleven signatures, comprising three single base substitutions, four double base substitutions and three indels, accounted for 85% of the mutational burden [45]. Single-base substitutions of note included SBS 1, 5 and 18, all of which are clock-like and exhibit a linear relationship with age. Intriguingly, the mutational burden for all three SBSs differed based on anatomic location along the colon. For SBS1, mean mutation rate across individuals ranged from 16.8 mutations per year (95% confidence interval 15.2 – 18.3) in the ascending colon and cecum, to 12.8 mutations per year (95% CI, 10.6 – 14.9) in the descending and sigmoid colon. Overall, the average mutation rate for SBS 1, 5 and 18 was 43.6 mutations per crypt per year. Mutations in cancer genes were also observed in phenotypically normal colonic epithelium; here termed 'normal drivers' (Lee-Six et al., 2019). These genes included truncating mutations in *AXIN2*, and *STAG2*, and hotspot mutations in *PIK3CA* (E542K, R38H)*, ERRB2* (R678Q, V842I, T862A)*, ERBB3 (*R475W, R667L) and the gene of interest in this study, *FBXW7*. Hotspot mutations in R505C and R658Q were found in *FBXW7*. FBXW7 is a key component of the Skp1-Cdc53/Cullin-F-box-protein complex (SCF/β-TrCP), an E3 ubiquitin ligase responsible for marking proteins by ubiquitination, targeting them for proteasomal degradation (Yeh, Bellon and Nicot, 2018). Substrates of FBXW7 are key regulators of cellular gene transcription, cell cycle progression or activators of signaling pathways, thus making *FBXW7* a key tumour suppressor gene. The significance of *FBXW7* mutations in the context of normal tissues however remains uncharacterized. Given normal driver genes such as *FBXW7* appear to acquire mutations before the acquisition of accepted initiating mutations associated with colorectal cancer such as *APC*, *TP53*, *SMAD4* or *KRAS*, this raises the possibility that normal driver mutations could play a critical role in early cancer evolution.

The detection of driver mutations in the colon is not unique,and sequencing studies in a number of different tissue types has yielded their own set of normal drivers. Martincorena et al. (Martincorena et al., 2015) performed ultradeep targeted sequencing of the coding exons of 74 known skin and other cancer genes in sun-exposed but macroscopically normal skin from patients undergoing blepharoplasty. *NOTCH1* was the most frequently mutated cancer gene observed in normal skin, and was present in 57.1 clones/cm^3 (95% confidence interval 51 to 61/cm^3). *NOTCH1* is a known cancer gene associated with squamous cell carcinoma (SCC). Other mutations present in both normal skin as well as SCC included *NOTCH2*, *NOTCH3*, *TP53* as well as *FAT1*. Overall, the presence of driver mutations appeared to confer a slight advantage in terms of average clone sizes. In comparison to selectively neutral average clone sizes which measured 15mm^2 ($CI_{95\%}$ 0.13 to 0.17), clone sizes for *NOTCH1*, and *FGFR3* were 0.23mm^2 (q= 0.0002) and 0.69mm^2 (q= 0.0007) respectively.

A similar protocol used for human skin was applied to the human oesophagus (Martincorena et al., 2018). Through ultradeep sequencing of the same 74 cancer genes, 14 cancer genes were discovered to be present in histologically normal human oesophagus. Closer examination of each genetic mutation yields unique clues to the role of such driver mutations in normal epithelium. In the case of *NOTCH1*, the mutation is present in 30 – 80% of normal oesophageal epithelium, but present in only 10% of oesophageal squamous cell carcinoma (oSCC) (Song et al., 2014). Conversely, *TP53* is present in more than 95% of oSCC but present in only a few cells in normal oesophageal epithelium.

The landscape of driver mutations in normal endometrium has also been demonstrated by Moore et al. (Moore et al., 2020). Twelve genes were found to be positively selected during whole genome sequencing of individual endometrial glands. By constructing phylogenetic trees and comparing mutational patterns between clades, the authors were able to determine that driver mutations in the endometrium are likely to have occurred in the first decades of life, thereafter demonstrating an age-related increase in the burden of such driver mutations.

Returning back to the colon, I propose three possibilities which may account for the role of driver mutations in normal colonic epithelium, and which are explored in my experiments in this Theme.

1. The first possibility is that these mutations are the earliest molecular changes which sets a cell on its path towards cancer. Given that such mutations occur in only about 1% of histologically normal crypts, uncovering these mutations has only been made possible recently by technologies such as whole genome sequencing. I believe that this hypothesis is unlikely given that phylogenetic trees of mutations have not demonstrated these mutations as being clonal.

2. A second possibility is the converse, that these driver mutations have no bearing or relationship to cancer. These driver mutations may result in clonal expansions, but perhaps not to the point that cancer is formed. In fact, there is evidence that a diffusion-like process may accommodate crypt expansions to a point without altering the gross histology (Olpe et al., 2021). This was observed when comparing *KDM6A* and *STAG2* mutant clones with *KRAS* mutant clones. Being a driver mutation, *KRAS* mutant clones resulted in a 17-fold increase in crypt fission, whereas *KDM6A* and *STAG2* mutants increased fission by 5- and 3-fold respectively. The authors calculated a 12-fold threshold beyond which crypt architecture could no longer diffuse to accommodate the increased fission rates, resulting in a morphologic change in tissue architecture. It is therefore possible that such driver mutations in normal tissue, while exerting a positive selection effect, have no relationship to carcinogenesis.

This raises the possibility that the timing of a mutation is therefore critical. *FBXW7* is a known CRC gene but is often found to be subclonal. In the presence of an *APC* mutation, the presence of *FBXW7* could lend a cumulative proliferative advantage which might have been

abrogated had *FBXW7* occurred first. There exists evidence for such phenomena in the literature, albeit not in CRC. In adrenocortical tumours, overexpression of oncogenic Ras before loss of p53 results in highly malignant tumours whereas the reverse order gives rise to benign tumours only (Herbet et al., 2012). In chronic myeloproliferative neoplasms, disease symptoms characteristics could be stratified between patients who had *JAK2*-first versus *TET2*-first mutations (Kent and Green, 2017).

3. A third possibility is that driver mutations in normal epithelium exert an indirect but crucial collaborative effect in colorectal carcinogenesis. Evidence for clonal cooperation have been described. Mutant KRAS is resistant to cetuximab due to the production of TGFα and amphiregulin, which is also secreted out from cells. When media from KRAS mutant cells were transferred to wild-type KRAS cells, cetuximab was observed to be less effective than expected (Hobor et al., 2014). Direct interactions between subclones have also been observed, albeit not in CRC. In a mouse model of small cell lung cancer (SCLC), two clonal populations were observed to cause increased cellular proliferation (Calbo et al., 2011). One population expressed neuroendocrine markers such as synaptophysin (Syp), achaete-scute complex homolog 1(Ash1), and neural cell adhesion molecule (NCAM), while the other population expressed mesenchymal markers and were low in NCAM. *In vitro* co-cultures of both clones resorted in much greater proliferation of cancer tissue than individually grown clones. Moreover, phenotypic differences were seen between co-cultured and non-co-cultured cells, with co-cultured cells more accurately resembling the cancer phenotype. In addition, when cells were transferred in vivo into a mouse model, abdominal metastases were prevalent in the co-cultured cohort.

Two recent articles have demonstrated clonal competition in the key Wnt pathway in CRC. In an article by van Neerven et al., a negative feedback loop in APC mutant organoids involving an upregulation of WNT antagonist genes such as Notum, Wif1, and Dkk2 was observed (van

Neerven et al., 2021b). The secretion of these Wnt antagonists resulted in Wnt-pathway downregulation amongst wild-type cells. Interestingly, this effect could be reversed when Wnt was reactivated using GSK3β inhibitors. The competitive inhibition of wild-type cells by APC mutants was confirmed in another article by Flanagan et al. using a mouse model instead (Flanagan et al., 2021a). NOTUM inhibitor treatment resulted in reduced fixation of APC mutants.

Notably, the *FBXW7* mutation introduced in my work differs from the hotspot mutation at R505C and R658Q observed in normal colonic tissue. This was due to technical limitations originating from the CRISPR strategy adopted. To achieve a high CRISPR targeting efficacy, I utilized a multi-guide approach which necessitated the presence of three different sgRNA sequences, therefore requiring three corresponding Cas9 PAM sequences in close proximity to each other (Chan, Collins and Buczacki, 2023). As such, my CRISPR strategy targeted a region at exon 2. Based on the "just enough" hypothesis, I could predict that the effect of generating bi-allelic loss of FBXW7 activity is likely to result in an excessive reduction in SCF complex efficacy to render the cellular phenotype less than ideal for cellular proliferation. This appears to correspond to previous work in mice using gut specific knockout of Fbxw7 which resulted in truncating mutations at exon 5, disrupting the F-box, and similarly resulting in complete loss of functional protein (Sancho et al., 2010). Here, the authors noted that this model of Fbxw7 mutation did not result in increased polyposis in the intestines of mice, and required concurrent loss of Apc for polyps to be formed, findings which were recapitulated in my experiments. Overall, while my model of *FBXW7* mutation differed from what is found *in vivo*, my focus on the effects of the *FBXW7* mutation on epistasis and mutational order remains a novel finding. I believe that changing the *FBXW7* mutation could affect the magnitude of the transcriptional consequence described in my work, and could form the basis for future studies.

A key component of my experiments in this theme concern the effects of epigenetic plasticity in cancer evolution. In models of early pancreatic cancer, *KRAS*-mutant cells were shown to

demonstrate distinct chromatin accessibility patterns which could predict divergence into either a benign or malignant fate (Burdziak et al., 2023). In colorectal cancer, the impact of epigenetic changes on tumour initiation are only beginning to be uncovered. Johnstone et al. highlighted the importance of the three-dimensional architecture, in the form of chromatin loops, topologically activating domains and large-scale compartments, of DNA in restricting malignant progression (Johnstone et al., 2020). A more recent study by Heide et al. found chromatin modifier genes are actively mutated in the early initiation of colorectal cancer, and can result in the accumulation of further genetic mutations which herald carcinogenesis (Heide et al., 2022). Specific to my experiments, I was able to establish that epigenetic plasticity resulted in a shift of the adult cell towards a fetal state. This fetal intestinal cell state remains hitherto poorly defined in the literature and appears limited to a state of unlimited cellular regeneration induced during cellular injury or irradiation (Fernandez Vallone et al., 2016, Mustata et al., 2013), and in cancer, has been purported to result in a cancer stem cell state owing to its regenerative and proliferative abilities.

In this study, I used adult human colonic organoids and CRISPR-Cas9 gene editing to explore the cell autonomous and gene-interactive effects of an *FBXW7* mutation. Combining gene editing with organoid culture fashions a powerful tool ideal for studying the functional consequences of driver mutations in normal tissue for it allows specific mutations to be introduced into tissue devoid of genetic perturbation. By sequentially introducing multiple distinct mutations in the same cell, I also aimed to determine relationships arising between cells with an *FBXW7* mutation and the acquisition of subsequent cancer driver mutations. I found that cells with an *FBXW7* deletion were transcriptionally similar to wildtype cells but subtly primed towards a foetal/YAP-active state. Further, I observed that an initial *FBXW7* deletion dampened the transcriptional effect of a subsequent *APC* mutation also resulting in retention of a near-wildtype phenotype. Whereas when I reversed this order of mutations, despite identical genotype, I generated *bona fide* cancer organoids. I uncovered this profound mutational order-based cellular plasticity to be driven by epigenetic differences in both global

and local chromatin accessibility. Here I provide the first direct eukaryotic evidence for the existence of epistasis, and that order of mutations, as much as the mutations themselves, are key in determining their effect.

Finally, it is crucial to acknowledge that the *FBXW7* deletion introduced in my work differs from the hotspot mutation at R505C and R658Q observed in normal colonic tissue. This was due to technical limitations originating from the CRISPR strategy adopted. To achieve a high CRISPR targeting efficacy, I utilized a multi-guide approach which necessitated the presence of three different sgRNA sequences, therefore requiring three corresponding Cas9 PAM sequences in close proximity to each other (Chan, Collins and Buczacki, 2023). As such, my CRISPR strategy targeted a region at exon 2. I discuss the implications of my *FBXW7* mutant model in the Discussion section of this Chapter.

Results

Mutant *FBXW7*$^{-/-}$ recapitulate known effects on downstream targets

Patient-derived adult stem cell wildtype organoids (W) were generated from phenotypically normal colon tissue of patients undergoing surgery for colorectal cancer. CRISPR-Cas9 editing was used to generate *FBXW7*$^{-/-}$ organoids (F) (Chan, Collins and Buczacki, 2023). Our gene editing strategy utilizes three sgRNAs to introduce large deletions at exon 2 of the *FBXW7* gene. Following CRISPR-Cas9 gene editing, successful gene editing was confirmed using Sanger sequencing between PCR primers as stated in the Methods found in Chapter 3. F and W organoids possessed similar morphology (Fig. 23a). I validated knockout of the *FBXW7* gene using western blot (Fig. 23b), as well as immunofluorescence using anti-FBXW7 antibodies (Fig. 23c). I also performed western blot on the substrate targets of FBXW7: p-c-Jun, cyclin E, p-c-Myc, and Notch intracellular domain (NICD) and showed upregulation of all protein substrates on knockout of FBXW7 (Fig. 23d).

To characterise transcriptional differences between F and W organoids, bulk-RNA sequencing (RNA-seq) on both organoid types was performed (Fig. 23e). Using a cut-off of $\log_2FC \leq -1.5$ or ≥ 1.5, and adjusted *p*-value < 0.05, F and W were transcriptionally similar with only two genes, *HLA-DQB1* and *GSTM1*, found to be differentially expressed.

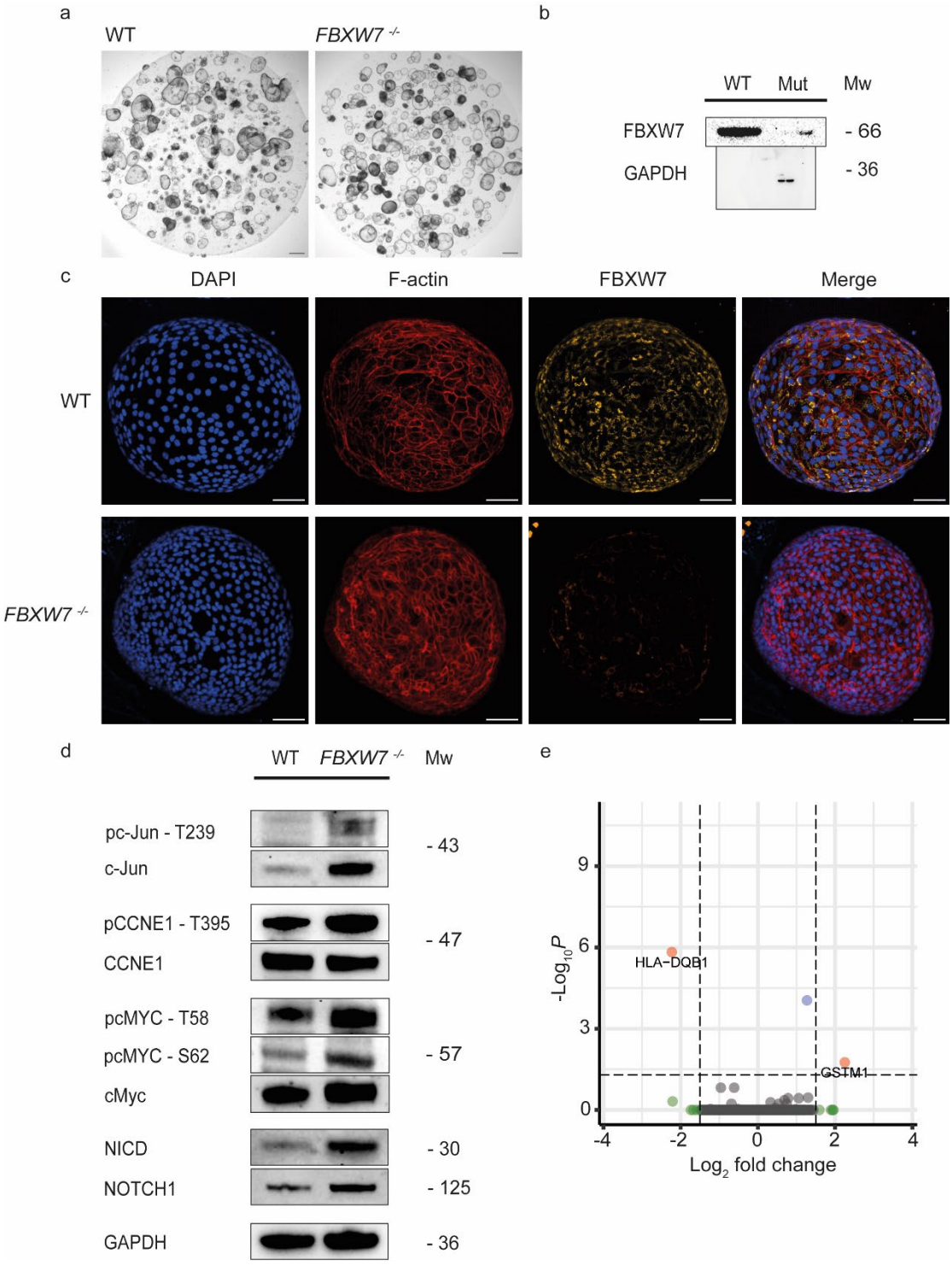

a
WT
FBXW7 -/-
b
WT
Mut
Mw
FBXW7
- 66
GAPDH
- 36
c
DAPI
F-actin
FBXW7
Merge
WT
FBXW7 -/-
d
WT
FBXW7 -/-
Mw
pc-Jun - T239
c-Jun
- 43
pCCNE1 - T395
CCNE1
- 47
pcMYC - T58
pcMYC - S62
cMyc
- 57
NICD
- 30
NOTCH1
- 125
GAPDH
- 36
e
-Log10P
9
6
3
0
HLA-DQB1
GSTM1
-4
-2
0
2
4
Log2 fold change

Figure 20: Characterisation of W and F organoids.

a. Brightfield microscopy of W and F organoids revealed no observable phenotypic differences. Scale bars represent 500μm. b. Western blot validation showing loss of FBXW7 protein in F organoids. c. Immunofluorescence of W and F organoids with DAPI nuclear stain (blue), F-actin (red), FBXW7 (orange). Scale bars represent 100μm. d. FBXW7 targets the phosphorylated substrates and ubiquitinates these substrates for proteasomal degradation. This included phosphorylated cJun, phosphorylated CCNE1, phosphosylated cMyc, and notch intracellular domain (NICD). The non-phosphorylated substrates were affected to varying degrees. While upregulation of cJun, was observed, there was no change in the quantities of CCNE1, cMYC and NOTCH. e. Volcano plot from bulk RNAseq of F vs W organoids revealed minimal differential expression of genes. Only HLA-DQB1 was found to be significantly downregulated, and GSTM1 significantly upregulated in F organoids. All experiments were performed with N = 3 biological replicates.

FBXW7$^{-/-}$ mutants do not exhibit a competitive proliferative advantage over wildtype organoids

Given the transcriptional similarities between W and F organoids, I interrogated the possibility of a competitive or cooperative effect by co-culturing F organoids with W organoids. This follows from the earlier work in CRC cell lines presented in this thesis where I had shown DNA damage via an AKAP8-mediated mechanism in wildtype cells adjacent to *FBXW7*$^{-/-}$ cells (Chan et al., 2023). Similarly, murine *Apc*$^{-/-}$ intestinal organoids co-cultured with wildtype organoids have been shown to demonstrate increased proliferation as a result of a supercompetitor effect driven by secretion of the Wnt-antagonist NOTUM from *Apc*$^{-/-}$ organoids (van Neerven et al., 2021b). As a positive control for our organoid co-culture model, I therefore generated human *APC*$^{-/-}$ organoids (A) and performed lentiviral transduction of A organoids with pCDH-CMV-mCherry-T2A-Puro (Addgene #72264). W organoids underwent lentiviral transduction with pCDH-EF1-copGFP-T2A-Puro (Addgene #72263). A and W organoids were co-cultured in equal proportion, and fluorescent imaging was performed on days 1, 3, 5 and 7 (Figure 24a). In A and W co-cultured organoids, mCherry-labelled A organoids demonstrated a significant proliferative advantage in concordance with previously described results performed in murine organoids (Figures 24a, 24c and 24d) (van Neerven et al., 2021b). F organoids were co-cultured with WT organoids in a similar fashion. F organoids underwent lentiviral transduction with pCDH-CMV-mCherry-T2A-Puro (Addgene #72264). In F and W co-cultured organoids, proliferation in both F and WT organoids did not exhibit any morphological differences (Figures 24b, 24e and 24f). This was further confirmed by an EdU proliferation assay which showed no difference in EdU positive cells between F and W organoids grown separately (Figure 24g).

I considered whether F mutants might be inducing a transcriptional response on co-cultured W organoids. In F and WT co-cultured organoids, I therefore performed fluorescent activated cell sorting (FACS) to isolate GFP and mCherry positive cells, representing W and F cells respectively. Sorted cells were submitted for bulk-RNAseq, and the transcriptome compared

to non-co-cultured cells. Again, I observed that transcriptomes between co-cultured W and non-co-cultured W (Figure 24h), and co-cultured F and non-co-cultured F organoids (Figure 24i) exhibited minimal differences. Taken collectively, these findings suggest that there is minimal effect of F organoids on neighbouring W cells, in contrast to the earlier findings in cell lines (Chan et al., 2023).

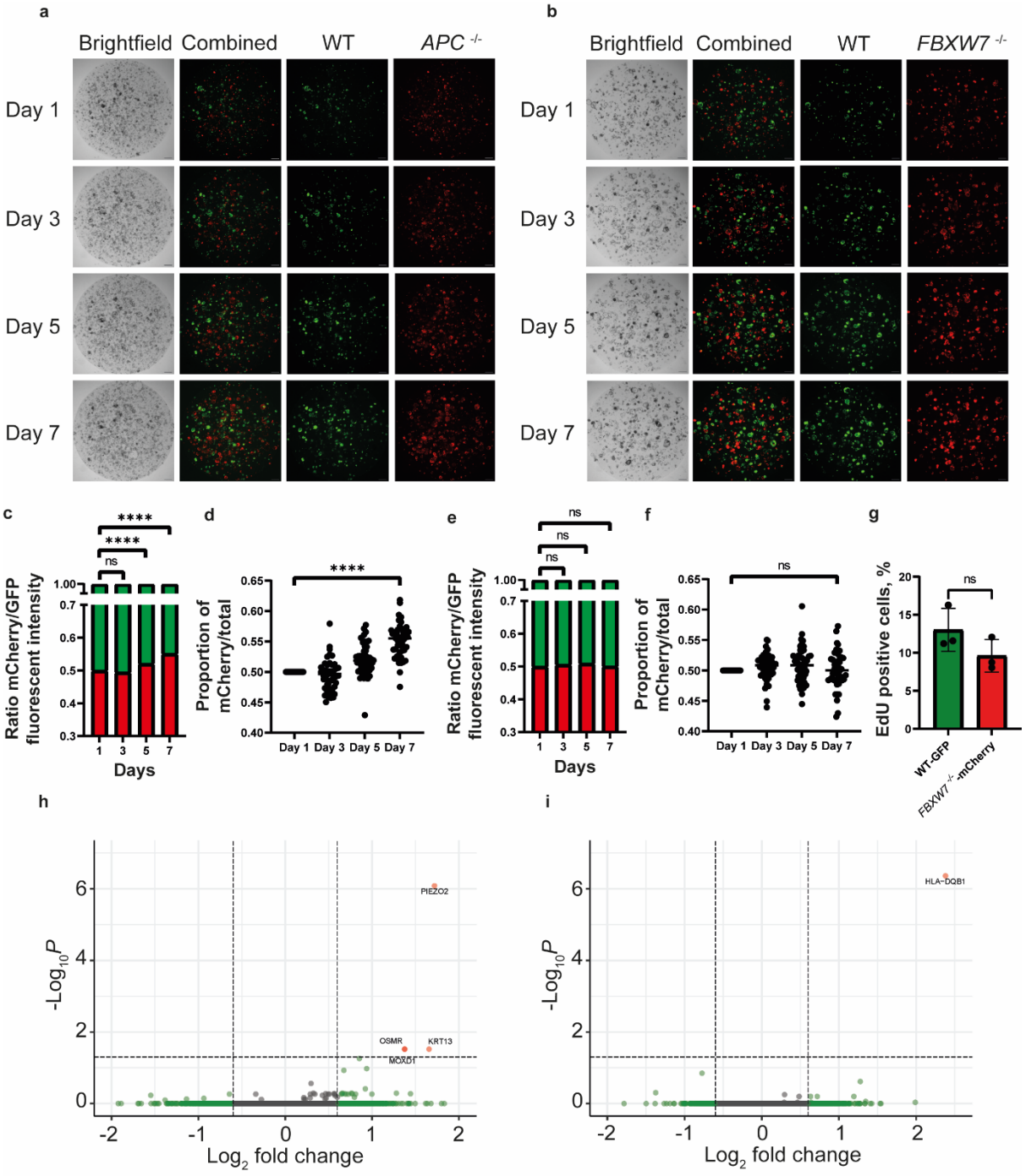

a
Brightfield
Combined
WT
APC -/-
Day 1
Day 3
Day 5
Day 7
b
Brightfield
Combined
WT
FBXW7 -/-
Day 1
Day 3
Day 5
Day 7
c
Ratio mCherry/GFP fluorescent intensity
Days
****
****
ns
d
Proportion of mCherry/total
Day 1 Day 3 Day 5 Day 7
****
e
Ratio mCherry/GFP fluorescent intensity
Days
ns
ns
ns
f
Proportion of mCherry/total
Day 1 Day 3 Day 5 Day 7
ns
g
EdU positive cells, %
ns
WT-GFP
FBXW7 -/- -mCherry
h
-Log10P
Log2 fold change
PIEZO2
OSMR
KRT13
MCXD1
i
-Log10P
Log2 fold change
HLA-DQB1

Figure 21: FBXW7$^{-/-}$ organoids do not influence growth of neighbouring wildtype organoids. *APC*$^{-/-}$ cells act as supercompetitors which suppress the growth of neighbouring wildtype cells, evidenced by proliferative advantage of A organoids over neighboring W organoids (a, c, d). A similar effect was not seen in F organoids co-cultured with W organoids (b, e, f). g. EdU assay for W and F organoids showed no significant differences in proliferation rate. Volcano plot from bulk RNAseq of co-cultured W versus non-co-cultured W (h) and co-cultured F versus non-co-cultured F (i) organoids showed that the effect of co-culture did not significantly influence the transcriptome. The values in this figure are mean ± sd, and statistical significance was measured by unpaired t-test. All experiments were performed with N = 3 biological replicates. A: *APC*$^{-/-}$ cells, W: wildtype cells, F: *FBXW7*$^{-/-}$ cells

FBXW7 mutation provides evidence of epistasis

In light of the lack of a cooperative/competitive interaction with wildtype cells, and building on the earlier finding that F organoids were transcriptionally similar to W organoids, an organoid model mirroring the polypoidal and cancerous phases of CRC carcinogenesis was generated (Figure 25a). An $APC^{-/-}$ mutation introduced into W organoids generated A organoids, representing the polypoidal, precancerous phase, while a $TP53^{-/-}$ mutation introduced into A organoids, therefore possessing a double knockout $APC^{-/-}$ / $TP53^{-/-}$ mutation, generated AT organoids and representing the cancerous phase. On each of these A and AT organoids, an *FBXW7* mutation was then introduced, generating AF and ATF organoids respectively (Figure 25b).

As previously described, A organoids were morphologically distinct from W organoids (Drost et al., 2015, Matano et al., 2015). A organoids grew large and cystic. The subsequent acquisition of a *TP53* mutation preserved this phenotype. Importantly, an additional *FBXW7* deletion on both the A and AT organoids also maintained this cystic morphology. To further characterise transcriptional differences between these different organoids, A, AF, AT and ATF organoids were submitted for bulk-RNAseq. A organoids were compared with AF organoids, while AT organoids were compared with ATF organoids. An epistatic effect of the *FBXW7* deletion in our comparison of various groups was observed. Together with the earlier comparison between F and W organoids which saw only 2 differentially expressed genes, I observed 87 differentially expressed genes between A and AF, and 294 between AT and ATF (Figures 25c, 25d, and 25e). In addition, between A and W, and AT and W, there were 420 and 455 differentially expressed genes respectively, highlighting the profound differences between introducing a mutation in *FBXW7* compared with other genes. Hierarchical clustering of these different groups showed that F and W organoids clustered together, while A, AF and ATF organoids formed separate clusters (Figure 25f). These results demonstrate that the same *FBXW7* deletion can have varying effects depending on its mutational background. More importantly, it suggests that an *FBXW7* deletion may have different

cellular effects depending on when the mutation was acquired, and that an early mutation will have fewer effects compared with a later mutation.

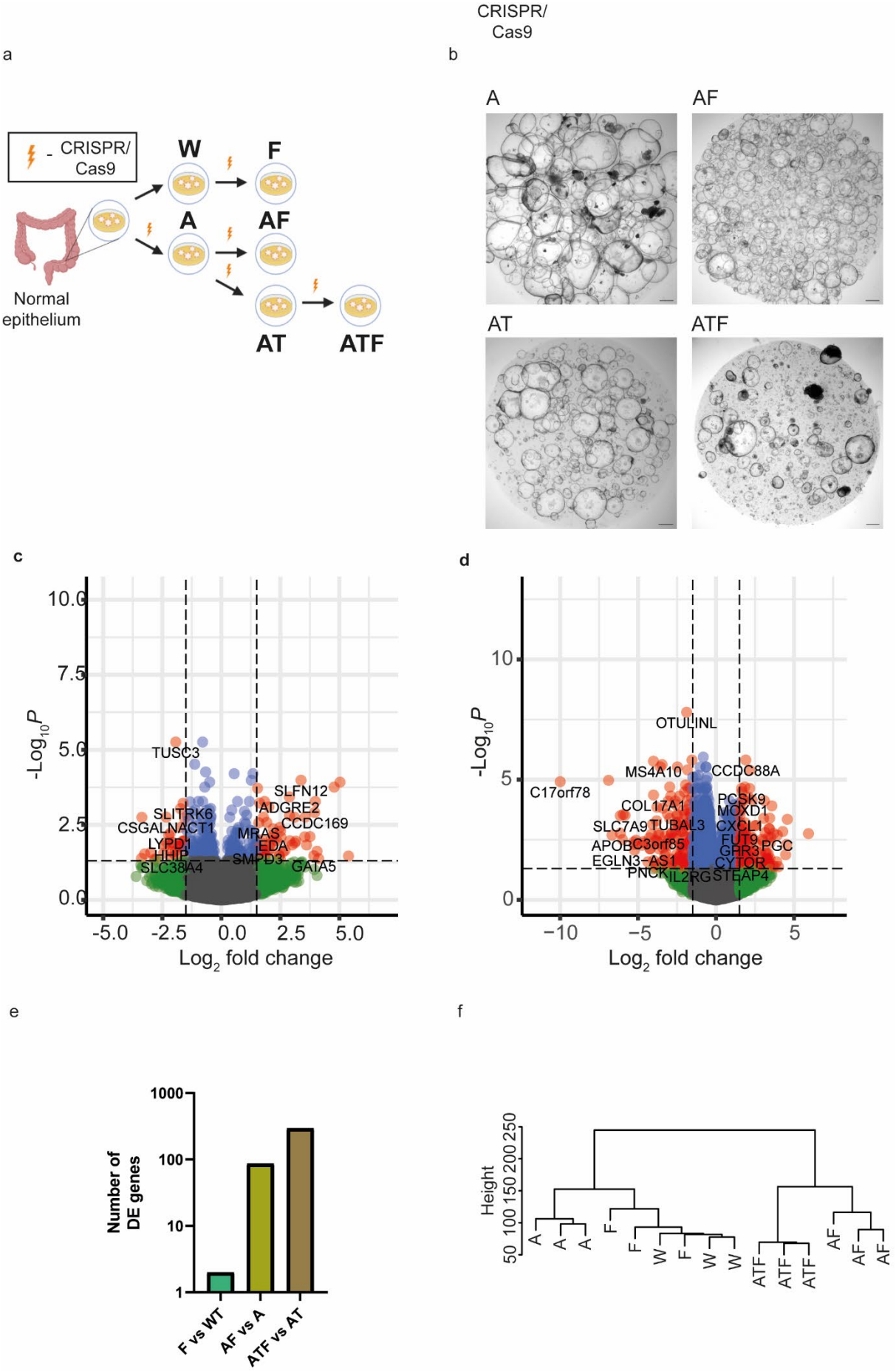

a
CRISPR/
Cas9
W
F
A
AF
AT
ATF
Normal
epithelium
b
CRISPR/
Cas9
A
AF
AT
ATF
c
-Log10P
Log2 fold change
TUSC3
SLFN12
ADGRE2
CCDC169
MRAS
EDA
SMPD3
GATA5
SLITRK6
CSGALNACT1
LYPD1
HHIP
SLC38A4
d
-Log10P
Log2 fold change
OTULINL
MS4A10
CCDC88A
C17orf78
COL17A1
PCSK9
MOXD1
SLC7A9
TUBAL3
CXCL1
FUT9
APOBC3orf85
GHR
PGC
EGLN3-AS1
CYTOR
PNCK
IL2RG
STEAP4
e
Number of
DE genes
F vs WT
AF vs A
ATF vs AT
f
Height

Figure 22: Effect of FBXW7 mutation is dependent on the mutational background.

a. Schematic showing the organoid models used in this analysis. W, A and AT organoids were generated to reflect normal, adenomatous and carcinomatous lesions. On each of these organoid types, an *FBXW7* deletion was introduced to generate F, AF and ATF organoids respectively. b. Brightfield microscope of A, AF, AT and ATF organoids. A organoids adopt a cystic phenotype which is representative of increased proliferation. AF organoids appear smaller in size compared with A organoids. AT organoids also recapitulate the cystic appearance of A organoids. ATF organoids appear less cystic than AT organoids. c. Volcano plot from bulk RNAseq of AF vs A organoids. d. Volcano plot from bulk RNAseq of ATF vs AT organoids. e. The number of differentially expressed genes on the addition of an *FBXW7* deletion changes based on the mutational background. While 2 genes were differentially expressed between W and F, 87 and 294 genes were differentially expressed between A and AF, and AT and ATF respectively. f. Hierarchical clustering of the different organoid types showed that W and F organoids clustered most similarly together. All experiments were performed with N = 3 biological replicates. W: wildtype cells, F: *FBXW7*$^{-/-}$ cells, A: *APC*$^{-/-}$ cells, AF: *FBXW7*$^{-/-}$/*APC*$^{-/-}$ cells, AT: *APC*$^{-/-}$/*TP53*$^{-/-}$ cells, ATF: *APC*$^{-/-}$/*TP53*$^{-/-}$/*FBXW7*$^{-/-}$ cells

To ensure observations were not an artifact of repeated iterations of gene editing, the AAVS1 safe harbour locus was targetted three times to replicate the triple-mutant ATF organoids. Morphologically, AAVS1-targetted organoids resembled wildtype organoids (Figure 26a). Bulk-RNAseq of AAVS1-targetted organoids compared with W organoids showed only two differentially expressed genes with a cut-off of $\log_2FC \leq -1.5$ or ≥ 1.5, and adjusted p-value < 0.05 (Figure 26b), greatly differing from the 493 differentially expressed genes between ATF and W organoids (Figure 26c). These findings confirm that epistatic effects described earlier were not simply a function of repeated iterations of CRISPR-Cas9 gene editing.

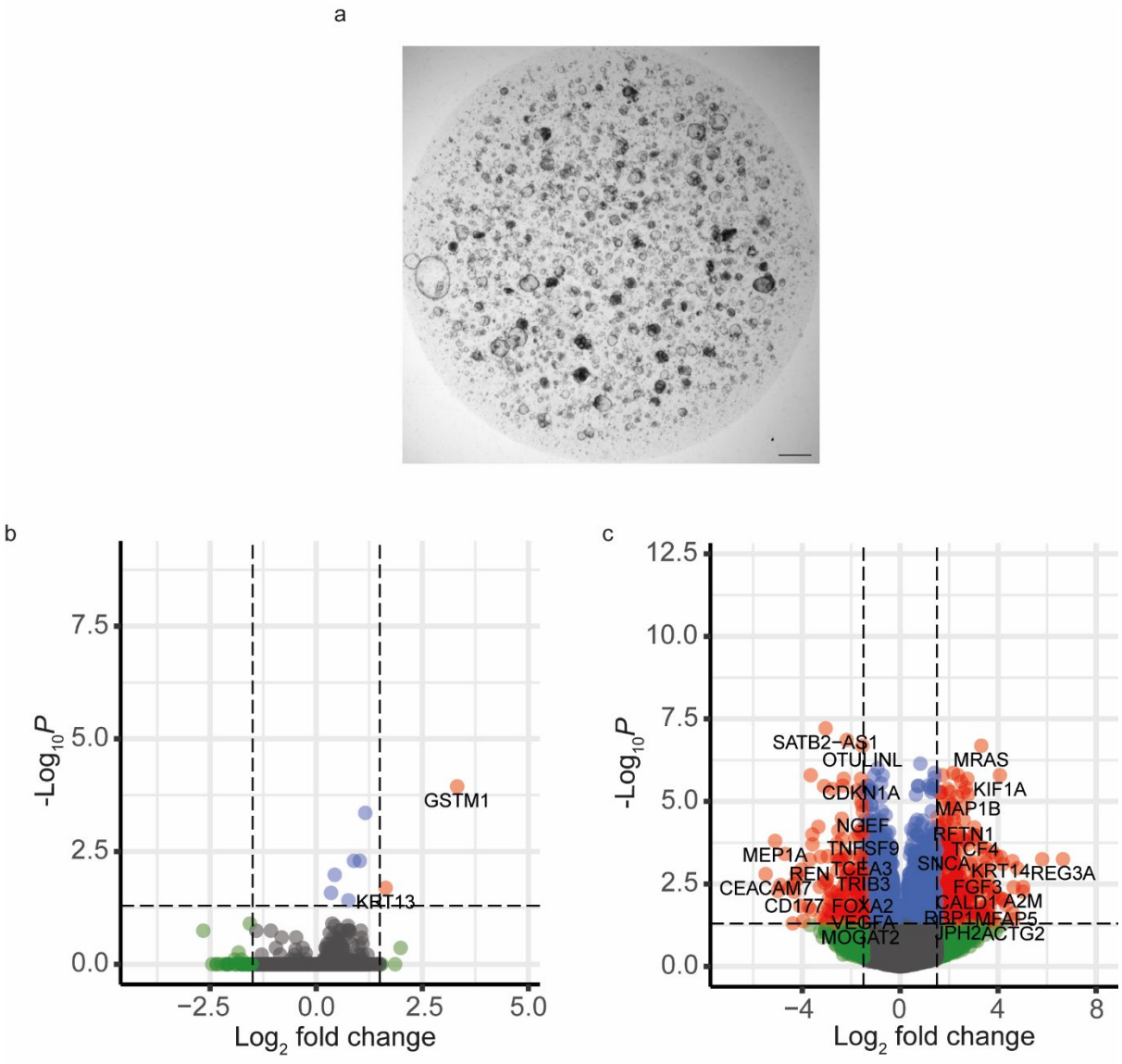


Figure 23: Effect of epistasis not an artefact of CRISPR/Cas9 gene editing.

a. Brightfield microscopy of organoids which had been gene edited at the AAVS1 loci three times to mimic triple-mutant ATF organoids. AAVS1 organoids did not adopt a cystic phenotype. b. Volcano plot from bulk RNAseq of AAVS1 vs W organoids. c. Volcano plot from bulk RNAseq of ATF vs W organoids. All experiments were performed with N = 3 biological replicates. W: wildtype cells, ATF: *APC*$^{-/-}$*/TP53*$^{-/-}$*/FBXW7*$^{-/-}$ cells

Order of *FBXW7* mutation determines its transcriptional effect

To ascertain the transcriptional effect of the order of mutation, I designed an experiment in which organoids possessed the same mutations, but differed in the sequence in which the mutations were introduced. As previously described, AF organoids possessed an *APC* mutation, followed by an *FBXW7* deletion. I generated FA organoids, which acquired an *FBXW7* deletion prior to the *APC* mutation (Figure 27a). Morphologically, FA organoids did not acquire the cystic phenotype characteristic of A or AF organoids (Figure 27b). I submitted AF and FA organoids for bulk-RNAseq. Principal component analyses of AF and FA organoids gene expression, together with W, A and F organoids revealed that FA organoids did not cluster with AF organoids (Figure 27c). Instead, FA organoids clustered with F and W organoids, while AF organoids clustered with A organoids. To further analyse the transcriptional effect of changing the order of mutations, I used a cut-off of $\log_2FC \leq -0.7$ or ≥ 0.7 to identify differentially expressed genes when AF and FA organoids groups were compared with wildtype organoids. This cut-off selected for 3335 genes in the AF group and 861 genes in the FA group (Figure 27d). As expected, there was significant overlap in the genes which were differentially expressed in both groups. In the FA group, 628 (72.9%) genes were also differentially expressed in the AF organoids. Further analyses of the gene expression changes for these 628 genes suggested a dampening effect of an earlier *FBXW7* deletion, such that the magnitude of the transcriptional perturbation was rendered less positive or negative (Figure 27e). Using GSEA, I also observed that some key signaling pathways involved in carcinogenesis such as the EMT pathway were significantly more active in FA compared with AF organoids (Normalised enrichment score (NES) 2.21, adjusted *p*-value= 1.70308×10^{-10}). Collectively, these findings promulgate our observation that mutational background affects the transcriptional effect of a mutation. Although our model adopted an *FBXW7* deletion instead of the *FBXW7* mutation more commonly found in normal tissue and cancer, our findings raise the possibility that *FBXW7* mutations in normal tissue may protect

against the future effect of an *APC* mutation, and by extension, could be protective against the initiation of CRC.

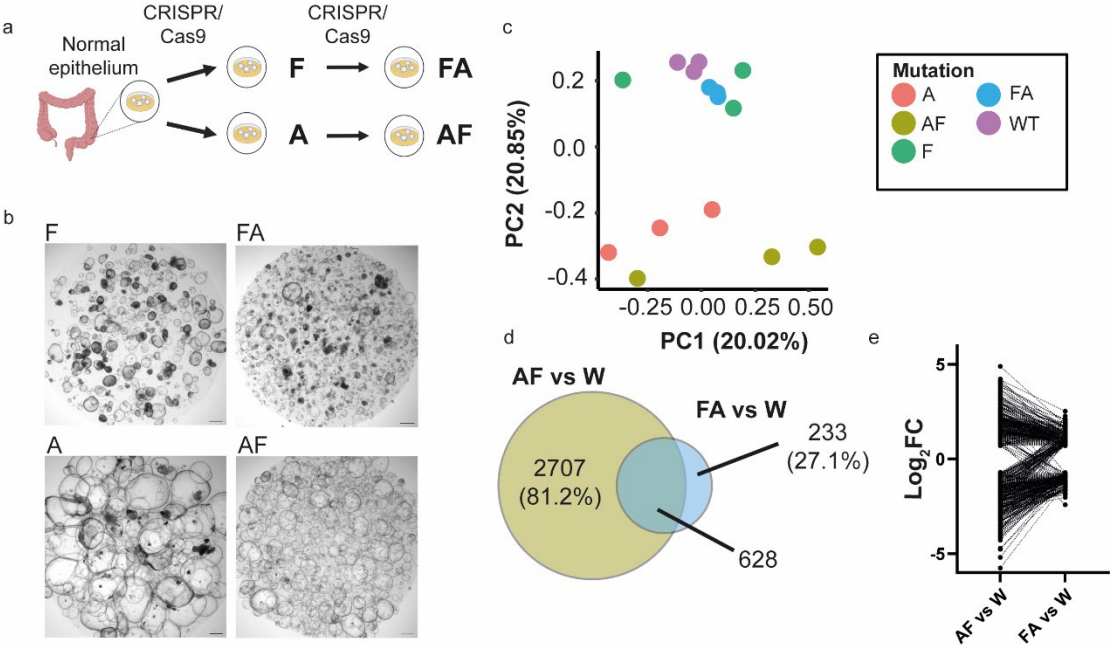


Figure 24: Effect of FBXW7 mutation is dependent on the order of mutation.

a. Schematic showing the organoid models used in this analysis. A and F organoids were generated separately. On the A organoid, an *FBXW7* mutation was introduced, generating AF organoids. On the F organoid, an *APC* mutation was introduced, generating FA organoids. AF and FA organoids possessed the same mutations but acquired the mutations in an inverse order. b. Brightfield microscope of A, AF, F and FA organoids. Although FA organoids acquired an *APC* mutation, this did not recapitulate the cystic phenotype observed in A or AF organoids. c. Principle-component analysis (PCA) showed not only that FA organoids clustered separately from AF organoids, FA organoids clustered together with W and F organoids, while AF organoids clustered closer to A organoids. d. Venn diagram depicting the overlap of differentially expressed genes at cut-off log2FC ≤ -0.7 or ≥ 0.7 between AF vs W and FA vs W comparisons. e. Analysis of these 628 genes showed there was a reduction in the magnitude by which the genes were differentially expressed, suggesting a dampening effect of changing the order of mutation on the transcriptome. All experiments were performed with N = 3 biological replicates. W: wildtype cells, FA: *FBXW7*$^{-/-}$/*APC*$^{-/-}$ cells, AF: *FBXW7*$^{-/-}$/*APC*$^{-/-}$ cells

scRNAseq data suggests dedifferentiation towards a fetal state in *FBXW7* mutant organoids

I endeavored to explain our finding that an *FBXW7* deletion was protective to consequences of a later acquisition of an *APC* mutation, considering our observation that mutational order affects transcriptomic expression. I hypothesized that subtle transcriptomic changes could be present in F organoids in small subpopulations of cells which might be masked by bulking during RNAseq. To overcome this, I performed single-cell RNA sequencing (scRNAseq) on F and W organoids. Clustering by Uniform Manifold Approximation and Projection (UMAP) showed 14 clusters (Fig. 28a). Notably cluster 13 was unique to *FBXW7* mutant cells (Fig. 28b). This cluster was marked by genes such as *KRT17* (log_2FC= -1.65, p= 2.52 x 10^{-35}), *TAGLN* (log_2FC= -1.33, p= 2.81 x 10^{-35}), and *MMP7* (log_2FC= -1.47, p= 7.70 x 10^{-34}) and an EMT signature (Figure 29). Importantly, I observed that wildtype cells expressed *FBXW7*, indicating that this gene is expressed in normal epithelium (Fig. 28c). Analysis by RNA velocity indicated that cluster 13 was an important terminal differentiation state, which was only present in *FBXW7* deleted cells (Fig. 28d). As expected, quantitative analysis between UMAP plots from F and W organoids demonstrated a statistically significant upregulation of the FA gene signature in F organoids relative to W organoids (Fig. 28e). I also observed increased expression of fetal gene signature (Fernandez Vallone et al., 2016) (Fig. 28f), and the YAP pathway gene signature (Gregorieff et al., 2015) (Fig. 28g) in F organoids. There was also a decreased expression of the intestinal stem cell signature (ISC) (Merlos-Suárez et al., 2011) (Fig. 28h), and decreased expression of the crypt-based columnar (CBC) stem cell signature (Vasquez et al., 2022) (Fig. 28i) in F organoids. I re-analysed bulk RNAseq data comparing F and W organoids using gene set enrichment analysis and found that fetal (Figure 30a) and YAP (Figure 30b) signatures correlated with our scRNAseq results. I did not observe a significant downregulation of the ISC (Figure 30c) and CBC (Figure 30d) signatures in bulk RNAseq as in scRNAseq. Likewise, I did not observe an FA gene signature in our bulk RNAseq comparison of F versus W (Figure 30e), which could be a function of minimal transcriptional perturbation observed in F organoids compared with W organoids. To

further explore the YAP pathway, I performed additional GSEA on FA vs W and AF vs W organoids. Intriguingly, this showed greater enrichment of the YAP signature in FA vs W (ES= 0.58) (Figure 31a) compared to AF vs W (ES= 0.36) (Figure 31b). Taken together, these findings point towards a reordering of the stem cell hierarchy, favouring pathways which recall a fetal regenerative/YAP active cell state.

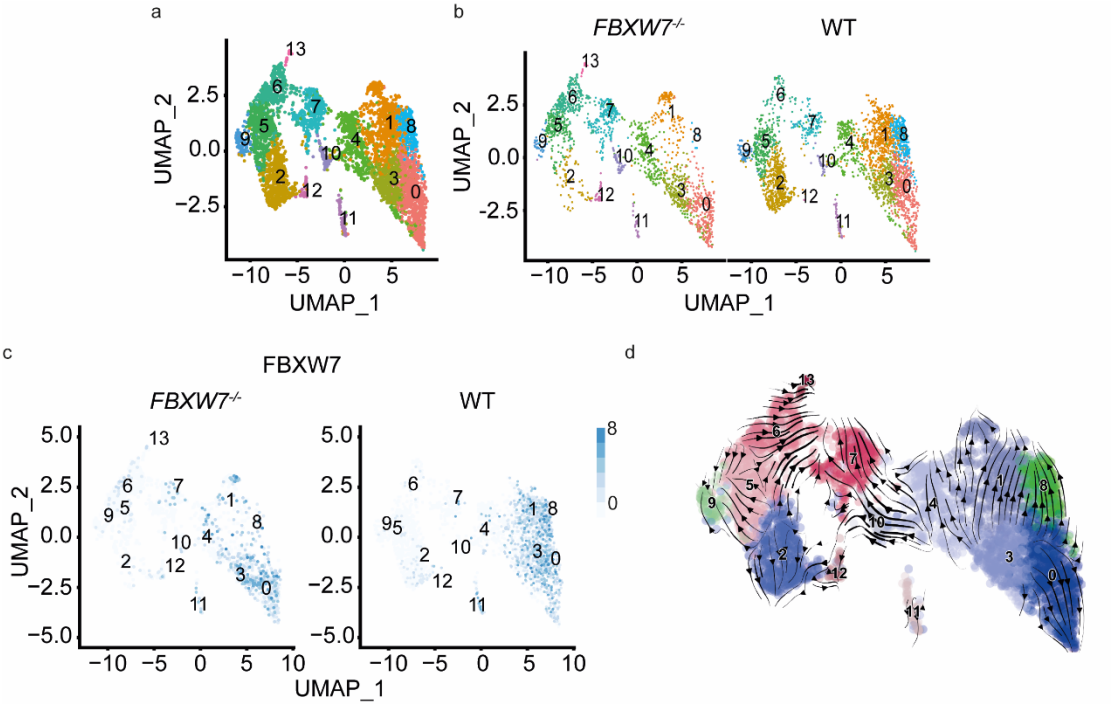


Figure 25: scRNAseq of W and F organoids

a. Uniform manifold approximation and projection (UMAP) analysis from scRNAseq of W and F organoids yielded 14 clusters. b. Cluster 13 was unique to F organoids. c. *FBXW7* is actively expressed in wildtype organoids with a reduction in expression in F organoids. All experiments were performed with N = 3 biological replicates. W: wildtype cells, F: *FBXW7*$^{-/-}$ cells

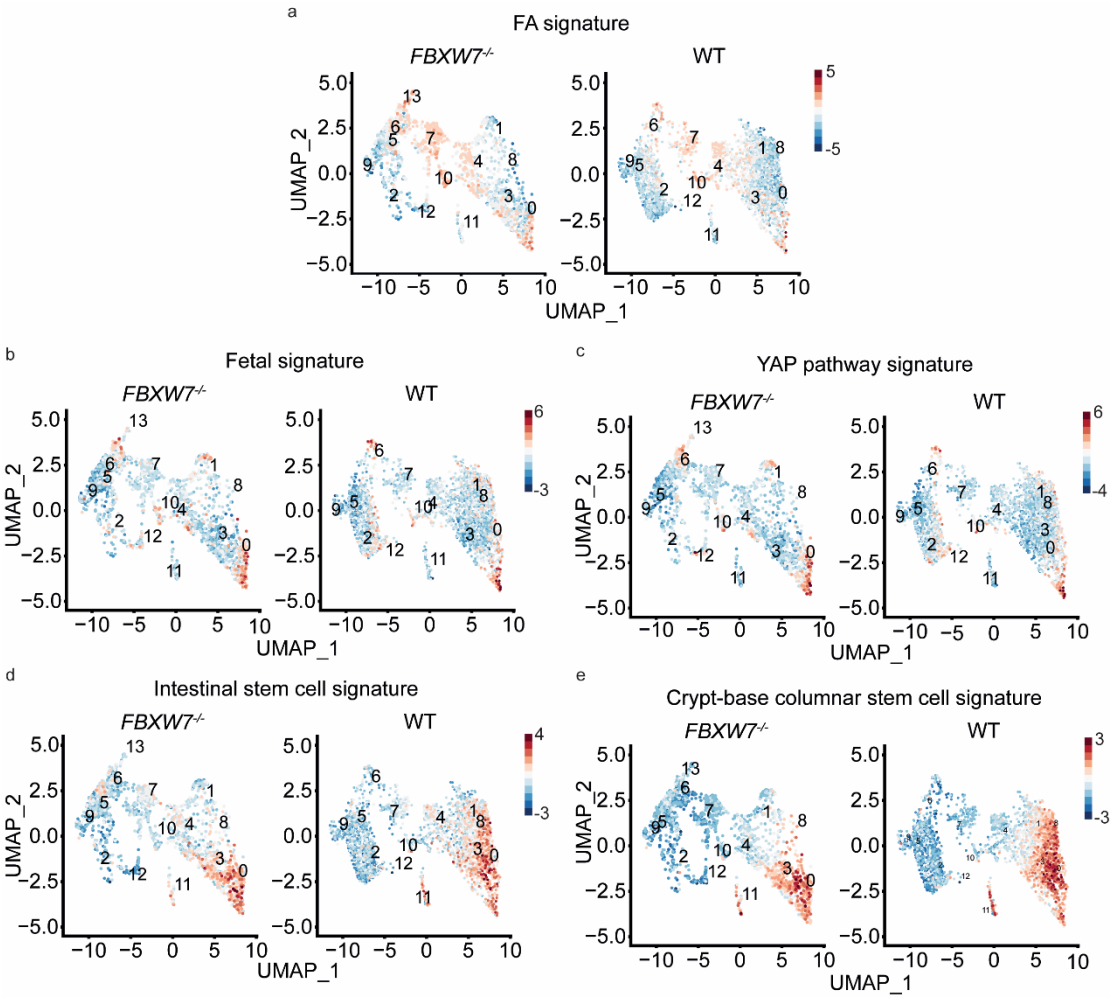


Figure 26: scRNAseq of W and F organoids demonstrated differences on gene signature and pathway analysis.

Chi-square analysis of W and F organoids demonstrated that F organoids were enriched for (a) FA gene signature ($p < 2.22 \times 10^{-16}$), (b) fetal signature ($p = 9.4754 \times 10^{-11}$), (c) YAP pathway signature ($p = 3.2134 \times 10^{-7}$), and depleted for (d) intestinal stem cell signature ($p = 1.4887 \times 10^{-6}$) and (e) crypt-based columnar stem cell signature ($p < 2.22 \times 10^{-16}$). All experiments were performed with N = 3 biological replicates. W: wildtype cells, F: *FBXW7*$^{-/-}$ cells

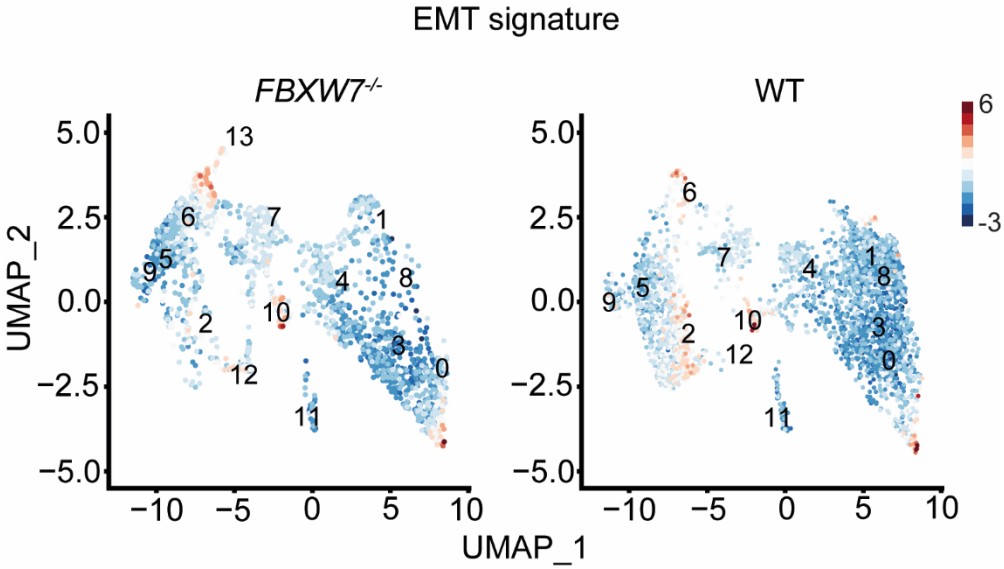
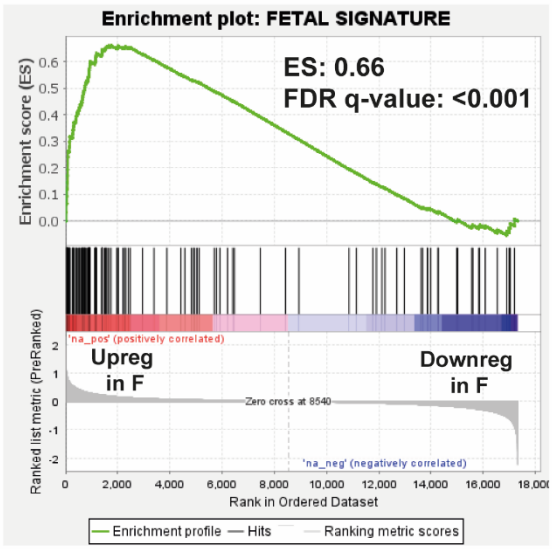


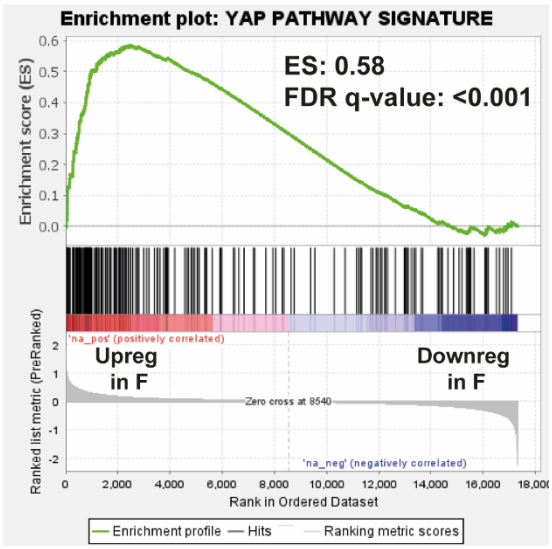
Figure 27: Cluster 13 in F organoids was marked by genes involved in EMT.

Cluster 13 was present only in *FBXW7* $^{-/-}$ organoids. The top five genes expressed in this cluster were *KRT17*, *TAGLN*, *MMP7*, *TUBA1A*, *VIM*, which have been found to be associated with EMT. All experiments were performed with N = 3 biological replicates.

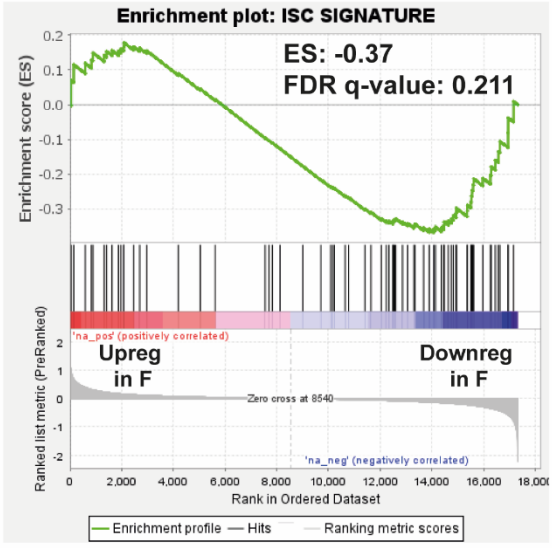
a


Enrichment plot: FETAL SIGNATURE
ES: 0.66
FDR q-value: <0.001
Enrichment score (ES)
'na_pos' (positively correlated)
Upreg in F
Downreg in F
Zero cross at 8540
'na_neg' (negatively correlated)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

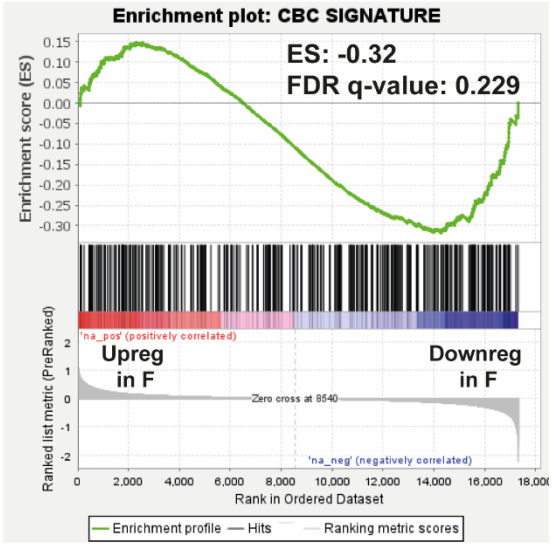
b


Enrichment plot: YAP PATHWAY SIGNATURE
ES: 0.58
FDR q-value: <0.001
Enrichment score (ES)
'na_pos' (positively correlated)
Upreg in F
Downreg in F
Zero cross at 8540
'na_neg' (negatively correlated)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

c


Enrichment plot: ISC SIGNATURE
ES: -0.37
FDR q-value: 0.211
Enrichment score (ES)
'na_pos' (positively correlated)
Upreg in F
Downreg in F
Zero cross at 8540
'na_neg' (negatively correlated)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

d


Enrichment plot: CBC SIGNATURE
ES: -0.32
FDR q-value: 0.229
Enrichment score (ES)
'na_pos' (positively correlated)
Upreg in F
Downreg in F
Zero cross at 8540
'na_neg' (negatively correlated)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

e

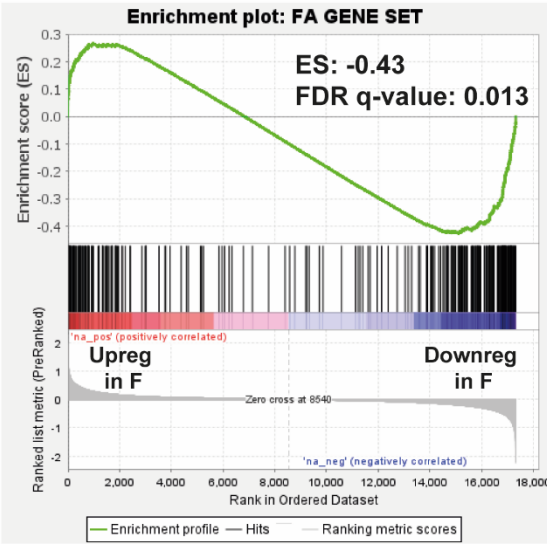

Enrichment plot: FA GENE SET
ES: -0.43
FDR q-value: 0.013
Enrichment score (ES)
'na_pos' (positively correlated)
Upreg in F
Downreg in F
Zero cross at 8540
'na_neg' (negatively correlated)
Ranked list metric (PreRanked)
Rank in Ordered Dataset
Enrichment profile
Hits
Ranking metric scores

Figure 28: GSEA of bulk RNAseq data of F vs W organoids using key pathways derived from scRNAseq analysis.

There was a significant upregulation of the fetal signature (Fernandez Vallone et al., 2016) (a), and YAP pathway (b) in F organoids compared with W organoids. There was a negative association in ISC (c) and CBC (d) signatures though this did not reach statistical significance. I however observed that our FA gene set was negatively correlated in F organoids (e). GSEA: Gene set enrichment analysis, W: wildtype cells, F: *FBXW7*$^{-/-}$ cells, ISC: intestinal stem cell, CBC: crypt-based columnar

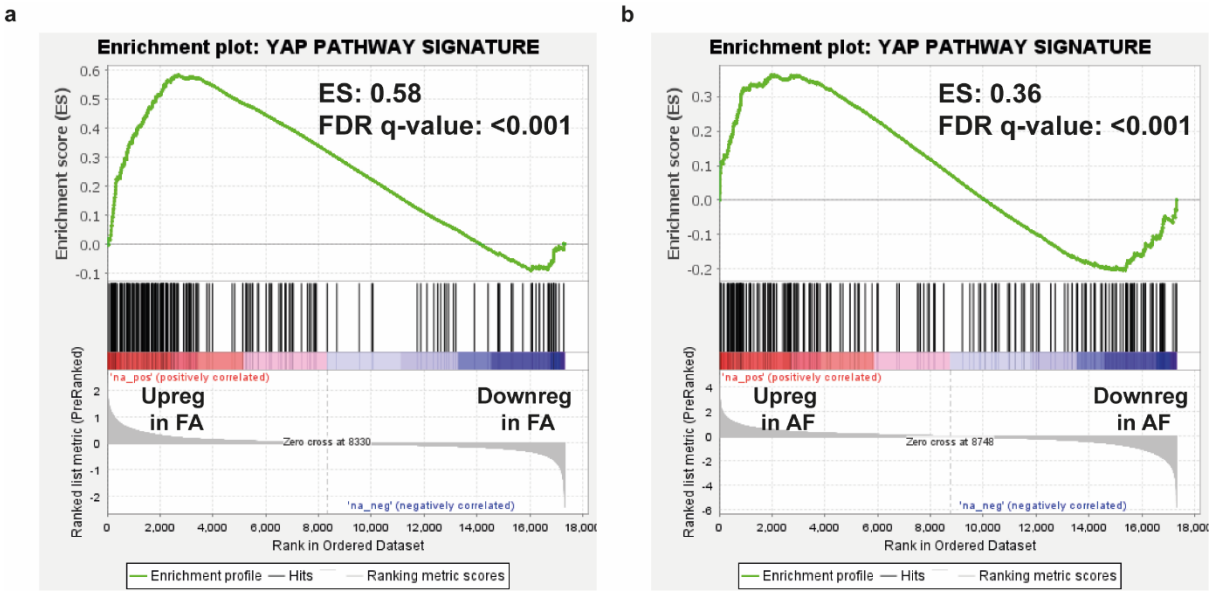


Figure 29: GSEA of bulk RNAseq data of FA vs W and AF vs W for the YAP pathway signature.

There was a stronger association between a. FA vs W organoids, then b. AF vs W organoids for the YAP pathway signature compared to AF vs W organoids.

Transcriptional effects of mutational order are driven by local and global changes in chromatin accessibility

Focus was turned towards uncovering the mechanism which resulted in FA and AF organoids exhibiting such profound transcriptional differences despite identical genotypes. I hypothesized that changes in chromatin accessibility induced by the earlier acquisition of an *FBXW7* mutation could influence the effects of a subsequent *APC* mutation. To interrogate this, assay for transposase-accessible chromatin with sequencing (ATACseq) on W, FA and AF organoids was performed. In total, 14 543 peaks differed between AF and W organoids, 16 990 peaks differed between FA and W organoids, and 2 156 peaks differed between AF and FA organoids. Reassuringly, these findings suggest that there was much similarity in the effect of mutational order on chromosome accessibility between AF and FA. More importantly, the overall accessibility of the whole genome was reduced in FA vs W organoids compared with AF vs W organoids (Figure 32a). An increased accessibility to genes in the CBC signature in AF vs W organoids compared with FA vs W organoids (Figure 32b) was observed in concordance with scRNAseq analysis performed on F and W organoids. Transcription factor motif scans of peaks detected at the transcriptional start sites including 5kb upstream and downstream was performed. *De novo* transcriptional factor motif scan using the GADEM algorithm on ATAC-seq peaks from FA vs W organoid comparison included “nCwGCmCwGn” (c) and "rGCAGGTGn” (d). The top 3 closest matches to each of these motifs as per JASPAR database, were *ZNF449*, *TEAD2*, and *TEAD1* and *SNAI1*, *SCRT2*, and *CTCFL* respectively. Intriguingly, TEAD1 and TEAD2 motifs were more accessible in FA vs W organoids relative to AF vs W organoids (Figure 32e). The TEAD transcriptional factor is the main regulatory factor for the Hippo signaling pathway (Lin, Park and Guan, 2017), which in turn activates YAP and TAZ. This concurred with scRNAseq analysis of F organoids which showed an upregulation of YAP pathway gene signature. I also observed increased chromatin accessibility of FA vs W organoids compared with AF vs W organoids at the SNAI1 transcription factor motif (Figure 32f). SNAI1 mediates EMT by repressing E-cadherin expression, therefore concurring with results observed in bulk RNAseq

and scRNAseq analysis of differentially expressed pathways. Taken together, these findings strongly implicate an epigenetic mechanism in which an early *FBXW7* deletion results in transcriptional quiescence and represses the effect of a later *APC* mutation.

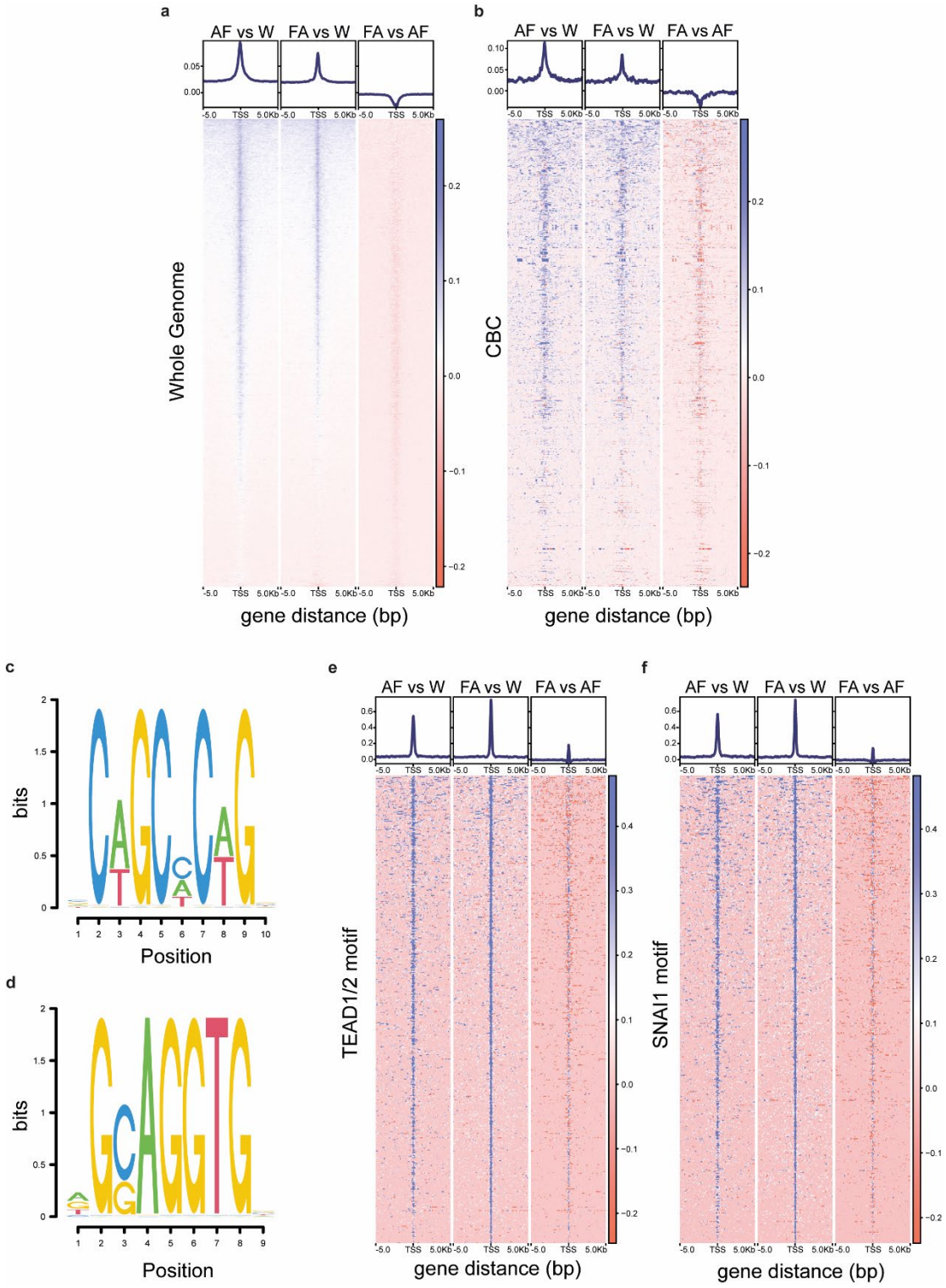

a
AF vs W
FA vs W
FA vs AF
Whole Genome
gene distance (bp)
b
CBC
gene distance (bp)
c
bits
Position
d
bits
Position
e
TEAD1/2 motif
gene distance (bp)
f
SNAI1 motif
gene distance (bp)

Figure 30: ATACseq analysis suggested cell state changes are due to alterations in chromatin accessibility.

ATACseq analysis suggested cell state changes are due to alterations in chromatin accessibility. a. Chromatin accessibility of the whole genome was decreased in FA vs W organoids compared with AF vs W organoids, which could account for the dampened transcriptional impact of FA vs AF organoids. b. Chromatin accessibility of CBC signature genes was increased in AF vs W organoids compared with FA vs W organoids, in concordance with expression of the CBC signature observed on scRNAseq. Motif plots of the two *de novo* motifs found enriched in the FA vs W organoid comparison were (c) “nCwGCmCwGn” and (d) "rGCAGGTGn”. Chromatin accessibility of TEAD1/2 (e) and SNAI1 (f) motifs were upregulated in FA vs W organoids compared with AF vs W organoids. All experiments were performed with N = 3 biological replicates. GSEA: Gene set enrichment analysis, W: wildtype cells, FA: *FBXW7*$^{-/-}$*/APC*$^{-/-}$ cells, AF: *FBXW7*$^{-/-}$*/APC*$^{-/-}$ cells, CBC: crypt-base columnar

Effects of our observed timing of mutation are recapitulated in patient data

To perform patient validation of these findings, two datasets to infer the transcriptional impact of the timing of a mutation were interrogated. From the National Cancer Institute TCGA-COAD database, 48 samples which were *FBXW7*^{mut} and 272 samples which were *FBXW7*^{WT} were identified. Single-sample gene set enrichment analysis (ssGSEA) was performed on each of these samples using the earlier derived FA gene signature (Figure 33a). Intriguingly, I observed a negative association between the FA_up gene signature (NES= -1.88, *p*-value= 2.22 x 10^{-7}), and a positive association between the FA_down gene signature (NES= 2.10, *p*-value= 2.12 x 10^{-8}) in a comparison of *FBXW7*^{mut} vs *FBXW7*^{WT}. As I was not able to determine when the acquisition of an *FBXW7* mutation occurred in cancers from the TCGA-COAD cohort, I hypothesise that the lack of my FA_up gene signature could possibly hint that *FBXW7* mutations in the TCGA-COAD database occurred at a more advanced stage after the acquisition of an *APC* mutation. In addition, ssGSEA of samples showed there was an inverse relationship between our FA_up and FA_down gene signatures (correlation score: -0.55, $p < 2.2 \times 10^{-16}$) (Figure 33b). Reassuringly, this suggests that our FA_up gene signature represents a novel fetal regenerative/YAP active cancer geneset to identify cancer cells in this state.

I further analysed transcriptional data from the recently published EPICC dataset (Heide et al., 2022). Here, the authors had performed single-gland RNAseq on a number of phenotypically normal glands. I analysed transcriptional data from phenotypically normal glands, and performed principal component analysis (PCA) on these glands (Figure 33c). Notably, one gland had an *FBXW7* mutation in the absence of an *APC* mutation, while another gland had an *APC* mutation in the absence of an *FBXW7* mutation. Two glands possessed an *FBXW7* as well as an *APC* mutation, though it was not possible to conclusively determine which mutation was acquired first. The sequenced *FBXW7* mutants were all R578* variants. Intriguingly, the gland with an *APC* mutation alone was located on the PCA diagram apart from all other glands. Notably, the two glands which had both *APC* and *FBXW7*

mutations also clustered closer to the wildtype and *FBXW7* mutant gland, suggesting that these two glands could have acquired the *FBXW7* mutation prior to the *APC* mutation, and therefore could explain why the glands still retained phenotypic normality. I acknowledge that these inferences have been drawn based on a limited number of glands, but are encouraged that existing patient-derived transcriptional data demonstrates concordance with the findings.

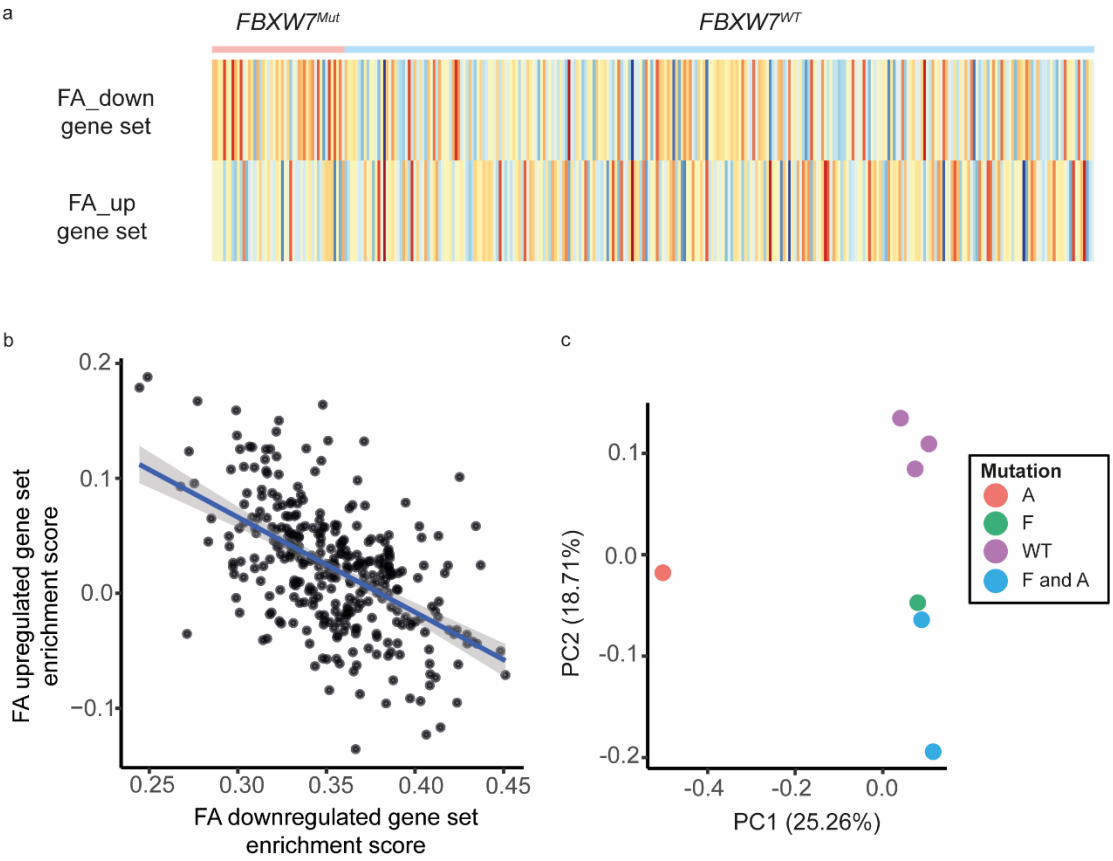


Figure 31: Validation of results in patient data.

a. Using TCGA-COAD data, single-sample gene set enrichment analysis (ssGSEA) of 48 samples which were *FBXW7*^{Mut} and 272 samples which were *FBXW7*^{WT} demonstrated a negative association with the FA_up gene set, and a positive association with the FA_down gene set in a comparison between *FBXW7*^{Mut} and *FBXW7*^{WT}. b. Moreover, there was an inverse relationship between the enrichment score for the FA up- and down-regulated gene sets (derived from FA vs AF organoids) across the 320 samples in the TCGA-COAD (correlation score: -0.55, $p < 2.2 \times 10^{-16}$). Together, these suggest that our FA gene signature represents a novel fetal regenerative/YAP active geneset to identify cells in this state. c. PCA of phenotypically normal intestinal glands from a recently published EPICC dataset (Heide et al., 2022) showed that a gland with both an *APC* and *FBXW7* mutation clustered closer to the *FBXW7*-mutant gland and apart from the *APC* gland, which could account for its phenotypic appearance. ssGSEA: single sample gene set enrichment analysis, FA: *FBXW7*$^{-/-}$/*APC*$^{-/-}$ cells, AF: *FBXW7*$^{-/-}$/*APC*$^{-/-}$ cells, TCGA-COAD: The Cancer Genome Atlas – Colon adenocarcinoma, PCA: principle component analysis

Discussion

This research provides preliminary evidence which can help to advance some core concepts in molecular biology. As mentioned, results from my experiments revealed the impact of *FBXW7* deletions on epistasis and expounds that the mutational order in which the *FBXW7* deletion is acquired has far-reaching consequences on the cell. Epistasis is a concept which suggests that genetic interactions result in unique molecular interactions which extend beyond the impact of each genetic perturbation if studied in isolation (Johnson, Reddy and Desai, 2023). Exploration of epistasis remains limited within the context of cancer. In a study of patients with essential thrombocytopenia, polycythemia vera and myelofibrosis, the order in which *JAK2* or *TET2* was mutated affected the clinical presentation of disease (Ortmann et al., 2015). In particular, an earlier *TET2* mutation prevented the *JAK2* mutation from upregulating genes associated with proliferation. In contrast, *JAK2* mutant first cells presented with cancer at a younger age, had polycythemia vera preferentially, and were more likely to have thrombotic events occurring. My organoid model, in conjunction with the CRISPR gene editing techniques, is an optimal model which to study epistasis and mutational order. As explored in my experiments, I am able to introduce multiple mutations in a specific sequence to identify the impact of altering the sequence on the transcriptional consequence of the cell. This would have been difficult to achieve in mouse models in which there lacks technology to introduce conditional knockouts beyond the Cre- and Dre- dual recombinase systems (Wang et al., 2022).

My research also described the importance of mutational order. Here, I showed first that an *FBXW7* deletion acquired on different mutational backgrounds conferred different transcriptional effects. Furthermore, I demonstrated that organoids with both an *FBXW7* and an *APC* mutation exhibited different transcriptional effects just by reordering the sequence by which these mutations were acquired. Taken together, these results suggest that understanding

the mutational background of a cancer at a single point in time provides only partial understanding of that cancer's evolutionary past and future. This could have implications in clinical practice. Targeted therapy in CRC relies on the presence of specific mutations such as *BRAF* in the case of Encorafenib, or *ERBB2* in Herceptin. Yet, a variegated response to chemotherapy exists in spite of patients possessing the requisite mutation. One possible explanation for this observation could be the mutational background, as has been shown in my study.

At this juncture, it is crucial to discuss the potential impact of my *FBXW7* biallelic truncating model vis a vis what was biologically detected on sequencing of the normal colon (Lee-Six et al., 2019). As mentioned in the introduction to this chapter, biallelic truncating mutations at *FBXW7* which in my model, essentially render the protein absent, are rare. Based on the "just enough" hypothesis, an optimal level of FBXW7 activity occurs with a single allelic missense mutation at one of the hotspots corresponding to the propellor tip region of the FBXW7 protein (Davis and Tomlinson, 2012). In cancer, this suggests that such a mutation most optimally results in cellular proliferation. It would therefore be reasonable to predict that the effect of generating bi-allelic loss of FBXW7 activity is likely to result in an excessive reduction in SCF complex efficacy such as to render the cellular phenotype less than ideal for cellular proliferation. Indeed, a single allelic propellor tip mutant at R482Q exhibited a stronger phenotype of polyposis versus a *FBXW7* $^{+/-}$ model in mice (Davis et al., 2014). Furthermore, this also corresponds to previous work in mice using gut specific knockout of Fbxw7 which resulted in truncating mutations at exon 5, disrupting the F-box, and similarly resulting in complete loss of functional protein (Sancho et al., 2010). Here, the authors noted that this model of Fbxw7 mutation did not result in increased polyposis in the intestines of mice, and required concurrent loss of Apc for polyps to be formed, findings which were recapitulated in my experiments. Overall, while my model of *FBXW7* mutation differed from what is found *in vivo*, my focus on the effects of the *FBXW7* mutation on epistasis and mutational order remains a novel finding. However, it is reasonable to suggest that changing

the *FBXW7* mutation could affect the magnitude of the transcriptional consequence described in my work. Therefore, similar to the abovementioned studies which have studied the impact of various *FBXW7* mutant genotypes on cellular proliferation, experiments should be conceived to study the effect of these various *FBXW7* mutant genotypes on transcriptional consequences.Another factor relating to the *FBXW7* gene is the present of three isoforms directed by alternative splicing at the 5`end of the gene. My targeting strategy was direct at the α-isoform because this is the most abundant *FBXW7* isoform. In mice, the β- and γ-isoforms were not detected in the intestine (Matsumoto, Onoyama and Nakayama, 2006). Notwithstanding this, I consider that there could be the possibility of β- and γ- isoforms of *FBXW7* which could still be present in my mutant cells. This raises the possibility that the transcriptional and experimental effects seen in my experiments relate solely to targeting of the α-isoform.

FBXW7 mutation notwithstanding, these findings were made possible only because of recent developments in organoid and gene editing technologies. Cell lines have been important models in cancer research, but are laden with hundreds or thousands of pre-existing mutations. Organoids derived from phenotypically normal tissue are however generally wildtype for driver mutations. Gene editing allows for mutational order to be directly observed *in vitro*. In this study, *FBXW7* and *APC* were used as exemplar mutations to study epistasis. Given the profound ITH observed in cancers, there is scope to extend the techniques and models used in my study to understand how the sequential and variable acquisition of driver mutations affects cancer phenotypes. For example, one question could be whether a *TP53* mutation acquired before *KRAS* results in a different phenotype from *KRAS* acquired before *TP53*.

My study also sought to understand the role of *FBXW7* deletion in phenotypically normal colonic tissue. Here, I demonstrated that *FBXW7* deletion in normal tissue resulted in a subtle reorientation of the cellular state towards a fetal state and promoted genes associated with the YAP/TAZ signaling pathway. While ATACseq on F and A organoids were not performed per

se, analysis in FA and AF organoids demonstrated that FA organoids had increased chromatin accessibility at TEAD1 and TEAD2 transcription factor binding sites, suggesting that YAP/TAZ signaling was upregulated via such an epigenetic mechanism. More importantly, my results demonstrated that one contributing factor for the decreased transcriptional perturbation observed in FA organoids relative to AF organoids could be explained by the generally more closed chromatin accessibility state in FA organoids. Overall, these findings were in concordance with two very recently published papers. In the first paper, an association between TEAD-family transcriptional factor accessibility was observed in cells which had acquired a more fetal-like state (Pikkupeura et al., 2023). In the second paper, an unbiased CRISPR-screen revealed *FBXW7* as a top target regulator of the fetal intestinal state (Hansen et al., 2023).This fetal intestinal cell state remains hitherto poorly defined in the literature and appears limited to a state of unlimited cellular regeneration induced during cellular injury or irradiation (Fernandez Vallone et al., 2016, Mustata et al., 2013), and in cancer, has been purported to result in a cancer stem cell state owing to its regenerative and proliferative characteristics. Findings from my research therefore stand at some odds with existing literature which purports that the fetal state is actually pro-oncogenic. I hypothesise a few possibilities for this discrepancy. Firstly, the fetal state remains poorly defined molecularly, with gene sets describing the fetal state exhibiting minimal overlap with one another. This suggests that the field has not reached a point of maturity at which the fetal state can be accurately defined. In my study, I have utilized one approach which defined a fetal state (Fernandez Vallone et al., 2016) based on 692 genes which were upregulated at least 4-fold in gastric and intestinal organoids compared with adult intestinal organoids. One way to further strengthen my argument that an *FBXW7* mutation results in a fetal state would be to do the converse of what had been done in my experiments. This is to say that fetal organoids may be harvested, and my *FBXW7* gene set can instead be analysed based upon the transcriptional signature obtained from the fetal organoids. This would constitute further work. Another potential reason for the discrepancy between my proposed protective effect of the fetal state on carcinogenesis is that hitherto evaluations of the fetal state have been

undertaken on a background of intestinal insult (Fernandez Vallone et al., 2016, Vasquez et al., 2022, Mustata et al., 2013) whereas in my experiments, the organoids remains phenotypically normal except for an *FBXW7* mutation. Environmental insult in the form of tissue injury or radiation could provide an additional insult which primes the phenotype towards oncogenesis. Again, this is speculative, and would constitute subsequent work.

Nonetheless, these findings from my experiments point towards a protective role of *FBXW7* mutations in normal epithelial tissue. The acquisition of an early *FBXW7* deletion results in a generally more closed chromatin accessibility state which decreases overall transcriptional activity. At regions where there is increased chromatin accessibility such as at TEAD1 and TEAD2 binding sites, some pathways such as YAP/TAZ signaling are upregulated. More generally, there is also a reorientation towards a more fetal-like intestinal state. The early *FBXW7* deletion abrogates the transcriptional perturbation which accompanies a subsequent *APC* mutation, and could therefore potentially result in a lower likelihood of future cancer occurrence.

At this point, a number of limitations relating to the experimental design need to be acknowledged. Firstly, truncating deletions in *FBXW7* were generated in my organoid model instead of the missense mutation described during whole genome sequencing of normal colonic epithelial tissue (Lee-Six et al., 2019). Missense mutations are also the most common *FBXW7* mutation type in CRC found in CRC (Davis and Tomlinson, 2012). Missense mutations however require DNA repair by homologous recombination and therefore suffer from a lower efficacy in human colonic organoids. I also did not formally sequence specimens for the prevalence of CRC driver mutations, although I anticipate this to be somewhat reduced given the rarity of such mutations in normal epithelia. Afforded more time, I could also consider using bulk RNAseq data to determine the presence of such driver mutations. Finally, multiple colonic specimens were used throughout this research to generate organoids which could possibly introduce patient to patient heterogeneity. This was unfortunately a function of human colonic organoids undergoing natural senescence after

about 12 subcultures. I was however comforted that PCA clustered organoids with the same genotype together, instead of clustering based on patient origin. This suggests that the effects of introducing a new genotype superseded the heterogeneity arising from patient to patient variability.

With more time, I would have liked to have built on these findings with further experiments. Firstly, I would like to directly compare phenotypes associated with malignancy in AF and FA organoids. Presently, I have used transcriptional consequences from both bulk RNAseq and scRNAseq as a readout for changes in cellular state. It would be meaningful to directly measure differences in proliferation between AF and FA organoids, and in turn comparing with F and A organoids, to establish the consequences of mutational order and epistasis on actual phenotypes. I would also consider further validation of this effect of an increase in YAP/TAZ signaling by overexpressing YAP/TAZ in the wildtype organoids and observing for transcriptional similarity to the results which I have observed in my present experiment. Likewise, YAP/TAZ signaling could be blocked in either A or AF organoids. Together, these experiments would provide further strength to the observation that YAP/TAZ signaling is associated with a fetal state, and that this upregulated fetal state is critical in abrogating the effects of a later *APC* mutation. Moreover, I would also be keen to elucidate the mechanism by which an *FBXW7* deletion results in the changes in chromatin accessibility observed. This would reveal a hitherto unknown function of the *FBXW7* gene. This could be performed by a proteomic or phospho-proteomic screen using mass spectrometry and the validation of potential targets and protein-protein interactions using co-immunoprecipitation. Finally, the implications of this work could extend to other genes as well as other cancer types. Thus, further experiments conducted over a wider mutational spectrum and in different cancers would establish a future translational application of mutational order and epistasis in clinical practice.

Chapter 6: Conclusion

I have divided my work in this thesis into two Themes, both using *FBXW7* as the candidate gene for my experiments. Both of these themes the spectrum of oncogenesis from normal tissue to cancer. Theme 1 evaluates the impact of an *FBXW7* mutant cell on its neighbouring cells, while Theme 2 evaluates the characteristic and function of *FBXW7* deletion in normal colonic tissue.

In Theme 1, I introduced an *FBXW7* mutation in two different cell lines. I observed that *FBXW7* mutant cells proliferated slower than *FBXW7* wildtype cells. In addition, I observed that mutant cells upregulated oxidative phosphorylation, and also had an increased burden of DNA damage. Intriguingly, I observed that wildtype cells co-cultured with mutant cells using a Transwell model exhibited increased DNA damage too. Proteomic analysis revealed that AKAP8 was secreted into the media of *FBXW7* mutant cells, and was absent in wildtype cells. Wildtype cells overexpressing AKAP8 exhibited upregulation of DNA damage, which was abrogated when wildtype cells were co-cultured with *FBXW7* cells in which *AKAP8* had also been knocked out. Taken together, I concluded that AKAP8 secreted by *FBXW7* mutant cells were taken up by wildtype cells, resulting in increased DNA damage in wildtype cells.

In Theme 2, I generated organoid models bearing different mutational backgrounds to represent the different stages of carcinogenesis – normal, precancerous, and cancer. These different stages were represented by organoids generated from phenotypically normal tissue, denoted as W organoids. CRISPR-Cas9 gene edited W organoids targeting the *APC* gene generated A organoids, representing the precancerous stage. Finally, CRISPR-Cas9 gene edited A organoids targeting the *TP53* gene generated AT organoids. An *FBXW7* deletion was introduced on W, A and AT organoids, generating F, AF and ATF organoids respectively. Intriguingly, although the same *FBXW7* deletion was introduced, the transcriptional consequence of this mutation varied depending on the mutational background of the organoid. In addition, organoids which possessed the same mutations but differed only

in the order in which the organoids were generated similarly possessed different transcriptional consequences. I sought to determine the mechanistic effect of *FBXW7* deletion on wildtype organoids using scRNAseq and ATACseq. Together, my experiments suggested that *FBXW7* deletions resulted in a reordering of the cell state towards a fetal state via YAP/TAZ signaling. I observed a global reduction of chromatin accessibility in *FBXW7* deleted organoids. Although the chromatin was rendered less accessible in general, unbiased detection of motif sequences which were more accessible identified TEAD1/TEAD2, a key mediator of YAP/TAZ signaling, suggesting that TEAD accessibility in *FBXW7* mutants could be responsible for the observed upregulation in YAP/TAZ and fetal gene signatures in *FBXW7* deleted organoids.

In this concluding chapter, I analyze my findings from the perspective of its contribution towards the understanding of ITH, the timing of a mutation, and implications on cancer screening.

FBXW7 mutations in different settings as an exemplar mutation to glean insights into intratumoral heterogeneity

ITH is a broad term which can refer to genetic, transcriptional, or epigenetic heterogeneity. Multiregional sequencing of cancers has firmly established the presence of genetic ITH in CRC (Sottoriva et al., 2015), while single-cell sequencing analyses has allowed for the tracking of transcriptional and epigenetic changes along the continuum from phenotypic normalcy, to precancerous adenomas, and finally in cancer (Becker et al., 2022). Yet, significant questions which could have profound implications on tumour initiation and growth remain. Some of these questions include defining specific triggers which give rise to ITH, and what fraction of ITH has a bearing on the phenotype of the cancer, rather than simply being a byproduct of mutagenesis. In my thesis, I have conducted experiments addressing two different themes. While each individual chapter has discussed the findings of those

experiments within each theme, here, I consider broader implications of my research on ITH in CRC.

Insights into the origin and functions of ITH

Endogenous and exogenous causes of DNA damage can give rise to genetic ITH. Exogenous sources include ultraviolet irradiation or carcinogen exposure such as cigarette smoke. Endogenous sources could be related to specific mutations which result in an accumulation of other mutant genes, such as in *POLE* or *POLD* mutants (Ma et al., 2022), or cells which have defective mismatch repair genes. As reflected in the assignment of aetiologies for unique mutational signatures, these molecular processes incompletely account for the entire spectrum of mutational insults, suggesting that other processes remain undiscovered. At present, causes of endogenous DNA damage have been attributed to cell intrinsic effects arising as a result of genetic mutations. In Theme 1, I demonstrated that *FBXW7* mutations increase the burden of DNA damage in a cell intrinsic way. More importantly, I demonstrated that cells in the vicinity of *FBXW7* mutant cells which are wildtype for *FBXW7* also possess a modest but significant increase in DNA damage burden. This highlights a mechanism of DNA damage acquisition which is not cell intrinsic, and suggests that DNA damage burden could far exceed the quantum expected if accounting for the mutant cell alone. This also raises the possibility that other sources of DNA damage exists which are endogenous but cell autonomous, and hypothetically could become a source of genetic ITH in the evolution towards cancer.

Genetic ITH is not the only source of heterogeneity, which encompasses transcriptional as well as epigenetic heterogeneity. In Theme 2, I demonstrate that a single *APC* mutation can be a source of significant transcriptional perturbation, which has recently also been characterised elsewhere (Bala et al., 2023). My results also suggest that mutations in *FBXW7* reduce global chromatin accessibility, dampening the transcriptional perturbation which would otherwise have accompanied an *APC* mutation. *FBXW7* instead increases the

accessibility of some regions which reorient the cell towards a fetal state and upregulates YAP/TAZ signalling. Crucially, on bulk RNAseq analysis of the transcriptional consequences between AF and FA organoids, focusing on genes which were commonly differentially expressed between both groups, I observed that there was a reduction in the log2FC values between AF vs W, and FA vs W. In other words, when comparing genes which were identified as being differentially expressed in *both* AF vs W and FA vs W, there was a tendency for the log2FC value to tend towards zero, suggesting a reduction in the magnitude in which the gene becomes differentially expressed. One possible explanation which my data strongly hints at, is that an earlier *FBXW7* mutation abrogates the effect of a later *APC* mutation. Given that *APC* is often regarded as the founder mutation which results in polyposis in CRC, it appears that the transcriptional ITH observed in *APC* mutant organoids is markedly reduced when an earlier *FBXW7* mutation occurs first. More broadly, these findings raise the possibility that specific mutations have the potential to inhibit or express transcriptional and epigenetic ITH, in turn leading to changes in cellular state. This therefore hints at the origins of transcriptional and epigenetic ITH being associated with certain gatekeepers which could either hinder or facilitate ITH.

Together, observations from my experiments propose a diverse role which mutations may play in modulating the induction of ITH. Some mutations may induce cell adjacent DNA damage, which could be a catalyst for DNA damage in those wildtype cells. Other mutations could have gatekeeping functions by altering chromatin accessibility. Further research will be needed to uncover whether these observations may be more broadly applied to other genes and in other cancers.

Recently, a meta-analysis of 1163 tumour samples from 24 tumour types has revealed 41 consensus meta-programs of ITH, which describe the range of intratumoral transcriptional programs observed to be perturbed in different cancers (Gavish et al., 2023). These 41 meta-programs can be broadly classified into eleven transcriptional hallmarks. The two most common transcriptional hallmarks were found to be relating to cell cycle and cellular stress.

Notably, these two hallmarks were observed to be common in nearly all 24 tumour types. Other transcriptional hallmarks were context specific and could be found only in a subset of tumour types. For example, interferon response was present in many solid cancers, but was almost absent in sarcomas. The ability to describe patterns of transcriptional ITH across different cancer types suggests that the perturbation of transcription is non-random, and could actually account for phenotypic characteristics of the cancer. Taking this idea further, one could hypothesise that both genetic and transcriptional ITH therefore contribute to cancer phenotype. In CRC, while significant work has been done to characterise the effect of cells harbouring mutations in the driver mutations on adjacent cells, less has been done on other mutations. This leaves out a not insignificant proportion of the ITH relating to mutations in non-driver mutations.

In theme 1, I investigated the effect of an *FBXW7* mutation on adjacent cells, and described an interaction between them. *FBXW7* is mutated in up to 17% of CRC, and is therefore a relatively common mutation found in CRC. The TCGA-COAD PanCancer atlas suggests that other commonly mutated genes include *MUC16* (incidence in CRC 27.3%), and *RYR2* (19.7%) (cBioPortal, 2023). *MUC16* mutations in CRC are associated with worse prognosis and survival, and has been shown to upregulate JAK2/STAT3 signaling in CRC, resulting in increased cellular proliferation (Liu et al., 2022). *RYR2* mutations in esophageal adenocarcinoma was found to upregulate signaling pathways associated with enhanced anti-tumour immunity, and had a better prognosis than *RYR2* wildtype counterparts (Liu et al., 2021). While these are not by any means direct evidence of a competitive or collaborative effect of mutant cells on adjacent cells, future work could be targetted towards investigating whether such commonly occurring mutations in CRC possess such a cell adjacent effect.

The function of ITH is perhaps also hinted at in Theme 2. Here, I demonstrated an epistatic effect such that the same *FBXW7* mutation results in highly varied transcriptional effects depending on the mutational background. A mutational background of increased genetic ITH therefore could result in subsequent mutations having a greater transcriptional impact.

Genetic ITH therefore acts as a catalyst for even greater transcriptional perturbation. By extension, one could hypothesise that the degree of ITH could be measured as a clinical marker of the extent in which that cell is on an evolutionary trajectory towards cancer, and be used to alert physicians to adopt more aggressive cancer screening strategies. More research to evaluate these hypotheses will need to be conducted.

This research therefore may be used to suggest a number of functions of ITH. On one hand, consideration of the individual mutations occurring within heterogeneity must be considered, as driver mutations could be imparting unique characteristics on adjacent cells. Furthermore, the degree of ITH could impact on epigenetic traits such as chromatin accessibility, and have implications on the consequences of subsequent mutations.

Epistasis and the timing of a mutation suggests one-off sequencing incompletely characterises the cell

Moving on from ITH, one of the principal findings from Theme 2 is that the timing of a mutation is critical in understanding the molecular characteristics of a cell. In Theme 2, whether an *FBXW7* mutation occurred before or after an *APC* mutation had profound implications on the transcriptional consequence of that cell. When the *FBXW7* mutation occurred before an *APC* mutation, less genes were differentially expressed compared to when an *APC* mutation occurred before an *FBXW7* mutation. This finding suggests that one-off sequencing incompletely characterises the expected phenotype of a cell. Furthermore, my experiments demonstrated that the mutational background is equally important when understanding the impact of a single mutant gene. Epistasis suggests that simply determining whether a gene is mutant provides an incomplete understanding about the effects of that gene being mutant. It is as important to understand that gene's potential interactions with other genes.

As an example, *BRAF* mutant CRC represents a subtype of CRC with one of the poorest prognosis as well as poorest response to chemotherapy . In studies which evaluate the response of patients with *BRAF* mutations to different treatment modalities, a highly variable response to treatment has been observed. Only 26% of CRC patients with *BRAF* mutations undergoing triple-therapy of encorafenib, binimetenib and cetuximab registered a response to this treatment, suggesting that possessing a mutation alone does not guarantee efficacy to treatment (Kopetz et al., 2019). Results from my experiments suggest that the timing in which the *BRAF* mutation was acquired could play a critical role in explaining why some cancers respond to therapy while others do not. This is to say that a snapshot of the mutations present in a patient's cancer does not provide sufficient information to determine that patient's response to chemotherapy because the timing in which the critical mutation occurs could impact on treatment response. Further experiments which predict the timing in which mutations were acquired may be essential to gauge treatment response.

Implications on screening for cancer

One of the aspects which most excites me returns to the original question which I had sought to answer in formulating experiments in Theme 2 – the role of driver mutations in normal colonic tissue. In the esophagus, *NOTCH1* mutations are highly prevalent and may be identified in up to 80% of phenotypically normal esophageal cells (Ng and Chan, 2023). A recent study provided evidence that *NOTCH1* mutations in normal esophageal epithelium conferred a beneficial effect due to accelerated clonal expansion, allowing *NOTCH1* mutant cells to colonize the epithelium while maintaining normal cellular function and behaviour (Abby et al., 2023). The increased proliferation of mutant *NOTCH1* cells therefore acted as a guardian against other mutations, which need to acquire mutations which can outcompete the proliferation rate of *NOTCH1* mutant cells.

Likewise, in the colon, my experiments have suggested a different mechanism in which *FBXW7* mutant cells could be protective against cancer, albeit via a different mechanism. I believe that this could have profound implications on cancer screening. The identification of driver mutations in normal colonic tissue could potentially hint at an individual's risk profile against future CRC, although such an implication is decades from realisation. More specifically, my results hint at chromatin accessibility being implicated in cancer initiation, raising the possibility that chromatin accessibility could be used as a screening strategy for cancer. Already, there is evidence from the chromatin accessibility profile of 410 tumour samples spanning 23 cancer types that chromatin accessibility profiles may be used to determine features of cancer such as cancer immune evasion and DNA damage and repair regulation (Corces et al., 2018). In pancreatic cancer, chromatin accessibility profiles have also been used to study the sensitivity of individual cancers to various chemotherapy drug sensitivities (Shi et al., 2022). It is possible that future CRC screening would involve not only a fecal immunoglobulin test, but also a test of the accessibility of chromatin structure.

Conclusion

Together, I have studied the impact of an *FBXW7* mutation under different conditions. In the first theme, I studied *FBXW7* mutations as a component of genetic ITH, particularly given a relatively high incidence of mutant *FBXW7* in CRC. In the second theme, I studied the impact of an *FBXW7* mutation in normal tissue, and discuss the importance of the timing of a mutation as well as mutational background on determining the effect of a gene's mutation on the cell. More fundamentally, results from Theme 2 have advanced the understanding of the role of driver mutations in normal colonic tissue.

References

TCGA-COAD [Online]. National Cancer Institute GDC Data Portal. Available: https://portal.gdc.cancer.gov/projects/TCGA-COAD [Accessed 7 March 2023].

Aaltonen, L. A., Salovaara, R., Kristo, P., Canzian, F., Hemminki, A., Peltomäki, P., Chadwick, R. B., Kääriäinen, H., Eskelinen, M., Järvinen, H., Mecklin, J. P. & De La Chapelle, A. 1998. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *N Engl J Med,* **338,** 1481-7.

Abby, E., Dentro, S. C., Hall, M. W. J., Fowler, J. C., Ong, S. H., Sood, R., Herms, A., Piedrafita, G., Abnizova, I., Siebel, C. W., Gerstung, M., Hall, B. A. & Jones, P. H. 2023. Notch1 mutations drive clonal expansion in normal esophageal epithelium but impair tumor growth. *Nat Genet,* **55,** 232-245.

Adane, B., Alexe, G., Seong, B. K. A., Lu, D., Hwang, E. E., Hnisz, D., Lareau, C. A., Ross, L., Lin, S., Dela Cruz, F. S., Richardson, M., Weintraub, A. S., Wang, S., Iniguez, A. B., Dharia, N. V., Conway, A. S., Robichaud, A. L., Tanenbaum, B., Krill-Burger, J. M., Vazquez, F., Schenone, M., Berman, J. N., Kung, A. L., Carr, S. A., Aryee, M. J., Young, R. A., Crompton, B. D. & Stegmaier, K. 2021. STAG2 loss rewires oncogenic and developmental programs to promote metastasis in Ewing sarcoma. *Cancer Cell,* **39,** 827-844.e10.

Alexandrov, L. B., Nik-Zainal, S., Wedge, D. C., Aparicio, S. A., Behjati, S., Biankin, A. V., Bignell, G. R., Bolli, N., Borg, A., Børresen-Dale, A. L., Boyault, S., Burkhardt, B., Butler, A. P., Caldas, C., Davies, H. R., Desmedt, C., Eils, R., Eyfjörd, J. E., Foekens, J. A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinski, M., Jäger, N., Jones, D. T., Jones, D., Knappskog, S., Kool, M., Lakhani, S. R., López-Otín, C., Martin, S., Munshi, N. C., Nakamura, H., Northcott, P. A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J. V., Puente, X. S., Raine, K., Ramakrishna, M., Richardson, A. L., Richter, J., Rosenstiel, P., Schlesner, M.,

Schumacher, T. N., Span, P. N., Teague, J. W., Totoki, Y., Tutt, A. N., Valdés-Mas, R., Van Buuren, M. M., Van 'T Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L. R., Zucman-Rossi, J., Futreal, P. A., Mcdermott, U., Lichter, P., Meyerson, M., Grimmond, S. M., Siebert, R., Campo, E., Shibata, T., Pfister, S. M., Campbell, P. J. & Stratton, M. R. 2013. Signatures of mutational processes in human cancer. *Nature,* **500,** 415-21.

Aquilina, G., Hess, P., Branch, P., Macgeoch, C., Casciano, I., Karran, P. & Bignami, M. 1994. A mismatch recognition defect in colon carcinoma confers DNA microsatellite instability and a mutator phenotype. *Proc Natl Acad Sci U S A,* **91,** 8905-9.

Arabi, A., Ullah, K., Branca, R. M., Johansson, J., Bandarra, D., Haneklaus, M., Fu, J., Ariës, I., Nilsson, P., Den Boer, M. L., Pokrovskaja, K., Grandér, D., Xiao, G., Rocha, S., Lehtiö, J. & Sangfelt, O. 2012. Proteomic screen reveals Fbw7 as a modulator of the NF-κB pathway. *Nat Commun,* **3,** 976.

Babaei-Jadidi, R., Li, N., Saadeddin, A., Spencer-Dene, B., Jandke, A., Muhammad, B., Ibrahim, E. E., Muraleedharan, R., Abuzinadah, M., Davis, H., Lewis, A., Watson, S., Behrens, A., Tomlinson, I. & Nateri, A. S. 2011. FBXW7 influences murine intestinal homeostasis and cancer, targeting Notch, Jun, and DEK for degradation. *J Exp Med,* **208,** 295-312.

Bailey, M. H., Tokheim, C., Porta-Pardo, E., Sengupta, S., Bertrand, D., Weerasinghe, A., Colaprico, A., Wendl, M. C., Kim, J., Reardon, B., Kwok-Shing Ng, P., Jeong, K. J., Cao, S., Wang, Z., Gao, J., Gao, Q., Wang, F., Liu, E. M., Mularoni, L., Rubio-Perez, C., Nagarajan, N., Cortés-Ciriano, I., Zhou, D. C., Liang, W. W., Hess, J. M., Yellapantula, V. D., Tamborero, D., Gonzalez-Perez, A., Suphavilai, C., Ko, J. Y., Khurana, E., Park, P. J., Van Allen, E. M., Liang, H., Lawrence, M. S., Godzik, A., Lopez-Bigas, N., Stuart, J., Wheeler, D., Getz, G., Chen, K., Lazar, A. J., Mills, G. B., Karchin, R. & Ding, L. 2018. Comprehensive Characterization of Cancer Driver Genes and Mutations. *Cell,* **174,** 1034-1035.

Bala, P., Rennhack, J. P., Aitymbayev, D., Morris, C., Moyer, S. M., Duronio, G. N., Doan, P., Li, Z., Liang, X., Hornick, J. L., Yurgelun, M. B., Hahn, W. C. & Sethi, N. S. 2023. Aberrant cell state plasticity mediated by developmental reprogramming precedes colorectal cancer initiation. *Sci Adv,* **9,** eadf0927.

Bando, T., Kato, Y., Ihara, Y., Yamagishi, F., Tsukada, K. & Isobe, M. 1999. Loss of heterozygosity of 14q32 in colorectal carcinoma. *Cancer Genet Cytogenet,* **111,** 161-5.

Barber, L. J., Davies, M. N. & Gerlinger, M. 2015. Dissecting cancer evolution at the macro-heterogeneity and micro-heterogeneity scale. *Curr Opin Genet Dev,* **30,** 1-6.

Bardi, G., Sukhikh, T., Pandis, N., Fenger, C., Kronborg, O. & Heim, S. 1995. Karyotypic characterization of colorectal adenocarcinomas. *Genes Chromosomes Cancer,* **12,** 97-109.

Barker, N., Van Es, J. H., Kuipers, J., Kujala, P., Van Den Born, M., Cozijnsen, M., Haegebarth, A., Korving, J., Begthel, H., Peters, P. J. & Clevers, H. 2007. Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature,* **449,** 1003-7.

Becker, W. R., Nevins, S. A., Chen, D. C., Chiu, R., Horning, A. M., Guha, T. K., Laquindanum, R., Mills, M., Chaib, H., Ladabaum, U., Longacre, T., Shen, J., Esplin, E. D., Kundaje, A., Ford, J. M., Curtis, C., Snyder, M. P. & Greenleaf, W. J. 2022. Single-cell analyses define a continuum of cell state and composition changes in the malignant transformation of polyps to colorectal cancer. *Nat Genet,* **54,** 985-995.

Bird, A. P. 1986. CpG-rich islands and the function of DNA methylation. *Nature,* **321,** 209-13.

Blum-Barnett, E., Madrid, S., Burnett-Hartman, A., Mueller, S. R., Mcmullen, C. K., Dwyer, A. & Feigelson, H. S. 2019. Financial burden and quality of life among early-onset colorectal cancer survivors: A qualitative analysis. *Health Expect,* **22,** 1050-1057.

Boland, C. R. & Goel, A. 2010. Microsatellite instability in colorectal cancer. *Gastroenterology,* **138,** 2073-2087 e3.

Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bigas, M. A., Fodde, R., Ranzani, G. N. & Srivastava, S. 1998. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res,* **58,** 5248-57.

Bond, C. E., Liu, C., Kawamata, F., Mckeone, D. M., Fernando, W., Jamieson, S., Pearson, S. A., Kane, A., Woods, S. L., Lannagan, T. R. M., Somashekar, R., Lee, Y., Dumenil, T., Hartel, G., Spring, K. J., Borowsky, J., Fennell, L., Bettington, M., Lee, J., Worthley, D. L., Leggett, B. A. & Whitehall, V. L. J. 2018. Oncogenic BRAF mutation induces DNA methylation changes in a murine model for human serrated colorectal neoplasia. *Epigenetics,* **13,** 40-48.

Bond, C. E., Mckeone, D. M., Kalimutho, M., Bettington, M. L., Pearson, S. A., Dumenil, T. D., Wockner, L. F., Burge, M., Leggett, B. A. & Whitehall, V. L. 2016. RNF43 and ZNRF3 are commonly altered in serrated pathway colorectal tumorigenesis. *Oncotarget,* **7,** 70589-70600.

Boopathy, G. T. K. & Hong, W. 2019. Role of Hippo Pathway-YAP/TAZ Signaling in Angiogenesis. *Front Cell Dev Biol,* **7,** 49.

Borowsky, J., Dumenil, T., Bettington, M., Pearson, S. A., Bond, C., Fennell, L., Liu, C., Mckeone, D., Rosty, C., Brown, I., Walker, N., Leggett, B. & Whitehall, V. 2018. The role of APC in WNT pathway activation in serrated neoplasia. *Mod Pathol,* **31,** 495-504.

Bose, D., Zimmerman, L. J., Pierobon, M., Petricoin, E., Tozzi, F., Parikh, A., Fan, F., Dallas, N., Xia, L., Gaur, P., Samuel, S., Liebler, D. C. & Ellis, L. M. 2011. Chemoresistant colorectal cancer cells and cancer stem cells mediate growth and survival of bystander cells. *Br J Cancer,* **105,** 1759-67.

Boyle, T., Keegel, T., Bull, F., Heyworth, J. & Fritschi, L. 2012. Physical activity and risks of proximal and distal colon cancers: a systematic review and meta-analysis. *J Natl Cancer Inst,* **104,** 1548-61.

Burdziak, C., Alonso-Curbelo, D., Walle, T., Reyes, J., Barriga, F. M., Haviv, D., Xie, Y., Zhao, Z., Zhao, C. J., Chen, H. A., Chaudhary, O., Masilionis, I., Choo, Z. N., Gao, V., Luan, W., Wuest, A., Ho, Y. J., Wei, Y., Quail, D. F., Koche, R., Mazutis, L., Chaligné, R., Nawy, T., Lowe, S. W. & Pe'er, D. 2023. Epigenetic plasticity cooperates with cell-cell interactions to direct pancreatic tumorigenesis. *Science,* **380,** eadd5327.

Cahill, D. P., Lengauer, C., Yu, J., Riggins, G. J., Willson, J. K., Markowitz, S. D., Kinzler, K. W. & Vogelstein, B. 1998. Mutations of mitotic checkpoint genes in human cancers. *Nature,* **392,** 300-3.

Calbo, J., Van Montfort, E., Proost, N., Van Drunen, E., Beverloo, H. B., Meuwissen, R. & Berns, A. 2011. A functional role for tumor cell heterogeneity in a mouse model of small cell lung cancer. *Cancer Cell,* **19,** 244-56.

Candelli, A., Holthausen, J. T., Depken, M., Brouwer, I., Franker, M. A., Marchetti, M., Heller, I., Bernard, S., Garcin, E. B., Modesti, M., Wyman, C., Wuite, G. J. & Peterman, E. J. 2014. Visualization and quantification of nascent RAD51 filament formation at single-monomer resolution. *Proc Natl Acad Sci U S A,* **111,** 15090-5.

Candiello, E., Reato, G., Verginelli, F., Gambardella, G., A, D. A., Calandra, N., Orzan, F., Iuliano, A., Albano, R., Sassi, F., Luraghi, P., Comoglio, P. M., Bertotti, A., Trusolino, L. & Boccaccio, C. 2023. MicroRNA 483-3p overexpression unleashes invasive growth of metastatic colorectal cancer via NDRG1 downregulation and ensuing activation of the ERBB3/AKT axis. *Mol Oncol,* **17,** 1280-1301.

Capone, E., Tryggvason, T., Cela, I., Dufrusine, B., Pinti, M., Del Pizzo, F., Gunnarsdottir, H. S., Grottola, T., De Laurenzi, V., Iacobelli, S., Lattanzio, R. & Sala, G. 2023. HER-3 surface expression increases in advanced colorectal cancer representing a potential therapeutic target. *Cell Death Discov,* **9,** 400.

Carnegie, G. K., Means, C. K. & Scott, J. D. 2009. A-kinase anchoring proteins: from protein complexes to physiology and disease. *IUBMB Life,* **61,** 394-406.

Cbioportal. 2023. *Colorectal Adenocarcinoma (TCGA, Pancancer Atlas)* [Online]. Available: https://www.cbioportal.org/study/summary?id=coadread_tcga_pan_can_atlas_2018 [Accessed 9 May 2023].

Chan, A. O., Broaddus, R. R., Houlihan, P. S., Issa, J. P., Hamilton, S. R. & Rashid, A. 2002. CpG island methylation in aberrant crypt foci of the colorectum. *Am J Pathol,* **160,** 1823-30.

Chan, D. K. H. & Buczacki, S. J. A. 2021. Tumour heterogeneity and evolutionary dynamics in colorectal cancer. *Oncogenesis,* **10,** 53.

Chan, D. K. H., Collins, S. D. & Buczacki, S. J. A. 2023. Generation and immunofluorescent validation of gene knockouts in adult human colonic organoids using multi-guide RNA CRISPR-Cas9. *STAR Protoc,* **4,** 101978.

Chan, D. K. H., Mandal, A., Hester, S., Yu, Z., Higgins, G. S., Kessler, B. M., Fischer, R. & Buczacki, S. J. A. 2023. Biallelic FBXW7 knockout induces AKAP8-mediated DNA damage in neighbouring wildtype cells. *Cell Death Discov,* **9,** 200.

Chan, D. S., Lau, R., Aune, D., Vieira, R., Greenwood, D. C., Kampman, E. & Norat, T. 2011. Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. *PLoS One,* **6,** e20456.

Chang, S. C., Lin, J. K., Lin, T. C. & Liang, W. Y. 2005. Loss of heterozygosity: an independent prognostic factor of colorectal cancer. *World J Gastroenterol,* **11,** 778-84.

Chen, K., Liu, X. H., Wang, F. R., Liu, H. P., Huang, Z. P. & Chen, X. 2018. The prognostic value of decreased NDRG1 expression in patients with digestive system cancers: A meta-analysis. *Medicine (Baltimore),* **97,** e12455.

Cleary, A. S., Leonard, T. L., Gestl, S. A. & Gunther, E. J. 2014. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature,* **508,** 113-7.

Clevers, H. 2016. Modeling Development and Disease with Organoids. *Cell,* **165,** 1586-1597.

Clevers, H., Loh, K. M. & Nusse, R. 2014. Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science,* **346,** 1248012.

Clister, T., Greenwald, E. C., Baillie, G. S. & Zhang, J. 2019. AKAP95 Organizes a Nuclear Microdomain to Control Local cAMP for Regulating Nuclear PKA. *Cell Chem Biol,* **26,** 885-891.e4.

Collas, P., Le Guellec, K. & Taskén, K. 1999. The A-kinase-anchoring protein AKAP95 is a multivalent protein with a key role in chromatin condensation at mitosis. *J Cell Biol,* **147,** 1167-80.

Colom, B., Herms, A., Hall, M. W. J., Dentro, S. C., King, C., Sood, R. K., Alcolea, M. P., Piedrafita, G., Fernandez-Antoran, D., Ong, S. H., Fowler, J. C., Mahbubani, K. T., Saeb-Parsy, K., Gerstung, M., Hall, B. A. & Jones, P. H. 2021. Mutant clones in normal epithelium outcompete and eliminate emerging tumours. *Nature,* **598,** 510-514.

Corces, M. R., Granja, J. M., Shams, S., Louie, B. H., Seoane, J. A., Zhou, W., Silva, T. C., Groeneveld, C., Wong, C. K., Cho, S. W., Satpathy, A. T., Mumbach, M. R., Hoadley, K. A., Robertson, A. G., Sheffield, N. C., Felau, I., Castro, M. a. A., Berman, B. P., Staudt, L. M., Zenklusen, J. C., Laird, P. W., Curtis, C., Greenleaf, W. J. & Chang, H. Y. 2018. The chromatin accessibility landscape of primary human cancers. *Science,* **362**.

Cornish, A. J., Gruber, A. J., Kinnersley, B., Chubb, D., Frangou, A., Caravagna, G., Noyvert, B., Lakatos, E., Wood, H. M., Arnedo-Pac, C., Culliford, R., Househam, J., Cross, W., Sud, A., Law, P., Leathlobhair, M. N., Hawari, A., Thorn, S., Sherwood, K., Gül, G., Fernandez-Tajes, J., Zapata, L., Alexandrov, L. B., Murugaesu, N., Sosinsky, A., Mitchell, J., Consortium, G. E. R., Lopez-Bigas, N., Quirke, P., Church, D. N., Tomlinson, I. P. M., Sottoriva, A., Graham, T. A., Wedge, D. C. & Houlston, R. S.

2022. Whole genome sequencing of 2,023 colorectal cancers reveals mutational landscapes, new driver genes and immune interactions. *bioRxiv***,** 2022.11.16.515599.

Counter, C. M., Hirte, H. W., Bacchetti, S. & Harley, C. B. 1994. Telomerase activity in human ovarian carcinoma. *Proc Natl Acad Sci U S A,* **91,** 2900-4.

Crockett, S. D. & Nagtegaal, I. D. 2019. Terminology, Molecular Features, Epidemiology, and Management of Serrated Colorectal Neoplasia. *Gastroenterology,* **157,** 949-966 e4.

Cross, W., Kovac, M., Mustonen, V., Temko, D., Davis, H., Baker, A. M., Biswas, S., Arnold, R., Chegwidden, L., Gatenbee, C., Anderson, A. R., Koelzer, V. H., Martinez, P., Jiang, X., Domingo, E., Woodcock, D. J., Feng, Y., Kovacova, M., Maughan, T., Consortium, S. C., Jansen, M., Rodriguez-Justo, M., Ashraf, S., Guy, R., Cunningham, C., East, J. E., Wedge, D. C., Wang, L. M., Palles, C., Heinimann, K., Sottoriva, A., Leedham, S. J., Graham, T. A. & Tomlinson, I. P. M. 2018a. The evolutionary landscape of colorectal tumorigenesis. *Nat Ecol Evol,* **2,** 1661-1672.

Cross, W., Kovac, M., Mustonen, V., Temko, D., Davis, H., Baker, A. M., Biswas, S., Arnold, R., Chegwidden, L., Gatenbee, C., Anderson, A. R., Koelzer, V. H., Martinez, P., Jiang, X., Domingo, E., Woodcock, D. J., Feng, Y., Kovacova, M., Maughan, T., Jansen, M., Rodriguez-Justo, M., Ashraf, S., Guy, R., Cunningham, C., East, J. E., Wedge, D. C., Wang, L. M., Palles, C., Heinimann, K., Sottoriva, A., Leedham, S. J., Graham, T. A. & Tomlinson, I. P. M. 2018b. The evolutionary landscape of colorectal tumorigenesis. *Nat Ecol Evol,* **2,** 1661-1672.

Cruk. *Bowel cancer incidence statistics* [Online]. Available: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer/incidence#heading-Four [Accessed 11 June 2023].

Cui, D., Xiong, X., Shu, J., Dai, X., Sun, Y. & Zhao, Y. 2020. FBXW7 Confers Radiation Survival by Targeting p53 for Degradation. *Cell Rep,* **30,** 497-509.e4.

Davies, H., Bignell, G. R., Cox, C., Stephens, P., Edkins, S., Clegg, S., Teague, J., Woffendin, H., Garnett, M. J., Bottomley, W., Davis, N., Dicks, E., Ewing, R., Floyd,

Y., Gray, K., Hall, S., Hawes, R., Hughes, J., Kosmidou, V., Menzies, A., Mould, C., Parker, A., Stevens, C., Watt, S., Hooper, S., Wilson, R., Jayatilake, H., Gusterson, B. A., Cooper, C., Shipley, J., Hargrave, D., Pritchard-Jones, K., Maitland, N., Chenevix-Trench, G., Riggins, G. J., Bigner, D. D., Palmieri, G., Cossu, A., Flanagan, A., Nicholson, A., Ho, J. W., Leung, S. Y., Yuen, S. T., Weber, B. L., Seigler, H. F., Darrow, T. L., Paterson, H., Marais, R., Marshall, C. J., Wooster, R., Stratton, M. R. & Futreal, P. A. 2002. Mutations of the BRAF gene in human cancer. *Nature,* **417,** 949-54.

Davis, H., Irshad, S., Bansal, M., Rafferty, H., Boitsova, T., Bardella, C., Jaeger, E., Lewis, A., Freeman-Mills, L., Giner, F. C., Rodenas-Cuadrado, P., Mallappa, S., Clark, S., Thomas, H., Jeffery, R., Poulsom, R., Rodriguez-Justo, M., Novelli, M., Chetty, R., Silver, A., Sansom, O. J., Greten, F. R., Wang, L. M., East, J. E., Tomlinson, I. & Leedham, S. J. 2015. Aberrant epithelial GREM1 expression initiates colonic tumorigenesis from cells outside the stem cell niche. *Nat Med,* **21,** 62-70.

Davis, H., Lewis, A., Behrens, A. & Tomlinson, I. 2014. Investigation of the atypical FBXW7 mutation spectrum in human tumours by conditional expression of a heterozygous propellor tip missense allele in the mouse intestines. *Gut,* **63,** 792-9.

Davis, H. & Tomlinson, I. 2012. CDC4/FBXW7 and the 'just enough' model of tumourigenesis. *J Pathol,* **227,** 131-5.

Davis, R. J., Gönen, M., Margineantu, D. H., Handeli, S., Swanger, J., Hoellerbauer, P., Paddison, P. J., Gu, H., Raftery, D., Grim, J. E., Hockenbery, D. M., Margolin, A. A. & Clurman, B. E. 2018. Pan-cancer transcriptional signatures predictive of oncogenic mutations reveal that Fbw7 regulates cancer cell oxidative metabolism. *Proc Natl Acad Sci U S A,* **115,** 5462-5467.

De Sousa E Melo, F., Kurtova, A. V., Harnoss, J. M., Kljavin, N., Hoeck, J. D., Hung, J., Anderson, J. E., Storm, E. E., Modrusan, Z., Koeppen, H., Dijkgraaf, G. J., Piskol, R. & De Sauvage, F. J. 2017. A distinct role for Lgr5(+) stem cells in primary and metastatic colon cancer. *Nature,* **543,** 676-680.

Di Giacomo, S., Sollazzo, M., De Biase, D., Ragazzi, M., Bellosta, P., Pession, A. & Grifoni, D. 2017. Human Cancer Cells Signal Their Competitive Fitness Through MYC Activity. *Sci Rep,* **7,** 12568.

Dodson, G. E., Limbo, O., Nieto, D. & Russell, P. 2010. Phosphorylation-regulated binding of Ctp1 to Nbs1 is critical for repair of DNA double-strand breaks. *Cell Cycle,* **9,** 1516-22.

Doxsey, S. 1998. The centrosome--a tiny organelle with big potential. *Nat Genet,* **20,** 104-6.

Drost, J., Van Jaarsveld, R. H., Ponsioen, B., Zimberlin, C., Van Boxtel, R., Buijs, A., Sachs, N., Overmeer, R. M., Offerhaus, G. J., Begthel, H., Korving, J., Van De Wetering, M., Schwank, G., Logtenberg, M., Cuppen, E., Snippert, H. J., Medema, J. P., Kops, G. J. & Clevers, H. 2015. Sequential cancer mutations in cultured human intestinal stem cells. *Nature,* **521,** 43-7.

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D. & Naldini, L. 1998. A third-generation lentivirus vector with a conditional packaging system. *J Virol,* **72,** 8463-71.

Duval, A. & Hamelin, R. 2002. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res,* **62,** 2447-54.

Eide, T., Taskén, K. A., Carlson, C., Williams, G., Jahnsen, T., Taskén, K. & Collas, P. 2003. Protein kinase A-anchoring protein AKAP95 interacts with MCM2, a regulator of DNA replication. *J Biol Chem,* **278,** 26750-6.

Fang, M., Ou, J., Hutchinson, L. & Green, M. R. 2014. The BRAF oncoprotein functions through the transcriptional repressor MAFG to mediate the CpG Island Methylator phenotype. *Mol Cell,* **55,** 904-915.

Fawole, A. S., Simpson, D. J., Rajagopal, R., Elder, J., Holland, T. A., Fryer, A., Deakin, M., Elder, J. B. & Farrell, W. E. 2002. Loss of heterozygosity on chromosome 10q is associated with earlier onset sporadic colorectal adenocarcinoma. *Int J Cancer,* **99,** 829-33.

Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W. & Et Al. 1990. Identification of a chromosome 18q gene that is altered in colorectal cancers. *Science,* **247,** 49-56.

Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F., Negri, E., Straif, K., Romieu, I., La Vecchia, C., Boffetta, P. & Jenab, M. 2011. Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Ann Oncol,* **22,** 1958-1972.

Ferguson, D. O., Sekiguchi, J. M., Chang, S., Frank, K. M., Gao, Y., Depinho, R. A. & Alt, F. W. 2000. The nonhomologous end-joining pathway of DNA repair is required for genomic stability and the suppression of translocations. *Proc Natl Acad Sci U S A,* **97,** 6630-3.

Fernandez Vallone, V., Leprovots, M., Strollo, S., Vasile, G., Lefort, A., Libert, F., Vassart, G. & Garcia, M. I. 2016. Trop2 marks transient gastric fetal epithelium and adult regenerating cells after epithelial damage. *Development,* **143,** 1452-63.

Fichtner, M., Bozkurt, E., Salvucci, M., Mccann, C., Mcallister, K. A., Halang, L., Düssmann, H., Kinsella, S., Crawford, N., Sessler, T., Longley, D. B. & Prehn, J. H. M. 2020. Molecular subtype-specific responses of colon cancer cells to the SMAC mimetic Birinapant. *Cell Death Dis,* **11,** 1020.

Fischer, M. 2019. Conservation and divergence of the p53 gene regulatory network between mice and humans. *Oncogene,* **38,** 4095-4109.

Fischer, M. 2021. Mice Are Not Humans: The Case of p53. *Trends in Cancer,* **7,** 12-14.

Flanagan, D. J., Pentinmikko, N., Luopajärvi, K., Willis, N. J., Gilroy, K., Raven, A. P., Mcgarry, L., Englund, J. I., Webb, A. T., Scharaw, S., Nasreddin, N., Hodder, M. C., Ridgway, R. A., Minnee, E., Sphyris, N., Gilchrist, E., Najumudeen, A. K., Romagnolo, B., Perret, C., Williams, A. C., Clevers, H., Nummela, P., Lähde, M., Alitalo, K., Hietakangas, V., Hedley, A., Clark, W., Nixon, C., Kirschner, K., Jones, E. Y., Ristimäki, A., Leedham, S. J., Fish, P. V., Vincent, J. P., Katajisto, P. &

Sansom, O. J. 2021a. NOTUM from Apc-mutant cells biases clonal competition to initiate cancer. *Nature,* **594,** 430-435.

Flanagan, D. J., Pentinmikko, N., Luopajärvi, K., Willis, N. J., Gilroy, K., Raven, A. P., Mcgarry, L., Englund, J. I., Webb, A. T., Scharaw, S., Nasreddin, N., Hodder, M. C., Ridgway, R. A., Minnee, E., Sphyris, N., Gilchrist, E., Najumudeen, A. K., Romagnolo, B., Perret, C., Williams, A. C., Clevers, H., Nummela, P., Lähde, M., Alitalo, K., Hietakangas, V., Hedley, A., Clark, W., Nixon, C., Kirschner, K., Jones, E. Y., Ristimäki, A., Leedham, S. J., Fish, P. V., Vincent, J. P., Katajisto, P. & Sansom, O. J. 2021b. NOTUM from Apc-mutant cells biases clonal competition to initiate cancer. *Nature*.

Galindo-Moreno, M., Giráldez, S., Limón-Mortés, M. C., Belmonte-Fernández, A., Reed, S. I., Sáez, C., Japón, M., Tortolero, M. & Romero, F. 2019. SCF(FBXW7)-mediated degradation of p53 promotes cell recovery after UV-induced DNA damage. *Faseb j,* **33,** 11420-11430.

García-Nieto, P. E., Morrison, A. J. & Fraser, H. B. 2019. The somatic mutation landscape of the human body. *Genome Biol,* **20,** 298.

Gavish, A., Tyler, M., Greenwald, A. C., Hoefflin, R., Simkin, D., Tschernichovsky, R., Galili Darnell, N., Somech, E., Barbolin, C., Antman, T., Kovarsky, D., Barrett, T., Gonzalez Castro, L. N., Halder, D., Chanoch-Myers, R., Laffy, J., Mints, M., Wider, A., Tal, R., Spitzer, A., Hara, T., Raitses-Gurevich, M., Stossel, C., Golan, T., Tirosh, A., Suvà, M. L., Puram, S. V. & Tirosh, I. 2023. Hallmarks of transcriptional intratumour heterogeneity across a thousand tumours. *Nature,* **618,** 598-606.

Gialeli, C., Theocharis, A. D. & Karamanos, N. K. 2011. Roles of matrix metalloproteinases in cancer progression and their pharmacological targeting. *Febs j,* **278,** 16-27.

Giannakis, M., Mu, X. J., Shukla, S. A., Qian, Z. R., Cohen, O., Nishihara, R., Bahl, S., Cao, Y., Amin-Mansour, A., Yamauchi, M., Sukawa, Y., Stewart, C., Rosenberg, M., Mima, K., Inamura, K., Nosho, K., Nowak, J. A., Lawrence, M. S., Giovannucci, E. L., Chan, A. T., Ng, K., Meyerhardt, J. A., Van Allen, E. M., Getz, G., Gabriel, S. B.,

Lander, E. S., Wu, C. J., Fuchs, C. S., Ogino, S. & Garraway, L. A. 2016. Genomic Correlates of Immune-Cell Infiltrates in Colorectal Carcinoma. *Cell Rep,* **15,** 857-865.

Giardiello, F. M., Allen, J. I., Axilbund, J. E., Boland, C. R., Burke, C. A., Burt, R. W., Church, J. M., Dominitz, J. A., Johnson, D. A., Kaltenbach, T., Levin, T. R., Lieberman, D. A., Robertson, D. J., Syngal, S. & Rex, D. K. 2014. Guidelines on genetic evaluation and management of Lynch syndrome: a consensus statement by the US Multi-Society Task Force on colorectal cancer. *Gastroenterology,* **147,** 502-26.

Grady, W. M. & Carethers, J. M. 2008. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology,* **135,** 1079-99.

Gregorieff, A., Liu, Y., Inanlou, M. R., Khomchuk, Y. & Wrana, J. L. 2015. Yap-dependent reprogramming of Lgr5(+) stem cells drives intestinal regeneration and cancer. *Nature,* **526,** 715-8.

Guiu, J., Hannezo, E., Yui, S., Demharter, S., Ulyanchenko, S., Maimets, M., Jørgensen, A., Perlman, S., Lundvall, L., Mamsen, L. S., Larsen, A., Olesen, R. H., Andersen, C. Y., Thuesen, L. L., Hare, K. J., Pers, T. H., Khodosevich, K., Simons, B. D. & Jensen, K. B. 2019. Tracing the origin of adult intestinal stem cells. *Nature,* **570,** 107-111.

Haince, J. F., Mcdonald, D., Rodrigue, A., Déry, U., Masson, J. Y., Hendzel, M. J. & Poirier, G. G. 2008. PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J Biol Chem,* **283,** 1197-208.

Hansen, S. L., Larsen, H. L., Pikkupeura, L. M., Maciag, G., Guiu, J., Müller, I., Clement, D. L., Mueller, C., Johansen, J. V., Helin, K., Lerdrup, M. & Jensen, K. B. 2023. An organoid-based CRISPR-Cas9 screen for regulators of intestinal epithelial maturation and cell fate. *Sci Adv,* **9,** eadg4055.

Hastie, N. D., Dempster, M., Dunlop, M. G., Thompson, A. M., Green, D. K. & Allshire, R. C. 1990. Telomere reduction in human colorectal carcinoma and with ageing. *Nature,* **346,** 866-8.

He, Y., Van't Veer, L. J., Mikolajewska-Hanclich, I., Van Velthuysen, M. L., Zeestraten, E. C., Nagtegaal, I. D., Van De Velde, C. J. & Marijnen, C. A. 2009. PIK3CA mutations predict local recurrences in rectal cancer patients. *Clin Cancer Res,* **15,** 6956-62.

Heide, T., Househam, J., Cresswell, G. D., Spiteri, I., Lynn, C., Mossner, M., Kimberley, C., Fernandez-Mateos, J., Chen, B., Zapata, L., James, C., Barozzi, I., Chkhaidze, K., Nichol, D., Gunasri, V., Berner, A., Schmidt, M., Lakatos, E., Baker, A. M., Costa, H., Mitchinson, M., Piazza, R., Jansen, M., Caravagna, G., Ramazzotti, D., Shibata, D., Bridgewater, J., Rodriguez-Justo, M., Magnani, L., Graham, T. A. & Sottoriva, A. 2022. The co-evolution of the genome and epigenome in colorectal cancer. *Nature,* **611,** 733-743.

Herbet, M., Salomon, A., Feige, J. J. & Thomas, M. 2012. Acquisition order of Ras and p53 gene alterations defines distinct adrenocortical tumor phenotypes. *PLoS Genet,* **8,** e1002700.

Herzog, F., Primorac, I., Dube, P., Lenart, P., Sander, B., Mechtler, K., Stark, H. & Peters, J. M. 2009. Structure of the anaphase-promoting complex/cyclosome interacting with a mitotic checkpoint complex. *Science,* **323,** 1477-81.

Hiramoto, T., Nakanishi, T., Sumiyoshi, T., Fukuda, T., Matsuura, S., Tauchi, H., Komatsu, K., Shibasaki, Y., Inui, H., Watatani, M., Yasutomi, M., Sumii, K., Kajiyama, G., Kamada, N., Miyagawa, K. & Kamiya, K. 1999. Mutations of a novel human RAD54 homologue, RAD54B, in primary cancer. *Oncogene,* **18,** 3422-6.

Hoang, M. L., Kinde, I., Tomasetti, C., Mcmahon, K. W., Rosenquist, T. A., Grollman, A. P., Kinzler, K. W., Vogelstein, B. & Papadopoulos, N. 2016. Genome-wide quantification of rare somatic mutations in normal human tissues using massively parallel sequencing. *Proc Natl Acad Sci U S A,* **113,** 9846-51.

Hobor, S., Van Emburgh, B. O., Crowley, E., Misale, S., Di Nicolantonio, F. & Bardelli, A. 2014. TGFα and amphiregulin paracrine network promotes resistance to EGFR blockade in colorectal cancer cells. *Clin Cancer Res,* **20,** 6429-38.

Hoeijmakers, J. H. 2001. Genome maintenance mechanisms for preventing cancer. *Nature,* **411,** 366-74.

Hofmarcher T, L. P. The Cost of Cancers of the Digestive System in Europe. Lund, Sweden: IHE.

Househam, J., Heide, T., Cresswell, G. D., Spiteri, I., Kimberley, C., Zapata, L., Lynn, C., James, C., Mossner, M., Fernandez-Mateos, J., Vinceti, A., Baker, A. M., Gabbutt, C., Berner, A., Schmidt, M., Chen, B., Lakatos, E., Gunasri, V., Nichol, D., Costa, H., Mitchinson, M., Ramazzotti, D., Werner, B., Iorio, F., Jansen, M., Caravagna, G., Barnes, C. P., Shibata, D., Bridgewater, J., Rodriguez-Justo, M., Magnani, L., Sottoriva, A. & Graham, T. A. 2022. Phenotypic plasticity and genetic control in colorectal cancer evolution. *Nature,* **611,** 744-753.

Hu, X. T., Xing, W., Zhao, R. S., Tan, Y., Wu, X. F., Ao, L. Q., Li, Z., Yao, M. W., Yuan, M., Guo, W., Li, S. Z., Yu, J., Ao, X. & Xu, X. 2020. HDAC2 inhibits EMT-mediated cancer metastasis by downregulating the long noncoding RNA H19 in colorectal cancer. *J Exp Clin Cancer Res,* **39,** 270.

Hu, Y. & Smyth, G. K. 2009. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods,* **347,** 70-8.

Huang, D., Sun, W., Zhou, Y., Li, P., Chen, F., Chen, H., Xia, D., Xu, E., Lai, M., Wu, Y. & Zhang, H. 2018. Mutations of key driver genes in colorectal cancer progression and metastasis. *Cancer Metastasis Rev,* **37,** 173-187.

Huen, M. S., Sy, S. M. & Chen, J. 2010. BRCA1 and its toolbox for the maintenance of genome integrity. *Nat Rev Mol Cell Biol,* **11,** 138-48.

Hung, M. C., Schechter, A. L., Chevray, P. Y., Stern, D. F. & Weinberg, R. A. 1986. Molecular cloning of the neu gene: absence of gross structural alteration in oncogenic alleles. *Proc Natl Acad Sci U S A,* **83,** 261-4.

Jass, J. R. 2007. Classification of colorectal cancer based on correlation of clinical, morphological and molecular features. *Histopathology,* **50,** 113-30.

Je, I. J., Vermeulen, L., Meijer, G. A. & Dekker, E. 2015. Serrated neoplasia-role in colorectal carcinogenesis and clinical implications. *Nat Rev Gastroenterol Hepatol,* **12,** 401-9.

Jiang, Y., Ben, Q., Shen, H., Lu, W., Zhang, Y. & Zhu, J. 2011. Diabetes mellitus and incidence and mortality of colorectal cancer: a systematic review and meta-analysis of cohort studies. *Eur J Epidemiol,* **26,** 863-76.

Jin, D. Y., Spencer, F. & Jeang, K. T. 1998. Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. *Cell,* **93,** 81-91.

Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol,* **7,** 335-46.

Johnson, M. S., Reddy, G. & Desai, M. M. 2023. Epistasis and evolution: recent advances and an outlook for prediction. *BMC Biol,* **21,** 120.

Johnstone, S. E., Reyes, A., Qi, Y., Adriaens, C., Hegazi, E., Pelka, K., Chen, J. H., Zou, L. S., Drier, Y., Hecht, V., Shoresh, N., Selig, M. K., Lareau, C. A., Iyer, S., Nguyen, S. C., Joyce, E. F., Hacohen, N., Irizarry, R. A., Zhang, B., Aryee, M. J. & Bernstein, B. E. 2020. Large-Scale Topological Changes Restrain Malignant Progression in Colorectal Cancer. *Cell,* **182,** 1474-1489.e23.

Kambara, T., Simms, L. A., Whitehall, V. L., Spring, K. J., Wynter, C. V., Walsh, M. D., Barker, M. A., Arnold, S., Mcgivern, A., Matsubara, N., Tanaka, N., Higuchi, T., Young, J., Jass, J. R. & Leggett, B. A. 2004. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. *Gut,* **53,** 1137-44.

Karczmarski, J., Goryca, K., Pachlewski, J., Dabrowska, M., Pysniak, K., Paziewska, A., Kulecka, M., Lenarcik, M., Mroz, A., Mikula, M. & Ostrowski, J. 2019. Mutation Profiling of Premalignant Colorectal Neoplasia. *Gastroenterol Res Pract,* **2019,** 2542640.

Kawaguchi, Y., Newhook, T. E., Tran Cao, H. S., Tzeng, C. D., Chun, Y. S., Aloia, T. A., Dasari, A., Kopetz, S. & Vauthey, J. N. 2021. Alteration of FBXW7 is Associated with Worse Survival in Patients Undergoing Resection of Colorectal Liver Metastases. *J Gastrointest Surg,* **25,** 186-194.

Kawashita, Y., Morine, Y., Ikemoto, T., Saito, Y., Iwahashi, S., Yamada, S., Higashijima, J., Imura, S., Ogawa, H., Yagi, T. & Shimada, M. 2017. Loss of Fbxw7 expression is a predictor of recurrence in colorectal liver metastasis. *J Hepatobiliary Pancreat Sci,* **24,** 576-583.

Kent, D. G. & Green, A. R. 2017. Order Matters: The Order of Somatic Mutations Influences Cancer Evolution. *Cold Spring Harb Perspect Med,* **7**.

Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack, A. D., Fan, H., Shen, H., Way, G. P., Greene, C. S., Liu, Y., Akbani, R., Feng, B., Donehower, L. A., Miller, C., Shen, Y., Karimi, M., Chen, H., Kim, P., Jia, P., Shinbrot, E., Zhang, S., Liu, J., Hu, H., Bailey, M. H., Yau, C., Wolf, D., Zhao, Z., Weinstein, J. N., Li, L., Ding, L., Mills, G. B., Laird, P. W., Wheeler, D. A., Shmulevich, I., Monnat, R. J., Jr., Xiao, Y. & Wang, C. 2018. Genomic and Molecular Landscape of DNA Damage Repair Deficiency across The Cancer Genome Atlas. *Cell Rep,* **23,** 239-254.e6.

Kong, X. Y., Zhang, D. C., Zhuang, W. X., Hua, S. H., Dai, Y., Yuan, Y. Y., Feng, L. L., Huang, Q., Teng, B. G., Yu, X. Y., Liu, W. Z. & Zhang, Y. X. 2016. AKAP95 promotes cell cycle progression via interactions with cyclin E and low molecular weight cyclin E. *Am J Transl Res,* **8,** 811-26.

Koo, B. K., Spit, M., Jordens, I., Low, T. Y., Stange, D. E., Van De Wetering, M., Van Es, J. H., Mohammed, S., Heck, A. J., Maurice, M. M. & Clevers, H. 2012. Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature,* **488,** 665-9.

Kopan, R. & Ilagan, M. X. 2009. The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell,* **137,** 216-33.

Kopetz, S., Grothey, A., Yaeger, R., Van Cutsem, E., Desai, J., Yoshino, T., Wasan, H., Ciardiello, F., Loupakis, F., Hong, Y. S., Steeghs, N., Guren, T. K., Arkenau, H. T., Garcia-Alfonso, P., Pfeiffer, P., Orlov, S., Lonardi, S., Elez, E., Kim, T. W., Schellens, J. H. M., Guo, C., Krishnan, A., Dekervel, J., Morris, V., Calvo Ferrandiz,

A., Tarpgaard, L. S., Braun, M., Gollerkeri, A., Keir, C., Maharry, K., Pickard, M., Christy-Bittel, J., Anderson, L., Sandor, V. & Tabernero, J. 2019. Encorafenib, Binimetinib, and Cetuximab in BRAF V600E-Mutated Colorectal Cancer. *N Engl J Med,* **381,** 1632-1643.

Korphaisarn, K., Morris, V. K., Overman, M. J., Fogelman, D. R., Kee, B. K., Raghav, K. P. S., Manuel, S., Shureiqi, I., Wolff, R. A., Eng, C., Menter, D., Hamilton, S. R., Kopetz, S. & Dasari, A. 2017. FBXW7 missense mutation: a novel negative prognostic factor in metastatic colorectal adenocarcinoma. *Oncotarget,* **8,** 39268-39279.

Kristensen, H., Thyø, A., Emmertsen, K. J., Smart, N. J., Pinkney, T., Warwick, A. M., Pang, D., Elfeki, H., Shalaby, M., Emile, S. H., Abdelkhalek, M., Zuhdy, M., Poskus, T., Dulskas, A., Horesh, N., Furnée, E. J. B., Verkuijl, S. J., Rama, N. J., Domingos, H., Maciel, J., Solis-Peña, A., Espín-Basany, E., Hidalgo-Pujol, M., Biondo, S., Sjövall, A. & Christensen, P. 2022. Surviving rectal cancer at the cost of a colostomy: global survey of long-term health-related quality of life in 10 countries. *BJS Open,* **6**.

Krupa, R., Sliwinski, T., Wisniewska-Jarosinska, M., Chojnacki, J., Wasylecka, M., Dziki, L., Morawiec, J. & Blasiak, J. 2011. Polymorphisms in RAD51, XRCC2 and XRCC3 genes of the homologous recombination repair in colorectal cancer--a case control study. *Mol Biol Rep,* **38,** 2849-54.

Lammi, L., Arte, S., Somer, M., Jarvinen, H., Lahermo, P., Thesleff, I., Pirinen, S. & Nieminen, P. 2004. Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *Am J Hum Genet,* **74,** 1043-50.

Lasko, D., Cavenee, W. & Nordenskjöld, M. 1991. Loss of constitutional heterozygosity in human cancer. *Annu Rev Genet,* **25,** 281-314.

Lee-Six, H., Olafsson, S., Ellis, P., Osborne, R. J., Sanders, M. A., Moore, L., Georgakopoulos, N., Torrente, F., Noorani, A., Goddard, M., Robinson, P., Coorens, T. H. H., O'neill, L., Alder, C., Wang, J., Fitzgerald, R. C., Zilbauer, M., Coleman, N., Saeb-Parsy, K., Martincorena, I., Campbell, P. J. & Stratton, M. R. 2019. The

landscape of somatic mutation in normal colorectal epithelial cells. *Nature,* **574,** 532-537.

Li, R., Di, L., Li, J., Fan, W., Liu, Y., Guo, W., Liu, W., Liu, L., Li, Q., Chen, L., Chen, Y., Miao, C., Liu, H., Wang, Y., Ma, Y., Xu, D., Lin, D., Huang, Y., Wang, J., Bai, F. & Wu, C. 2021. A body map of somatic mutagenesis in morphologically normal human tissues. *Nature,* **597,** 398-403.

Li, W., Hu, J., Shi, B., Palomba, F., Digman, M. A., Gratton, E. & Jiang, H. 2020. Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. *Nat Cell Biol,* **22,** 960-972.

Li, Y. & Benezra, R. 1996. Identification of a human mitotic checkpoint gene: hsMAD2. *Science,* **274,** 246-8.

Liang, Z. & Li, X. 2021. Identification of ANXA1 as a potential prognostic biomarker and correlating with immune infiltrates in colorectal cancer. *Autoimmunity,* **54,** 76-87.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P. & Tamayo, P. 2015. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst,* **1,** 417-425.

Lin, K. C., Park, H. W. & Guan, K. L. 2017. Regulation of the Hippo Pathway Transcription Factor TEAD. *Trends Biochem Sci,* **42,** 862-872.

Liu, H., Gu, X., Wang, G., Huang, Y., Ju, S., Huang, J. & Wang, X. 2019. Copy number variations primed lncRNAs deregulation contribute to poor prognosis in colorectal cancer. *Aging (Albany NY),* **11,** 6089-6108.

Liu, W., Dong, X., Mai, M., Seelan, R. S., Taniguchi, K., Krishnadath, K. K., Halling, K. C., Cunningham, J. M., Boardman, L. A., Qian, C., Christensen, E., Schmidt, S. S., Roche, P. C., Smith, D. I. & Thibodeau, S. N. 2000. Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating beta-catenin/TCF signalling. *Nat Genet,* **26,** 146-7.

Liu, Z., Gu, Y., Li, X., Zhou, L., Cheng, X., Jiang, H., Huang, Y., Zhang, Y., Xu, T., Yang, W. & Huang, Q. 2022. Mucin 16 Promotes Colorectal Cancer Development and Progression Through Activation of Janus Kinase 2. *Dig Dis Sci,* **67,** 2195-2208.

Liu, Z., Liu, L., Jiao, D., Guo, C., Wang, L., Li, Z., Sun, Z., Zhao, Y. & Han, X. 2021. Association of RYR2 Mutation With Tumor Mutation Burden, Prognosis, and Antitumor Immunity in Patients With Esophageal Adenocarcinoma. *Front Genet,* **12,** 669694.

Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. 2010. Intestinal stem cell replacement follows a pattern of neutral drift. *Science,* **330,** 822-5.

Losi, L., Baisse, B., Bouzourene, H. & Benhattar, J. 2005. Evolution of intratumoral genetic heterogeneity during colorectal cancer progression. *Carcinogenesis,* **26,** 916-22.

Ma, X., Dong, L., Liu, X., Ou, K. & Yang, L. 2022. POLE/POLD1 mutation and tumor immunotherapy. *J Exp Clin Cancer Res,* **41,** 216.

Ma, Y., Yang, Y., Wang, F., Zhang, P., Shi, C., Zou, Y. & Qin, H. 2013. Obesity and risk of colorectal cancer: a systematic review of prospective studies. *PLoS One,* **8,** e53916.

Mah, A. T., Yan, K. S. & Kuo, C. J. 2016. Wnt pathway regulation of intestinal stem cells. *J Physiol,* **594,** 4837-47.

Manceau, G., Marisa, L., Boige, V., Duval, A., Gaub, M. P., Milano, G., Selves, J., Olschwang, S., Jooste, V., Le Legrain, M., Lecorre, D., Guenot, D., Etienne-Grimaldi, M. C., Kirzin, S., Martin, L., Lepage, C., Bouvier, A. M. & Laurent-Puig, P. 2015. PIK3CA mutations predict recurrence in localized microsatellite stable colon cancer. *Cancer Med,* **4,** 371-82.

Maréchal, A. & Zou, L. 2013. DNA damage sensing by the ATM and ATR kinases. *Cold Spring Harb Perspect Biol,* **5**.

Martincorena, I., Fowler, J. C., Wabik, A., Lawson, A. R. J., Abascal, F., Hall, M. W. J., Cagan, A., Murai, K., Mahbubani, K., Stratton, M. R., Fitzgerald, R. C., Handford, P. A., Campbell, P. J., Saeb-Parsy, K. & Jones, P. H. 2018. Somatic mutant clones colonize the human esophagus with age. *Science,* **362,** 911-917.

Martincorena, I., Roshan, A., Gerstung, M., Ellis, P., Van Loo, P., Mclaren, S., Wedge, D. C., Fullam, A., Alexandrov, L. B., Tubio, J. M., Stebbings, L., Menzies, A., Widaa, S., Stratton, M. R., Jones, P. H. & Campbell, P. J. 2015. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science,* **348,** 880-6.

Maser, R. S. & Depinho, R. A. 2002. Connecting chromosomes, crisis, and cancer. *Science,* **297,** 565-9.

Matano, M., Date, S., Shimokawa, M., Takano, A., Fujii, M., Ohta, Y., Watanabe, T., Kanai, T. & Sato, T. 2015. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat Med,* **21,** 256-62.

Mathieu, N., Pirzio, L., Freulet-Marriere, M. A., Desmaze, C. & Sabatier, L. 2004. Telomeres and chromosomal instability. *Cell Mol Life Sci,* **61,** 641-56.

Matsumoto, A., Onoyama, I. & Nakayama, K. I. 2006. Expression of mouse Fbxw7 isoforms is regulated in a cell cycle- or p53-dependent manner. *Biochem Biophys Res Commun,* **350,** 114-9.

Mcclintock, B. 1941. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics,* **26,** 234-82.

Merlos-Suárez, A., Barriga, F. M., Jung, P., Iglesias, M., Céspedes, M. V., Rossell, D., Sevillano, M., Hernando-Momblona, X., Da Silva-Diz, V., Muñoz, P., Clevers, H., Sancho, E., Mangues, R. & Batlle, E. 2011. The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell,* **8,** 511-24.

Monaco, G., Van Dam, S., Casal Novo Ribeiro, J. L., Larbi, A. & De Magalhães, J. P. 2015. A comparison of human and mouse gene co-expression networks reveals conservation and divergence at the tissue, pathway and disease levels. *BMC Evol Biol,* **15,** 259.

Moore, L., Leongamornlert, D., Coorens, T. H. H., Sanders, M. A., Ellis, P., Dentro, S. C., Dawson, K. J., Butler, T., Rahbari, R., Mitchell, T. J., Maura, F., Nangalia, J., Tarpey,

P. S., Brunner, S. F., Lee-Six, H., Hooks, Y., Moody, S., Mahbubani, K. T., Jimenez-Linan, M., Brosens, J. J., Iacobuzio-Donahue, C. A., Martincorena, I., Saeb-Parsy, K., Campbell, P. J. & Stratton, M. R. 2020. The mutational landscape of normal human endometrial epithelium. *Nature,* **580,** 640-646.

Morgan, E., Arnold, M., Gini, A., Lorenzoni, V., Cabasag, C. J., Laversanne, M., Vignat, J., Ferlay, J., Murphy, N. & Bray, F. 2023. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. *Gut,* **72,** 338-344.

Morral, C., Stanisavljevic, J., Hernando-Momblona, X., Mereu, E., Álvarez-Varela, A., Cortina, C., Stork, D., Slebe, F., Turon, G., Whissell, G., Sevillano, M., Merlos-Suárez, A., Casanova-Martí, À., Moutinho, C., Lowe, S. W., Dow, L. E., Villanueva, A., Sancho, E., Heyn, H. & Batlle, E. 2020. Zonation of Ribosomal DNA Transcription Defines a Stem Cell Hierarchy in Colorectal Cancer. *Cell Stem Cell,* **26,** 845-861.e12.

Mustata, R. C., Vasile, G., Fernandez-Vallone, V., Strollo, S., Lefort, A., Libert, F., Monteyne, D., Pérez-Morga, D., Vassart, G. & Garcia, M. I. 2013. Identification of Lgr5-independent spheroid-generating progenitors of the mouse fetal intestinal epithelium. *Cell Rep,* **5,** 421-32.

Nagtegaal, I. D., Odze, R. D., Klimstra, D., Paradis, V., Rugge, M., Schirmacher, P., Washington, K. M., Carneiro, F. & Cree, I. A. 2020. The 2019 WHO classification of tumours of the digestive system. *Histopathology,* **76,** 182-188.

Ng, A. S. & Chan, D. K. H. 2023. Commonalities and differences in the mutational signature and somatic driver mutation landscape across solid and hollow viscus organs. *Oncogene,* **42,** 2713-2724.

Nguyen, L. H., Goel, A. & Chung, D. C. 2020. Pathways of Colorectal Carcinogenesis. *Gastroenterology,* **158,** 291-302.

Nicholson, A. M., Olpe, C., Hoyle, A., Thorsen, A. S., Rus, T., Colombé, M., Brunton-Sim, R., Kemp, R., Marks, K., Quirke, P., Malhotra, S., Ten Hoopen, R., Ibrahim, A., Lindskog, C., Myers, M. B., Parsons, B., Tavaré, S., Wilkinson, M., Morrissey, E. &

Winton, D. J. 2018. Fixation and Spread of Somatic Mutations in Adult Human Colonic Epithelium. *Cell Stem Cell,* **22,** 909-918.e8.

Nusse, Y. M., Savage, A. K., Marangoni, P., Rosendahl-Huber, A. K. M., Landman, T. A., De Sauvage, F. J., Locksley, R. M. & Klein, O. D. 2018. Parasitic helminths induce fetal-like reversion in the intestinal stem cell niche. *Nature,* **559,** 109-113.

O'brien, M. J., Yang, S., Clebanoff, J. L., Mulcahy, E., Farraye, F. A., Amorosino, M. & Swan, N. 2004. Hyperplastic (serrated) polyps of the colorectum: relationship of CpG island methylator phenotype and K-ras mutation to location and histologic subtype. *Am J Surg Pathol,* **28,** 423-34.

O'brien, M. J., Yang, S., Mack, C., Xu, H., Huang, C. S., Mulcahy, E., Amorosino, M. & Farraye, F. A. 2006. Comparison of microsatellite instability, CpG island methylation phenotype, BRAF and KRAS status in serrated polyps and traditional adenomas indicates separate pathways to distinct colorectal carcinoma end points. *Am J Surg Pathol,* **30,** 1491-501.

Ogino, S., Lochhead, P., Giovannucci, E., Meyerhardt, J. A., Fuchs, C. S. & Chan, A. T. 2014. Discovery of colorectal cancer PIK3CA mutation as potential predictive biomarker: power and promise of molecular pathological epidemiology. *Oncogene,* **33,** 2949-55.

Olpe, C., Khamis, D., Chukanova, M., Skoufou-Papoutsaki, N., Kemp, R., Marks, K., Tatton, C., Lindskog, C., Nicholson, A., Brunton-Sim, R., Malhotra, S., Ten Hoopen, R., Stanley, R., Winton, D. J. & Morrissey, E. 2021. A Diffusion-like Process Accommodates New Crypts During Clonal Expansion in Human Colonic Epithelium. *Gastroenterology,* **161,** 548-559.e23.

Ortmann, C. A., Kent, D. G., Nangalia, J., Silber, Y., Wedge, D. C., Grinfeld, J., Baxter, E. J., Massie, C. E., Papaemmanuil, E., Menon, S., Godfrey, A. L., Dimitropoulou, D., Guglielmelli, P., Bellosillo, B., Besses, C., Döhner, K., Harrison, C. N., Vassiliou, G. S., Vannucchi, A., Campbell, P. J. & Green, A. R. 2015. Effect of mutation order on myeloproliferative neoplasms. *N Engl J Med,* **372,** 601-612.

Pal, T., Permuth-Wey, J., Kumar, A. & Sellers, T. A. 2008. Systematic review and meta-analysis of ovarian cancers: estimation of microsatellite-high frequency and characterization of mismatch repair deficient tumor histology. *Clin Cancer Res,* **14,** 6847-54.

Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'arrigo, A., Truong, O., Hsuan, J. J. & Jiricny, J. 1995. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science,* **268,** 1912-4.

Parker, T. W. & Neufeld, K. L. 2020. APC controls Wnt-induced β-catenin destruction complex recruitment in human colonocytes. *Sci Rep,* **10,** 2957.

Penn, L. J., Brooks, M. W., Laufer, E. M. & Land, H. 1990. Negative autoregulation of c-myc transcription. *Embo j,* **9,** 1113-21.

Pikkupeura, L. M., Bressan, R. B., Guiu, J., Chen, Y., Maimets, M., Mayer, D., Schweiger, P. J., Hansen, S. L., Maciag, G. J., Larsen, H. L., Lõhmussaar, K., Pedersen, M. T., Teves, J. M. Y., Bornholdt, J., Benes, V., Sandelin, A. & Jensen, K. B. 2023. Transcriptional and epigenomic profiling identifies YAP signaling as a key regulator of intestinal epithelium maturation. *Sci Adv,* **9,** eadf9460.

Pino, M. S. & Chung, D. C. 2010. The chromosomal instability pathway in colon cancer. *Gastroenterology,* **138,** 2059-72.

Qi, Z., Redding, S., Lee, J. Y., Gibb, B., Kwon, Y., Niu, H., Gaines, W. A., Sung, P. & Greene, E. C. 2015. DNA sequence alignment by microhomology sampling during homologous recombination. *Cell,* **160,** 856-869.

Ranjha, L., Howard, S. M. & Cejka, P. 2018. Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma,* **127,** 187-214.

Raskov, H., Soby, J. H., Troelsen, J., Bojesen, R. D. & Gogenur, I. 2020. Driver Gene Mutations and Epigenetics in Colorectal Cancer. *Ann Surg,* **271,** 75-85.

Reilly, N. M., Novara, L., Di Nicolantonio, F. & Bardelli, A. 2019. Exploiting DNA repair defects in colorectal cancer. *Mol Oncol,* **13,** 681-700.

Remvikos, Y., Muleris, M., Salmon, R. J. & Dutrillaux, B. 1997. Colorectal carcinogenesis: from chromosomal evolution pathways to molecular pathogenesis. *Cancer Genet Cytogenet,* **93,** 63-73.

Romero-Pérez, L., Surdez, D., Brunet, E., Delattre, O. & Grünewald, T. G. P. 2019. STAG Mutations in Cancer. *Trends Cancer,* **5,** 506-520.

Sabharwal, S. S. & Schumacker, P. T. 2014. Mitochondrial ROS in cancer: initiators, amplifiers or an Achilles' heel? *Nat Rev Cancer,* **14,** 709-21.

Sachdeva, M., Zhu, S., Wu, F., Wu, H., Walia, V., Kumar, S., Elble, R., Watabe, K. & Mo, Y. Y. 2009. p53 represses c-Myc through induction of the tumor suppressor miR-145. *Proc Natl Acad Sci U S A,* **106,** 3207-12.

Saito, T., Niida, A., Uchi, R., Hirata, H., Komatsu, H., Sakimura, S., Hayashi, S., Nambara, S., Kuroda, Y., Ito, S., Eguchi, H., Masuda, T., Sugimachi, K., Tobo, T., Nishida, H., Daa, T., Chiba, K., Shiraishi, Y., Yoshizato, T., Kodama, M., Okimoto, T., Mizukami, K., Ogawa, R., Okamoto, K., Shuto, M., Fukuda, K., Matsui, Y., Shimamura, T., Hasegawa, T., Doki, Y., Nagayama, S., Yamada, K., Kato, M., Shibata, T., Mori, M., Aburatani, H., Murakami, K., Suzuki, Y., Ogawa, S., Miyano, S. & Mimori, K. 2018. A temporal shift of the evolutionary principle shaping intratumor heterogeneity in colorectal cancer. *Nat Commun,* **9,** 2884.

Sancho, R., Jandke, A., Davis, H., Diefenbacher, M. E., Tomlinson, I. & Behrens, A. 2010. F-box and WD repeat domain-containing 7 regulates intestinal cell lineage commitment and is a haploinsufficient tumor suppressor. *Gastroenterology,* **139,** 929-41.

Sato, T., Stange, D. E., Ferrante, M., Vries, R. G., Van Es, J. H., Van Den Brink, S., Van Houdt, W. J., Pronk, A., Van Gorp, J., Siersema, P. D. & Clevers, H. 2011. Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium. *Gastroenterology,* **141,** 1762-72.

Schell, M. J., Yang, M., Teer, J. K., Lo, F. Y., Madan, A., Coppola, D., Monteiro, A. N., Nebozhyn, M. V., Yue, B., Loboda, A., Bien-Willner, G. A., Greenawalt, D. M. &

Yeatman, T. J. 2016. A multigene mutation classification of 468 colorectal cancers reveals a prognostic role for APC. *Nat Commun,* **7,** 11743.

Schwitalla, S., Fingerle, A. A., Cammareri, P., Nebelsiek, T., Göktuna, S. I., Ziegler, P. K., Canli, O., Heijmans, J., Huels, D. J., Moreaux, G., Rupec, R. A., Gerhard, M., Schmid, R., Barker, N., Clevers, H., Lang, R., Neumann, J., Kirchner, T., Taketo, M. M., Van Den Brink, G. R., Sansom, O. J., Arkan, M. C. & Greten, F. R. 2013. Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell,* **152,** 25-38.

Secombe, J., Pierce, S. B. & Eisenman, R. N. 2004. Myc: a weapon of mass destruction. *Cell,* **117,** 153-6.

Shadel, G. S. & Horvath, T. L. 2015. Mitochondrial ROS signaling in organismal homeostasis. *Cell,* **163,** 560-9.

Shang, W., Yan, C., Liu, R., Chen, L., Cheng, D., Hao, L., Yuan, W., Chen, J. & Yang, H. 2021. Clinical significance of FBXW7 tumor suppressor gene mutations and expression in human colorectal cancer: a systemic review and meta-analysis. *BMC Cancer,* **21,** 770.

Shi, X., Li, Y., Yuan, Q., Tang, S., Guo, S., Zhang, Y., He, J., Zhang, X., Han, M., Liu, Z., Zhu, Y., Gao, S., Wang, H., Xu, X., Zheng, K., Jing, W., Chen, L., Wang, Y., Jin, G. & Gao, D. 2022. Integrated profiling of human pancreatic cancer organoids reveals chromatin accessibility features associated with drug sensitivity. *Nat Commun,* **13,** 2169.

Shimokawa, M., Ohta, Y., Nishikori, S., Matano, M., Takano, A., Fujii, M., Date, S., Sugimoto, S., Kanai, T. & Sato, T. 2017. Visualization and targeting of LGR5(+) human colon cancer stem cells. *Nature,* **545,** 187-192.

Sinicrope, F. A. 2018. Lynch Syndrome-Associated Colorectal Cancer. *N Engl J Med,* **379,** 764-773.

Siraj, A. K., Pratheeshkumar, P., Divya, S. P., Parvathareddy, S. K., Bu, R., Masoodi, T., Kong, Y., Thangavel, S., Al-Sanea, N., Ashari, L. H., Abduljabbar, A., Al-Homoud,

S., Al-Dayel, F. & Al-Kuraya, K. S. 2019. TGFβ-induced SMAD4-dependent Apoptosis Proceeded by EMT in CRC. *Mol Cancer Ther,* **18,** 1312-1322.

Snippert, H. J., Schepers, A. G., Van Es, J. H., Simons, B. D. & Clevers, H. 2014. Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion. *EMBO Rep,* **15,** 62-9.

Snippert, H. J., Van Der Flier, L. G., Sato, T., Van Es, J. H., Van Den Born, M., Kroon-Veenboer, C., Barker, N., Klein, A. M., Van Rheenen, J., Simons, B. D. & Clevers, H. 2010. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell,* **143,** 134-44.

Snover, D. C. 2011. Update on the serrated pathway to colorectal carcinoma. *Hum Pathol,* **42,** 1-10.

Song, Y., Li, L., Ou, Y., Gao, Z., Li, E., Li, X., Zhang, W., Wang, J., Xu, L., Zhou, Y., Ma, X., Liu, L., Zhao, Z., Huang, X., Fan, J., Dong, L., Chen, G., Ma, L., Yang, J., Chen, L., He, M., Li, M., Zhuang, X., Huang, K., Qiu, K., Yin, G., Guo, G., Feng, Q., Chen, P., Wu, Z., Wu, J., Ma, L., Zhao, J., Luo, L., Fu, M., Xu, B., Chen, B., Li, Y., Tong, T., Wang, M., Liu, Z., Lin, D., Zhang, X., Yang, H., Wang, J. & Zhan, Q. 2014. Identification of genomic alterations in oesophageal squamous cell cancer. *Nature,* **509,** 91-5.

Sottoriva, A., Kang, H., Ma, Z., Graham, T. A., Salomon, M. P., Zhao, J., Marjoram, P., Siegmund, K., Press, M. F., Shibata, D. & Curtis, C. 2015. A Big Bang model of human colorectal tumor growth. *Nat Genet,* **47,** 209-16.

Spring, K. J., Zhao, Z. Z., Karamatic, R., Walsh, M. D., Whitehall, V. L., Pike, T., Simms, L. A., Young, J., James, M., Montgomery, G. W., Appleyard, M., Hewett, D., Togashi, K., Jass, J. R. & Leggett, B. A. 2006. High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. *Gastroenterology,* **131,** 1400-7.

Stein, R. A. & Staros, J. V. 2000. Evolutionary analysis of the ErbB receptor and ligand families. *J Mol Evol,* **50,** 397-412.

Stirling, D. R., Swain-Bowden, M. J., Lucas, A. M., Carpenter, A. E., Cimini, B. A. & Goodman, A. 2021. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics,* **22,** 433.

Strand, M., Earley, M. C., Crouse, G. F. & Petes, T. D. 1995. Mutations in the MSH3 gene preferentially lead to deletions within tracts of simple repetitive DNA in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A,* **92,** 10418-21.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A. & Bray, F. 2021. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin,* **71,** 209-249.

Ter Steege, E. J. & Bakker, E. R. M. 2021. The role of R-spondin proteins in cancer biology. *Oncogene,* **40,** 6469-6478.

Thiagalingam, S., Laken, S., Willson, J. K., Markowitz, S. D., Kinzler, K. W., Vogelstein, B. & Lengauer, C. 2001. Mechanisms underlying losses of heterozygosity in human colorectal cancers. *Proc Natl Acad Sci U S A,* **98,** 2698-702.

Thibodeau, S. N., Bren, G. & Schaid, D. 1993. Microsatellite instability in cancer of the proximal colon. *Science,* **260,** 816-9.

Thirimanne, H. N., Wu, F., Janssens, D. H., Swanger, J., Diab, A., Feldman, H. M., Amezquita, R. A., Gottardo, R., Paddison, P. J., Henikoff, S. & Clurman, B. E. 2022. Global and context-specific transcriptional consequences of oncogenic Fbw7 mutations. *Elife,* **11**.

Thompson, L. L., Rutherford, K. A., Lepage, C. C. & Mcmanus, K. J. 2021. The SCF Complex Is Essential to Maintain Genome and Chromosome Stability. *Int J Mol Sci,* **22**.

Toyota, M., Ahuja, N., Ohe-Toyota, M., Herman, J. G., Baylin, S. B. & Issa, J. P. 1999. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A,* **96,** 8681-6.

Tripathi, V., Kaur, E., Kharat, S. S., Hussain, M., Damodaran, A. P., Kulshrestha, S. & Sengupta, S. 2019. Abrogation of FBW7α-dependent p53 degradation enhances p53's function as a tumor suppressor. *J Biol Chem,* **294,** 13224-13232.

Uchi, R., Takahashi, Y., Niida, A., Shimamura, T., Hirata, H., Sugimachi, K., Sawada, G., Iwaya, T., Kurashige, J., Shinden, Y., Iguchi, T., Eguchi, H., Chiba, K., Shiraishi, Y., Nagae, G., Yoshida, K., Nagata, Y., Haeno, H., Yamamoto, H., Ishii, H., Doki, Y., Iinuma, H., Sasaki, S., Nagayama, S., Yamada, K., Yachida, S., Kato, M., Shibata, T., Oki, E., Saeki, H., Shirabe, K., Oda, Y., Maehara, Y., Komune, S., Mori, M., Suzuki, Y., Yamamoto, K., Aburatani, H., Ogawa, S., Miyano, S. & Mimori, K. 2016. Integrated Multiregional Analysis Proposing a New Model of Colorectal Cancer Evolution. *PLoS Genet,* **12,** e1005778.

Uchida, S. & Sugino, T. 2023. ERBB2-Mutant Gastrointestinal Tumors Represent Heterogeneous Molecular Biology, Particularly in Microsatellite Instability, Tumor Mutation Burden, and Co-Mutated Genes: An In Silico Study. *Curr Issues Mol Biol,* **45,** 7404-7416.

Umar, A., Risinger, J. I., Hawk, E. T. & Barrett, J. C. 2004. Testing guidelines for hereditary non-polyposis colorectal cancer. *Nat Rev Cancer,* **4,** 153-8.

Vaghi, C., Mauri, G., Agostara, A. G., Patelli, G., Pizzutilo, E. G., Nakamura, Y., Yoshino, T., Siena, S. & Sartore-Bianchi, A. 2023. The predictive role of ERBB2 point mutations in metastatic colorectal cancer: A systematic review. *Cancer Treat Rev,* **112,** 102488.

Van Neerven, S. M., De Groot, N. E., Nijman, L. E., Scicluna, B. P., Van Driel, M. S., Lecca, M. C., Warmerdam, D. O., Kakkar, V., Moreno, L. F., Vieira Braga, F. A., Sanches, D. R., Ramesh, P., Ten Hoorn, S., Aelvoet, A. S., Van Boxel, M. F., Koens, L., Krawczyk, P. M., Koster, J., Dekker, E., Medema, J. P., Winton, D. J., Bijlsma, M. F., Morrissey, E., Léveillé, N. & Vermeulen, L. 2021a. Apc-mutant cells act as supercompetitors in intestinal tumour initiation. *Nature*.

Van Neerven, S. M., De Groot, N. E., Nijman, L. E., Scicluna, B. P., Van Driel, M. S., Lecca, M. C., Warmerdam, D. O., Kakkar, V., Moreno, L. F., Vieira Braga, F. A., Sanches, D. R., Ramesh, P., Ten Hoorn, S., Aelvoet, A. S., Van Boxel, M. F., Koens, L., Krawczyk, P. M., Koster, J., Dekker, E., Medema, J. P., Winton, D. J., Bijlsma, M. F., Morrissey, E., Léveillé, N. & Vermeulen, L. 2021b. Apc-mutant cells act as supercompetitors in intestinal tumour initiation. *Nature,* **594,** 436-441.

Vasquez, E. G., Nasreddin, N., Valbuena, G. N., Mulholland, E. J., Belnoue-Davis, H. L., Eggington, H. R., Schenck, R. O., Wouters, V. M., Wirapati, P., Gilroy, K., Lannagan, T. R. M., Flanagan, D. J., Najumudeen, A. K., Omwenga, S., Mccorry, A. M. B., Easton, A., Koelzer, V. H., East, J. E., Morton, D., Trusolino, L., Maughan, T., Campbell, A. D., Loughrey, M. B., Dunne, P. D., Tsantoulis, P., Huels, D. J., Tejpar, S., Sansom, O. J. & Leedham, S. J. 2022. Dynamic and adaptive cancer stem cell population admixture in colorectal neoplasia. *Cell Stem Cell,* **29,** 1213-1228.e8.

Velho, S., Oliveira, C., Ferreira, A., Ferreira, A. C., Suriano, G., Schwartz, S., Jr., Duval, A., Carneiro, F., Machado, J. C., Hamelin, R. & Seruca, R. 2005. The prevalence of PIK3CA mutations in gastric and colon cancer. *Eur J Cancer,* **41,** 1649-54.

Vilar, E. & Gruber, S. B. 2010. Microsatellite instability in colorectal cancer-the stable evidence. *Nat Rev Clin Oncol,* **7,** 153-62.

Viny, A. D., Bowman, R. L., Liu, Y., Lavallée, V. P., Eisman, S. E., Xiao, W., Durham, B. H., Navitski, A., Park, J., Braunstein, S., Alija, B., Karzai, A., Csete, I. S., Witkin, M., Azizi, E., Baslan, T., Ott, C. J., Pe'er, D., Dekker, J., Koche, R. & Levine, R. L. 2019. Cohesin Members Stag1 and Stag2 Display Distinct Roles in Chromatin Accessibility and Topological Control of HSC Self-Renewal and Differentiation. *Cell Stem Cell,* **25,** 682-696.e8.

Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura, Y., White, R., Smits, A. M. & Bos, J. L. 1988. Genetic alterations during colorectal-tumor development. *N Engl J Med,* **319,** 525-32.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr. & Kinzler, K. W. 2013. Cancer genome landscapes. *Science,* **339,** 1546-58.

Voorneveld, P. W., Kodach, L. L., Jacobs, R. J., Van Noesel, C. J., Peppelenbosch, M. P., Korkmaz, K. S., Molendijk, I., Dekker, E., Morreau, H., Van Pelt, G. W., Tollenaar, R. A., Mesker, W., Hawinkels, L. J., Paauwe, M., Verspaget, H. W., Geraets, D. T., Hommes, D. W., Offerhaus, G. J., Van Den Brink, G. R., Ten Dijke, P. & Hardwick, J. C. 2015. The BMP pathway either enhances or inhibits the Wnt pathway depending on the SMAD4 and p53 status in CRC. *Br J Cancer,* **112,** 122-30.

Vycital, O., Pitule, P., Hosek, P., Kriz, T., Treska, V. & Liska, V. 2019. Expression of Serpin B9 as a Prognostic Factor of Colorectal Cancer. *Anticancer Res,* **39,** 6063-6066.

Wang, H., He, L., Li, Y., Pu, W., Zhang, S., Han, X., Lui, K. O. & Zhou, B. 2022. Dual Cre and Dre recombinases mediate synchronized lineage tracing and cell subset ablation in vivo. *J Biol Chem,* **298,** 101965.

Wang, Z., Cummins, J. M., Shen, D., Cahill, D. P., Jallepalli, P. V., Wang, T. L., Parsons, D. W., Traverso, G., Awad, M., Silliman, N., Ptak, J., Szabo, S., Willson, J. K., Markowitz, S. D., Goldberg, M. L., Karess, R., Kinzler, K. W., Vogelstein, B., Velculescu, V. E. & Lengauer, C. 2004. Three classes of genes mutated in colorectal cancers with chromosomal instability. *Cancer Res,* **64,** 2998-3001.

Watson, P., Vasen, H. F. A., Mecklin, J. P., Bernstein, I., Aarnio, M., Järvinen, H. J., Myrhøj, T., Sunde, L., Wijnen, J. T. & Lynch, H. T. 2008. The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. *Int J Cancer,* **123,** 444-449.

Weisenberger, D. J., Siegmund, K. D., Campan, M., Young, J., Long, T. I., Faasse, M. A., Kang, G. H., Widschwendter, M., Weener, D., Buchanan, D., Koh, H., Simms, L., Barker, M., Leggett, B., Levine, J., Kim, M., French, A. J., Thibodeau, S. N., Jass, J., Haile, R. & Laird, P. W. 2006. CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet,* **38,** 787-93.

Welcker, M. & Clurman, B. E. 2008. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat Rev Cancer,* **8,** 83-93.

Welcker, M., Orian, A., Jin, J., Grim, J. E., Harper, J. W., Eisenman, R. N. & Clurman, B. E. 2004. The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc Natl Acad Sci U S A,* **101,** 9085-90.

Werner, B. & Sottoriva, A. 2018. Variation of mutational burden in healthy human tissues suggests non-random strand segregation and allows measuring somatic mutation rates. *PLoS Comput Biol,* **14,** e1006233.

Wright, W. E., Pereira-Smith, O. M. & Shay, J. W. 1989. Reversible cellular senescence: implications for immortalization of normal human diploid fibroblasts. *Mol Cell Biol,* **9,** 3088-92.

Wu, Y., Wang, Y., Yao, H., Li, H., Meng, F., Li, Q., Lin, X. & Liu, L. 2022. MNX1-AS1, a c-Myc induced lncRNA, promotes the Warburg effect by regulating PKM2 nuclear translocation. *J Exp Clin Cancer Res,* **41,** 337.

Yan, D., Wiesmann, M., Rohan, M., Chan, V., Jefferson, A. B., Guo, L., Sakamoto, D., Caothien, R. H., Fuller, J. H., Reinhard, C., Garcia, P. D., Randazzo, F. M., Escobedo, J., Fantl, W. J. & Williams, L. T. 2001. Elevated expression of axin2 and hnkd mRNA provides evidence that Wnt/beta -catenin signaling is activated in human colon tumors. *Proc Natl Acad Sci U S A,* **98,** 14973-8.

Yan, H. H. N., Lai, J. C. W., Ho, S. L., Leung, W. K., Law, W. L., Lee, J. F. Y., Chan, A. K. W., Tsui, W. Y., Chan, A. S. Y., Lee, B. C. H., Yue, S. S. K., Man, A. H. Y., Clevers, H., Yuen, S. T. & Leung, S. Y. 2017. RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. *Gut,* **66,** 1645-1656.

Yeh, C. H., Bellon, M. & Nicot, C. 2018. FBXW7: a critical tumor suppressor of human cancers. *Mol Cancer,* **17,** 115.

You, Z. & Bailis, J. M. 2010. DNA damage and decisions: CtIP coordinates DNA repair and cell cycle checkpoints. *Trends Cell Biol,* **20,** 402-9.

Yu, Z., Huang, H., Reim, A., Charles, P. D., Northage, A., Jackson, D., Parry, I. & Kessler, B. M. 2017. Optimizing 2D gas chromatography mass spectrometry for robust tissue, serum and urine metabolite profiling. *Talanta,* **165,** 685-691.

Yui, S., Azzolin, L., Maimets, M., Pedersen, M. T., Fordham, R. P., Hansen, S. L., Larsen, H. L., Guiu, J., Alves, M. R. P., Rundsten, C. F., Johansen, J. V., Li, Y., Madsen, C. D., Nakamura, T., Watanabe, M., Nielsen, O. H., Schweiger, P. J., Piccolo, S. & Jensen, K. B. 2018. YAP/TAZ-Dependent Reprogramming of Colonic Epithelium Links ECM Remodeling to Tissue Regeneration. *Cell Stem Cell,* **22,** 35-49.e7.

Yum, M. K., Han, S., Fink, J., Wu, S. S., Dabrowska, C., Trendafilova, T., Mustata, R., Chatzeli, L., Azzarelli, R., Pshenichnaya, I., Lee, E., England, F., Kim, J. K., Stange, D. E., Philpott, A., Lee, J. H., Koo, B. K. & Simons, B. D. 2021. Tracing oncogene-driven remodelling of the intestinal stem cell niche. *Nature*.

Yumimoto, K. & Nakayama, K. I. 2020. Recent insight into the role of FBXW7 as a tumor suppressor. *Semin Cancer Biol,* **67,** 1-15.

Zhang, Y., Liu, Z. & Zhang, B. 2022. Separate roles of LMAN1 and MCFD2 in ER-to-Golgi trafficking of FV and FVIII. *Blood Adv*.

Zhang, Y., Zhu, M., Zheng, C., Wei, W., Emmer, B. T. & Zhang, B. 2022. LMAN1-MCFD2 complex is a cargo receptor for the ER-Golgi transport of α1-antitrypsin. *Biochem J,* **479,** 839-855.

Zhou, C., Sun, H., Zheng, C., Gao, J., Fu, Q., Hu, N., Shao, X., Zhou, Y., Xiong, J., Nie, K., Zhou, H., Shen, L., Fang, H. & Lyu, J. 2018. Oncogenic HSP60 regulates mitochondrial oxidative phosphorylation to support Erk1/2 activation during pancreatic cancer cell growth. *Cell Death Dis,* **9,** 161.

Zhou, C. Z., Peng, Z. H., Zhang, F., Qiu, G. Q. & He, L. 2002. Loss of heterozygosity on long arm of chromosome 22 in sporadic colorectal carcinoma. *World J Gastroenterol,* **8,** 668-73.

Zhou, H., Neelakantan, D. & Ford, H. L. 2017. Clonal cooperativity in heterogenous cancers. *Semin Cell Dev Biol,* **64,** 79-89.

Zhu, G., Pei, L., Xia, H., Tang, Q. & Bi, F. 2021. Role of oncogenic KRAS in the prognosis, diagnosis and treatment of colorectal cancer. *Mol Cancer,* **20,** 143.