Somatic Mutation in Colorectal Cancer: A Window to Disease Pathogenesis



Abigail D'Cruz

Keble College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2020

Preface

This thesis is a collection of research work carried out by the author in the Wellcome Centre for Human Genetics, Nuffield Department of Medicine, University of Oxford, between October 2015 and January 2020, under the supervision of Professor Simon Leedham. The research described herein is original, though the work of others is cited with appropriate references within the text. No part of this research has previously been submitted for a degree at this or any other university.

Abstract

Cancer has long been considered to be a disease driven by mutation. In recent years, this theory has been updated to reflect increased understanding of the importance of additional factors influencing disease pathogenesis, such as the cell of origin and microenvironment. In this thesis, I explore somatic mutation acquisition and the effect of microenvironmental influence on human and animal models of colorectal tumourigenesis.

Analysis of mutations in healthy human colonic epithelium revealed a capacity to tolerate driver mutations, while inflamed epithelium exhibited increased mutation acquisition. I also investigated somatic mutation following exposure to the mutagen AOM in mice harbouring spontaneous colitis or a constitutively deranged intestinal microenvironment. Here, I show that these differences in underlying disease state influence mutation acquisition and provide novel evidence of an AOM-associated mutational signature in mice.

Although numerous studies have suggested links between morphogenetic signalling pathways and the DNA damage response (DDR), the influence of these pathways upon mutation protection mechanisms within the intestine remains unclear. To explore this, I studied the *Vil1-Grem1* (VG) mouse, a model of disrupted intestinal homeostasis in which aberrant epithelial expression of the BMP signalling antagonist Grem1 results in the formation of ectopic crypts. I provide evidence of impaired activation of DDR pathways within ectopic crypts following exposure to ionising radiation and go on to propose that a novel mechanism linking BMP signalling and DNA repair results in increased genomic instability within this compartment.

I also examined the effect of therapeutic restoration of cell fate determination on mutation acquisition following AOM exposure in VG mice. I demonstrate that therapeutic restoration of morphogen gradients in VG mice can rescue disease phenotype and confer protection against mutation acquisition, but may also result in a putative skew to DNA repair-associated mutation acquisition in these mice.

Finally, I provide evidence that disease progression in an immune-deficient model of inflammation-associated cancer occurs in the absence of mutation accumulation and is instead associated with dysregulated gene expression. Importantly, I demonstrate that abrogation of the inflammatory drive within this model rescues disease phenotype in colitis but not in polyps.

Overall, I have demonstrated that somatic mutation acquisition within the intestine varies and reflects the influence of the microenvironment, cell of origin and mutation protection mechanisms upon disease pathogenesis. This has translational relevance, given that the development of personalised therapies requires a holistic and accurate understanding of the mechanisms underpinning the vast heterogeneity of colorectal cancer.

Acknowledgements

To be given the time, space and support to independently explore novel ideas is the great privilege of academia and one that, for me, has been consistently underpinned by collaboration (and an inevitable smattering of existential crises). I therefore feel compelled to acknowledge not only the people who have contributed directly to the research presented here, but also those who have imbued my time in Oxford with true joy and so much laughter.

First and foremost, I want to thank my supervisor, Professor Simon Leedham. To be supervised by Simon is to experience first-hand the power of combining meaningful scientific enquiry with a patient-centric approach and a truly collaborative spirit. I am extremely grateful for his patience, support and insight over the last four years.

The members of the Leedham lab consistently made cycling up the hill worth it. Hayley Davis was instrumental to all of the mouse work within this project and is an unendingly patient, supportive and encouraging presence within the lab; I'm so grateful for all of her help and advice over the last four years. Ryan Schenck carried out all post-processing for the sequencing data utilised in this thesis; I am very fortunate to have had both his sense of humour and bioinformatic skills to drawn upon throughout this project. I'd also like to thank Gabriel, Eoghan, Martijn and Sujata for their sage advice and being great sounding boards, and Ester, Sulo and Nadia for brightening up the lab with laughter (and questionable bench hygiene).

The work presented here would not have been possible without the help of several collaborators and colleagues, both within the Wellcome Centre for Human Genetics and beyond. Within the Powrie lab at the Kennedy Institute of Rheumatology, the inimitable Dr Fanny Franchini and Dr Sarah McCuaig coordinated the IL-22 blockade experiment, while Dr Matt Jackson processed the raw RNAseq files to generate count data. A warm thank you also to Alina Janney, who contributed to my transcriptomic approach and spent many hours with me trying to figure out R pipelines over too much coffee. I'm also very grateful to Dr Viktor Koelzer for his time and expertise in introducing me to digital pathology and image analysis using the HALO platform. MIP sequencing was carried out by Laura Chegwidden and Dr Claire Palles within the Institute of Cancer and Genomic Sciences at the University of Birmingham. Dr Lai Mun Wang conducted all blinded pathology scoring with the utmost professionalism and reliability. Raquel Cuella-Martin and Cathy Oliviera helped me navigate the incredibly complex field of

DNA damage research and could always be relied upon for a smile and a chat when it all seemed too much. I must also note my sincere appreciation for the work of the research nurses within the endoscopy unit of the John Radcliffe Hospital, whose patience and efficiency enabled the successful collection of intestinal biopsies for this project.

Throughout my studies, I have been supported by funding from the University of Oxford Clarendon Fund, the Nuffield Department of Medicine Prize Studentship and the Keble College de Breyne and Ian Tucker scholarships. I am extremely grateful to these organisations for facilitating my time in Oxford, and would like to make particular mention here of my college, Keble, for providing an unfailingly welcoming environment and supporting me at every opportunity.

In the interests of word count, I will not list all of the social and sporting endeavours that have contributed so much to my time here, but anyone who knows me will understand why I feel compelled to single out the people and institutions of Cowley, coffee and rugby. The first time that Cowley Road felt like home took me by surprise, but it is now a feeling that is as natural and comforting to me as any other. I have lived with some of my best friends in Cowley and met many others while working at one of its institutions, Mostro Coffee. I also couldn't write this section without mentioning the Oxford University Rugby Football Club. Captaining OURFC in 2018 will always be one of the most significant experiences of my life, which may seem peculiar to those who view it merely as a university sports club but will make immediate sense to those fortunate enough to be a part of its storied history.

At the end of it all, though, this entire experience will always be defined by the people that have not only put up with me but managed to fill the last four years with warmth and sunshine (often in lieu of the English variety). The Oxonians: Arran, Bogdan, Kiah, Con, Jake, Bev, Claudia, Ro, Isa and Sydney; the Londoners: Holly, Helen, Shazia, Soph and Geoff; the Aussies back home: Ian, Pauly, Richard, Tij, Adelaide, Kate, Ciara, Hannah and Vlada; and the International Support Station: Ahmed, Palmo, Lewe, Ava, Chris and Ronald. They have believed in this even, and especially, when I didn't; I am extraordinarily grateful for their presence through all of it.

Last, but never least, my parents. Letty and William escaped poverty in Mumbai to build a life in which my sister and I were afforded every opportunity to succeed. Despite being separated by vast oceans and multiple time zones, I have never felt far from their unwavering support. This thesis is for them.

NB: The work presented here begins and ends with words from Tim Winton in his novel "Cloudstreet". My high school English teacher, Mr Dave Sharp, gave me a copy when I left Perth at 17 for university and it remains one of my most treasured possessions. It is because Dave encouraged me to not only challenge myself but every expectation I encountered that I have the opportunity to present this thesis - it is fitting to wrap it in a reminder of the value of great teachers and the endless possibilities they convey.

"Will you look at us by the river! The whole restless mob of us on spread blankets in the dreamy briny sunshine skylarking and chiacking about for one day, one clear, clean, sweet day in a good world in the midst of our living."

- Tim Winton, *Cloudstreet*

Contents

List of Tables **xii**

List of Figures **xiii**

List of Abbreviations **xix**

1 Introduction **1**

1.1 Intestinal architecture 1
1.1.1 The crypt and villus 1
1.1.2 Intestinal stem cells 2
1.2 Signalling pathways in intestinal homeostasis 4
1.2.1 WNT 4
1.2.2 Bone Morphogenetic Protein 6
1.2.3 TGF-β 7
1.2.4 Notch 8
1.2.5 Interleukins 9
1.2.6 Disruption of the intestinal microenvironment 10
1.2.6.1 Inflammation: IL-22 10
1.2.6.2 Disrupted morphogen gradients: HMPS 11
1.2.7 Mouse models of intestinal disease 12
1.3 Homeostatic protection against mutation 13
1.3.1 Spatial organisation and cell turnover 13
1.3.2 Neutral drift 14
1.4 Mutation control: The DNA damage response 15
1.4.1 Excision repair 16
1.4.2 Double strand break repair 17
1.4.2.1 Homologous recombination 18
1.4.2.2 Non-homologous end joining 18
1.4.2.3 Alternative end joining 19
1.4.3 DDR pathway choice 20
1.4.4 DDR in the gut 22
1.4.4.1 DDR along the crypt-villus axis 22

1.4.4.2 Irradiation-induced stem cell plasticity 23
1.4.4.3 The epithelial response to irradiation 25
1.4.5 Signalling pathways and DDR 26
1.5 Colorectal Cancer 29
1.5.1 The adenoma-carcinoma sequence 29
1.5.2 Added complexity: cell of origin and the microenvironment 30
1.5.3 Polyp pathways 31
1.5.4 Molecular characterisation of human CRC 33
1.5.5 The genotype-phenotype axis in intestinal disease 33
1.6 Outline of thesis 35

2 Materials and Methods **38**
2.1 Human Samples 39
2.1.1 Consent and ethics for human samples 39
2.1.2 Single and pooled crypt dissection of fresh biopsies 39
2.2 Mouse procedures 40
2.2.1 Mouse strains 40
2.2.2 Ear snips 41
2.2.3 Breeding 41
2.2.4 Administration of substances 41
2.2.4.1 Azoxymethane 41
2.2.4.2 Dextran Sodium Sulfate 41
2.2.4.3 *Helicobacter hepaticus* 42
2.2.4.4 *In vivo* antibody treatment 42
2.2.5 Irradiation 42
2.2.6 Gut preparation 42
2.2.6.1 Collection of small intestine crypt and villi 43
2.2.6.2 Collection of colon samples 43
2.3 Histology 43
2.3.1 Adenoma scoring 43
2.3.2 Processing 44
2.3.3 Embedding 44
2.3.4 Sectioning and staining 44
2.3.4.1 Haematoxylin and Eosin (H and E) staining . . . 44
2.3.5 Pathology scoring 45
2.4 DNA 45
2.4.1 DNA extraction 45
2.4.1.1 DNA extraction from tissue 45
2.4.1.2 DNA extraction from blood 46

2.4.2 Nucleic acid quality control 46
2.4.2.1 Nanodrop spectrophotometer 46
2.4.2.2 Qubit fluorometer 46
2.4.3 DNA sequencing 47
2.4.3.1 Molecular Inversion Probe Sequencing 47
2.4.3.2 Whole Exome Sequencing 51
2.4.3.3 Pre-processing 51
2.4.3.4 Merging multiplex sequences 51
2.4.3.5 Single Nucleotide Variant and Indel calling . . . 52
2.4.3.6 SNV/Indel post-processing 53
2.4.3.7 Driver mutation identification 54
2.4.3.8 Mutational signature analysis 54
2.5 RNA / cDNA 55
2.5.1 RNA extraction 55
2.5.2 RNA sequencing 56
2.5.3 Reverse transcription of RNA into cDNA 56
2.5.3.1 Pre-amplification of cDNA 56
2.5.4 Quantitative real-time PCR 57
2.5.4.1 Fluidigm 57
2.5.4.2 TaqMan micro-RNA assay 57
2.5.5 In-situ micro-RNA hybridisation 58
2.6 Protein 59
2.6.1 Immunohistochemistry 59
2.6.2 Digital pathology 60
2.6.2.1 HALO Image Analysis 60
2.6.2.2 Aperio ImageScope 61
2.6.2.3 QuPath 61

3 Underlying disease state impacts upon mutation acquisition 63
3.1 Introduction 64
3.1.1 Chapter aims 65
3.2 Evidence of selective pressures in normal and inflamed human colon 65
3.2.1 Mutation burden analysis 68
3.2.2 Characterisation of mutation types 69
3.2.3 Mutational signature analysis 72
3.2.4 Characterisation of driver gene mutations 75
3.2.5 Summary of findings 79
3.3 Evidence of selective pressures in mouse colon from varying disease states following uniform exposure to a mutagen 81

3.3.1 Survival analysis . . . 82
3.3.2 Characterisation of intestinal phenotype . . . 83
3.3.2.1 Macroscopic appearance . . . 83
3.3.2.2 Immunohistochemical analysis . . . 86
3.3.3 Genomic analysis . . . 89
3.3.3.1 Analysis of mutation burden . . . 90
3.3.3.2 Characterisation of mutation types . . . 92
3.3.3.3 Mutational signature analysis . . . 95
3.3.3.4 Characterisation of driver gene mutations . . . 101
3.3.4 Summary of findings . . . 105
3.4 Discussion . . . 106
3.4.1 Evidence of selective pressures in human colonic epithelium 106
3.4.2 Evidence of selective pressures in mouse colonic epithelium 109
3.4.3 Conclusion . . . 111

4 Cell intrinsic mutation susceptibility in *Vil1-Grem1* mice 113
4.1 Introduction . . . 114
4.2 Results . . . 116
4.2.1 IHC analysis of DDR pathway markers within the ectopic crypts of *Vil1-Grem1* mice . . . 116
4.2.1.1 DNA damage sensing within ectopic crypts . . . 117
4.2.1.2 DNA damage translation within ectopic crypts . 118
4.2.1.3 Activation of apoptotis within ectopic crypts . . . 120
4.2.2 Increased miR-96 expression in the SB1 villus region of VG mice . . . 124
4.2.3 Summary of findings . . . 125
4.3 Discussion and Future work . . . 128
4.3.1 Discussion . . . 128
4.3.1.1 Out with the old axis, in with the new . . . 130
4.3.1.2 Limitations . . . 131
4.3.1.3 Future Work . . . 131

5 The impact of therapy on mutation acquisition in *Vil1-Grem1* mice 133
5.1 Introduction . . . 134
5.2 Results: Anti-Grem1 in *Vil1-Grem1* mice . . . 136
5.2.1 Survival analysis . . . 136
5.2.2 Phenotypic characterisation . . . 137
5.2.3 Genomic analysis . . . 144
5.2.3.1 Mutation burden analysis . . . 144
5.2.3.2 Characterisation of mutation types . . . 145

5.2.3.3 Mutational signature analysis 147
5.2.3.4 Characterisation of driver gene mutations 152
5.2.4 Summary: Anti-Grem1 therapy in *Vil1-Grem1* mice 159
5.3 Discussion 162

6 The impact of therapy on mutation acquisition in 129SvEv.*Rag*$^{-/-}$ mice 166
6.1 Introduction 167
6.2 Results 168
6.2.1 Phenotypic characterisation 168
6.2.2 Genomic analysis 171
6.2.2.1 Mutation burden analysis of anti-IL-22-treated normal and polyp tissue 172
6.2.2.2 Characterisation of mutation types within anti-IL-22-treated normal and polyp tissue 175
6.2.2.3 Mutation signature analysis of anti-IL-22-treated normal and polyp tissue 177
6.2.3 Transcriptomic analysis 184
6.2.3.1 Characterisation of disease progression 184
6.2.3.2 Intestinal cell populations and signalling pathways 193
6.2.3.3 The effect of anti-IL-22 therapy on disease progression 199
6.2.3.4 Intestinal cell populations and signalling pathways in treated samples 206
6.2.4 Summary of findings 213
6.3 Discussion 214

7 Discussion 219
7.1 Summary of key findings 219
7.2 Discussion 220
7.2.1 Updating the genotype-phenotype model: persistence is key 220
7.2.2 The influence of persistence on disease progression 222
7.2.3 Exploiting mutation persistence for novel therapeutic strategies 223
7.2.4 Conclusion 224

Appendices

References 230

List of Tables

2.1 Genes targeted by MIP sequencing primers 50
2.2 'Variant Filter' program arguments 54
2.3 'Common Drivers in CRC' gene list 55
2.4 TaqMan probes 58
2.5 Primary antibodies used for immunohistochemistry 60
2.6 QuPath 'positive cell detection' algorithm settings 62

3.1 Patient information 66

4.1 Analysed proteins corresponding to the three stages of DDR . . . 117

5.1 Missense *Ctnnb1* mutations detected within *Vil1-Grem1* polyps . . 155
5.2 Missense mutations in genes involved in DNA repair 156

1 Gene sets utilised for gene set enrichment analysis 229

List of Figures

1.1 Architecture of the intestinal epithelium 3
1.2 Gradients of key signalling pathways within intestinal crypts . . . 6
1.3 *Vil1-Grem1* mouse phenotype 12
1.4 DNA damage repair pathways 21
1.5 DDR along the crypt-villus axis 23
1.6 The influence of TGF-β and BMP signalling upon DDR outcomes 28
1.7 The adenoma-carcinoma sequence 30
1.8 Model depicting 'conventional' and 'non-conventional' polyp pathways 32
1.9 Mutant clones in an average 1 cm^2 of normal eyelid skin 34
1.10 The genotype-phenotype axis in intestinal disease 35

3.1 Colon biopsy sites for each patient group 67
3.2 Total mutation burden per sample according to patient group . . 68
3.3 Characterisation of non-synonymous mutation burden in samples from all patient groups 70
3.4 Characterisation of mutation types in samples from each patient group. 71
3.5 Characterisation of mutational spectrum of samples within each patient group. 72
3.6 Mutational process contribution in samples from each patient group. 74
3.7 Distribution of non-synonymous driver gene mutations in samples from all patient groups. 76
3.8 Visualisation of the read depth and variant allele frequency of missense mutations. 77
3.9 Characterisation of missense variants within samples from each patient group. 78
3.10 Characterisation of genes harbouring the most missense mutations per patient group. 79
3.11 Comparison of missense mutations detected in healthy and inflamed samples. 80

3.12 Treatment scheme for control and experimental groups of wild-type and *Vil1-Grem1* mice. 82
3.13 Kaplan-Meier survival analysis of WT and VG mice exposed to AOM and DSS. 84
3.14 Macroscopic intestinal phenotype of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 85
3.15 Characterisation of colon histology in wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 88
3.16 Correlation analysis of colon polyp size and survival of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 89
3.17 Non-synonymous mutation burden of polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 91
3.18 Correlation analysis of colon polyp size and non-synonymous mutation burden of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 92
3.19 Correlation analysis of non-synonymous mutation burden and age of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. . . . 93
3.20 Characterisation of mutation types present in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 94
3.21 Characterisation of mutation types present in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 96
3.22 96-substitution classification of mutations in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 97
3.23 96-substitution classification of mutations in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 98
3.24 Analysis of mutational signature contribution in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. . 99
3.25 Analysis of mutational signature contribution in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 100
3.26 Distribution of driver gene mutations in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 102
3.27 Distribution of driver gene mutations in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 103
3.28 Representative images of IHC β-catenin staining in WT DSS AOM and VG AOM colon polyps. 103
3.29 Recurrently mutated driver genes within polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. 104
3.30 Comparison of missense mutations detected in polyps 105

4.1 Analysis of γH2AX expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 118
4.2 Analysis of pATM expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 119
4.3 Analysis of phospho-p53 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 120
4.4 Analysis of RAD51 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 121
4.5 Analysis of cleaved caspase-3 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 122
4.6 Analysis of BAG1 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. 123
4.7 Expression analysis of homeostatic miR-96 within WT and VG mice 126
4.8 Expression analysis of miR-96 in response to 10 Gy IR within WT and VG mice 127

5.1 Abrogation of *Vil1-Grem1* intestinal phenotype following anti-Grem1 antibody treatment 135
5.2 A model of therapy-driven protection against mutagenesis. 135
5.3 Treatment scheme for *Vil1-Grem1* mice treated with an anti-Grem1 antibody and AOM. 137
5.4 Kaplan-Meier survival plot (top) and analysis (bottom) for *Vil1-Grem1* mice treated with the mutagen AOM and an anti-Grem1 antibody. 138
5.5 Macroscopic intestinal phenotype of *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 139
5.6 Characterisation of small intestine phenotype of *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. 141
5.7 Characterisation of colon phenotype in *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 142
5.8 Correlation analysis of colon polyp size and mouse survival. . . . 143
5.9 Non-synonymous mutation burden of polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. 146
5.10 Correlation analysis of experimental polyps according to non-synonymous mutation burden and mouse survival. 147
5.11 Correlation analysis of polyp size and non-synonymous mutation burden. 148
5.12 Characterisation of mutation types present in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 149

5.13 Characterisation of mutational spectrum in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 150
5.14 Mutational process contribution in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 151
5.15 Distribution of driver gene mutations in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 152
5.16 Number of driver gene mutations within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. 153
5.17 Recurrently mutated driver genes within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. 154
5.18 Depth and variant allele frequency of missense mutations in DNA repair-associated genes 157
5.19 Average relative mutation signature contribution within polyps from each experimental model according to DNA repair mutations 158
5.20 Correlation analysis of polyp non-synonymous mutation burden and number of missense DNA repair-associated mutations. 159
5.21 Correlation analysis of polyp driver gene mutation burden and number of missense DNA repair mutations. 160
5.22 Average non-synonymous mutation burden within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 according to number of missense DNA repair mutations. 161
5.23 A model of the impact of anti-Grem1 therapy in *Vil1-Grem1* mice treated with AOM 164

6.1 Treatment scheme for experimental 129SvEv.*Rag*$^{-/-}$ mice 169
6.2 Characterisation of colon phenotype in experimental 129SvEv.*Rag*$^{-/-}$ mice 170
6.3 Characterisation of proliferative and progenitor cell number and apoptosis in 129SvEv.*Rag*$^{-/-}$+*Hh*+AOM mice treated with anti-IL-22 or isotype-mAb. 171
6.4 Representative image of an *Hh*+AOM mouse colonic polyp 172
6.5 Non-synonymous mutation burden in adjacent normal colon samples 173
6.6 Non-synonymous mutation burden in polyp samples. 174
6.7 Analysis of driver gene mutations according to non-synonymous mutation burden 175
6.8 Average non-synonymous mutation burden in normal adjacent samples and polyps from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice 176
6.9 Percentage contribution of mutation types in normal adjacent samples 177
6.10 Percentage contribution of mutation types in polyps 178

6.11 Average mutation type contribution in adjacent normal and polyp samples from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice 179
6.12 96-class signature analysis of mutations in normal samples 180
6.13 96-class signature analysis of mutations in polyps 180
6.14 Relative contribution of mutational signatures in normal samples 181
6.15 Relative contribution of mutational signatures in polyp samples . 183
6.16 Principal component analysis plot showing clustering of steady state, colitis and colitic polyp samples. 185
6.17 Volcano plot of genes differentially expressed in colitis samples compared to steady state samples 186
6.18 Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in colitis samples 187
6.19 Volcano plot of genes differentially expressed in colitic polyps compared to colitis samples 189
6.20 Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in colitic polyps 190
6.21 Gene set enrichment analysis plots showing enriched processes within colitis and colitic polyps 192
6.22 Normalized expression of genes involved in Il-22 signalling within steady state, colitis and colitic polyp samples. 194
6.23 Normalized expression of intestinal stem and proliferating cell markers within steady state, colitis and colitic polyp samples. . . 195
6.24 Normalized expression of markers of intestinal epithelial cell types within steady state, colitis and colitic polyp samples 196
6.25 Normalized expression of genes involved in Wnt signalling within steady state, colitis and colitic polyp samples 198
6.26 Normalized expression of genes involved in Notch signalling within steady state, colitis and colitic polyp samples 199
6.27 Principal component analysis plot showing clustering of steady state, colitis, colitic polyp, treated colitis and treated polyp samples. 200
6.28 Volcano plot of genes differentially expressed in anti-IL-22-treated colitis compared to untreated colitis 201
6.29 Volcano plot of genes differentially expressed in anti-IL-22-treated polyps compared to untreated polyps 202
6.30 Characterisation of the transcriptomic effect of anti-IL-22 therapy in colitis and colitic polyp samples. 203
6.31 Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in treated colitis 204
6.32 Heatmap of the top 100 pathways enriched in the lists of genes differentially expressed in anti-IL-22-treated polyps 205

6.33 GSEA of differentially expressed genes in anti-IL-22 treated colitis 206
6.34 Normalized expression of genes involved in Il-22 signalling within treated and untreated samples 208
6.35 Normalized expression of intestinal stem and proliferating cell markers within treated and untreated samples 209
6.36 Normalized expression of markers of intestinal epithelial cell types within treated and untreated samples 210
6.37 Normalized expression of genes involved in Wnt signalling within treated and untreated samples 211
6.38 Normalized expression of genes involved in Notch signalling within treated and untreated samples 212

7.1 Model depicting the factors that influence the persistence of mutated cells 221
2 Bar plot showing MIP sequencing read depth coverage per sample 227
3 Average small bowel (SB) and villus length of mice within each experimental group. 227
4 Characterisation of small bowel polyps in *Vil1-Grem1* mice 228

List of Abbreviations

ABC	Avidin/Biotin Containing
ADP	Adenosine Diphosphate
AHR	Aryl Hydrocarbon Receptor
ANNOVAR	ANNOtate VARiation tool
AOM	Azoxymethane
APC	Adenomatous Polyposis Coli
APOBEC	Apolipoprotein B mRNA Editing Catalytic Polypeptide-like
ATM	Ataxia-Telangiectasia Mutated
ATP	Adenosine Triphosphate
BAM	Binary Alignment Map
BER	Base Excision Repair
BMP	Bone Morphogenetic Protein
BSA	Bovine Serum Albumin
BWA	Burrows-Wheeler Alignment
CAC	Colitis-Associated Cancer
CFU	Colony Forming Units
COAD	Colon Adenocarcinoma
COREAD	Colorectal Adenocarcinoma
COSMIC	Catalogue Of Somatic Mutations In Cancer
CRC	Colorectal Cancer
DAB	3,3-Diaminobenzidine
DDR	DNA Damage Response
DE	Differentially Expressed
DEG	Differentially Expressed Gene
DFS	Disease-Free Survival
DIG	Digoxigenin
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
DSB	Double Strand Break

DSS	Dextran Sodium Sulfate
DTT	Dithiothreitol
ECF	Ectopic Crypt Foci
EDTA	Ethylenediaminetetraacetic Acid
FAM	Fluorescein Amidite
FAP	Familial Adenomatous Polyposis
FFPE	Formalin-Fixed Paraffin Embedded
GAPDH	Glyceraldehyde 3-Phosphate Dehydrogenase
GATK	Genomic Analysis Tool Kit
GEPIA	Gene Expression Profiling Interactive Analysis
GI	Gastrointestinal
GO	Gene Ontology
Hh	Helicobacter Hepaticus
HMPS	Hereditary Mixed Polyposis Syndrome
HNSCC	Head and Neck Squamous Cell Carcinoma
HR	Homologous Recombination
IBD	Inflammatory Bowel Disease
ICD	Intra-Cellular Domain
ICGC	International Cancer Genome Consortium
IFC	Integrated Fluidic Circuit
IHC	Immunohistochemistry
IL	Interleukin
IR	Ionising Radiation
ISC	Intestinal Stem Cell
ISH	In-Situ Hybridisation
JAK	Janus Kinase
KEGG	Kyoto Encyclopedia of Genes and Genomes
LFC	Log-Fold Change
LNA	Locked Nucleic Acid
MAPK	Mitogen-Activated Protein Kinase
MIP	Molecular Inversion Probe
MMEJ	Microhomology-Mediated End Joining
MMR	Mismatch Repair
MSI	Microsatellite Instability
MSS	Microsatellite Stable

NBF	Neutral Buffered Formalin
NER	Nucleotide Excision Repair
NHEJ	Non-Homologous End Joining
NHS	National Health Service
PARP	Poly-ADP Ribose Polymerase
PBS	Phosphate-Buffered Saline
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
PN	Product Number
PON	Panel Of Normals
PSC	Primary Sclerosing Cholangitis
PTEN	Phosphatase and Tensin Homolog
PUMA	p53-Upregulated Mediator of Apoptosis
RNA	Ribonucleic Acid
RPA	Replication Protein A
SAM	Sequence Alignment Map
SB	Small Bowel
SBS	Single Base Substution signature
SNV	Single Nucleotide Variant
SSA	Single Strand Annealing
SSC	Saline-Sodium Citrate
STAT	Signal Transducer and Activator of Transcription
TA	Transit Amplifying
TBE	Tris/Borate/EDTA
TCGA	The Cancer Genome Atlas
TCR	T-cell Receptor
TDRG	Tag-Defined Read Group
TGF	Transforming Growth Factor
UC	Ulcerative Colitis
UTR	Untranslated Region
UV	Ultra Violet
VCF	Variant Call File
VG	*Vil1-Grem1*
WES	Whole Exome Sequencing
WT	Wild-type

1 Introduction

1.1 Intestinal architecture

The human intestinal tract is a 4-6 m long tube comprised of an outer layer of smooth muscle lined with stroma and a sheet of epithelial cells. Rhythmic contractions of the smooth muscle layer result in the passage of food through the gut lumen. The epithelium provides dual absorptive and protective functions, mediating nutrient uptake and acting as a barrier against gut microbiota and dietary toxins. The majority of food digestion occurs within the small intestine, prior to passage to the large intestine for water removal before excretion from the rectum. In order to increase absorptive surface area, the epithelium of the small intestine consists of invaginations and protrusions referred to respectively as crypts of Lieberkuhn and villi. Villus length decreases along the intestinal tract so that the luminal surface of the large intestine and rectum is comprised of a flat surface epithelium harbouring only crypts.

1.1.1 The crypt and villus

Crypts and villi are the basic functional units of the intestine and their stereotypical architecture is maintained by constant, tightly regulated turnover of epithelial cells. These cells can be functionally divided into stem, transit amplifying (TA)

and differentiating cell compartments. The formation of these compartments is driven by an escalator of dividing and differentiating cells along the vertical crypt-villus axis. At the crypt base, a pool of 5-16 intestinal stem cells (ISCs) [1] gives rise to a population of 120-150 transit amplifying (TA) cells [2]. TA cells undergo a finite number of divisions before fully differentiating and are thus named for the constant upward movement of epithelial cells that arises as a result of this proliferative drive within crypts. As they move up the crypt-villus axis, cells differentiate and give rise to absorptive enterocytes and M cells, mucous-secreting goblet cells, hormone-secreting enteroendocrine cells and 'sentinel' tuft cells involved in immune-regulation [3]. Contrastingly, Paneth cells within the small intestine escape this upward movement and instead remain at the crypt base where they secrete anti-microbial proteins and are involved in maintenance of the crypt base stem cell niche [4, 5]. The cellular escalator culminates at the tip of the villus, where cells undergo apoptosis and are shed into the gut lumen. Homeostasis within the intestine therefore involves careful control of cell fate determination with an exquisite balance of proliferation and differentiation, resulting in rapid renewal of the entire epithelium every three days in mice and five days in humans [6].

1.1.2 Intestinal stem cells

Stem cells can be functionally defined as self-replicating and multipotent cells that are capable of long-term self-renewal and the generation of multiple cell lineages. Within the intestine, the 'Unitarian Theory of the Origin of the Epithelial Cells of the Gastrointestinal Tract' was proposed by Cheng and Leblond in 1974 [7]. This seminal work utilised phagosomes as markers to follow the evolution of crypt base columnar cells and demonstrate that all four (now six) differentiated intestinal cell lineages originated from the same precursor. The multipotency of cells at the intestinal crypt base was thus established. 14 years later, the use of chemical mutagenesis enabled visualisation of a flow of cells originating from the crypt base to the villus [8] and labelling of long-lived progenitor cells capable of giving

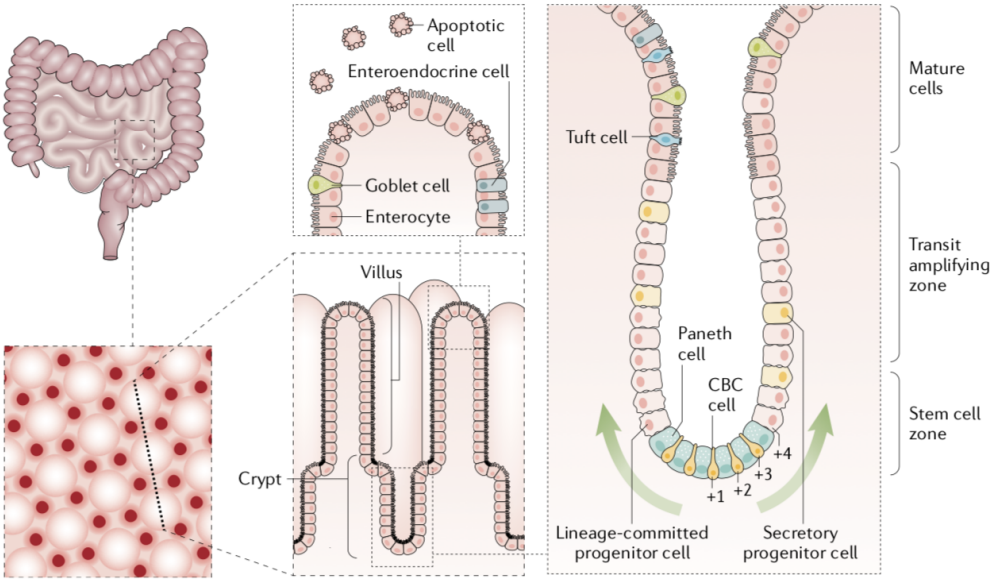


Figure 1.1: Architecture of the intestinal epithelium. The small intestine is organised into crypts and villi, with one villus surrounded by multiple crypts. Illustration of an intestinal cross-section reveals the structure of the crypt-villus unit, with stem cells at the crypt base giving rise to rapidly dividing and differentiating progenitor cells within the transit amplifying zone. Fully mature epithelial cells emerge from crypts and travel up towards the villus tip, where the cellular escalator culminates in the shedding of apoptosed cells into the gut lumen. Adapted from [1].

rise to all epithelial cell types [9]. In 2007, seminal work by Barker *et al* provided definitive proof of 'stemness' of crypt base columnar cells and established *Lgr5* as a highly specific marker for these cells. *Lgr5*$^{EGFP\text{-}IRES\text{-}CreERT2}$ mice were generated and crossed with *R26R-lacZ* reporter mice to generate offspring in which the administration of tamixofen results in activation of the Cre-ERT2 fusion enzyme. Subsequent continuous lacZ expression is thus detectable in Lgr5-expressing cells and their progeny. Cre-activation resulted in ribbons of lacZ-expressing cells that contained cells of all lineages, extended from crypts to villus tips within 5 days and exhibited lifelong persistence [10]. However, in addition to cells at the crypt base, those occupying the '+4' position within the crypt (situated immediately above the uppermost Paneth cell between the stem and TA cell zones) have long been described as additional ISC candidates due to their label-retaining capacity [11, 12]. Several studies have since identified *Lrig1* [13], *Hoxp1* [14], *Bmi1* [15] and *mTert* [16] as markers of additional stem cell populations within crypts that are capable

of reconstituting the epithelium following injury. Populations of stem cells at the +4 position were thus considered a quiescent, reserve stem cell pool capable of replenishing the crypt base ISC pool if required. However, elegant lineage tracing experiments have also identified populations of long-lived intestinal cells that can regain 'stemness' and regenerate damaged intestine. These 'early progenitor' cells express secretory markers and fall back into the stem cell niche following damage, where they potentially de-differentiate in order to develop full stem cell potential [4, 17]. Similarly, *Dll1*+ cells within the crypt are secretory progenitors that do not normally act as stem cells but are able to repopulate the crypt following intestinal damage [18]. These studies have raised tantalising questions regarding the capacity of the stem cell niche to reprogram fully differentiated cells and the existence of a 'point of no return' in terms of epithelial cell differentiation. It is clear that there is a high degree of plasticity within the intestine (expanded upon later in this chapter), with crypts harbouring multiple populations of ISCs and non-stem cells able to de-differentiate and regain 'stemness' to repopulate crypts. We have thus progressed from the use of highly-specific markers to define and interrogate intestinal stem cell populations within a unidirectional model of cell differentiation, to a broader understanding of context-dependent epithelial plasticity underpinned by dynamic, functionally-defined populations of stem cells.

1.2 Signalling pathways in intestinal homeostasis

The complex, cross-talking cell populations present within the intestine are governed by tightly controlled gradients of morphogenetic signalling pathways (Figure 1.2).

1.2.1 WNT

A wide range of evidence indicates that the WNT signalling pathway plays a crucial role in intestinal proliferation and the maintenance of ISCs. WNT signalling is based on the interaction of WNT glycoproteins with their transmembrane receptors. Non-canonical WNT signalling is induced by non-classical WNT

ligands and is defined by β-catenin-independent mechanisms of WNT signal transduction. Canonical WNT signalling, induced by classical WNT ligands, involves the stabilization and nuclear translocation of β-catenin and is the most relevant to stem cell signalling. In the absence of WNT activation, membranous β-catenin is involved in cell adhesion while cytosolic β-catenin is efficiently captured by a destruction complex consisting of the scaffold protein AXIN, adenomatous polyposis coli (APC) and glycogen synthase kinase-3β (GSK-3β). Phosphorylation of the N-terminus of β-catenin by this complex allows binding of an E3 ubiquitin ligase complex that targets β-catenin for proteosomal degradation. Upon activation of WNT signalling, AXIN is sequestered by the activated WNT receptor complex, thereby dismantling the degradation complex and allowing for the accumulation of cytoplasmic β-catenin. Subsequent nuclear translocation of β-catenin allows it to interact with members of the T-cell factor/lymphoid enhancer factor (TCF/LEF) family of transcription factors in order to drive transcription of WNT-target genes [19]. Nuclear localization of β-catenin is therefore a hallmark of active WNT signalling and, within the intestine, is localized to stem cells at the crypt base [20]. WNT ligands are expressed by both epithelial and mesenchymal cells, such as ISC-adjacent Paneth cells [5] and crypt-underlying *FOXL1*+ [21] and *Gli1*+ [22] mesenchymal cells, resulting in the spatial containment of ISCs to the crypt base . Despite the abundance of WNT ligands at the crypt base, their ability to activate WNT signalling is dependent upon R-spondin proteins. R-spondins are a family of four secreted, soluble proteins that bind LGR receptors such as LGR4 and LGR5, both expressed by ISCs [23]. R-spondin binding to LGRs recruits the transmembrane E3 ubiquitin ligases RNF43 or ZNRF3 [24]. Normally, unbound RNF43 and ZNRF3 ubiquitinate and target WNT receptors for degradation, so that the amount of unbound RNF43 or ZNRF3 controls the availability of WNT receptors. R-spondin binding to LGRs, and subsequent sequestration of RNF43 and ZNRF3, thus increases WNT sensitivity. R-spondin control of WNT receptor availability is thus referred to as ligand-independent WNT signalling. Mutations in genes encoding key WNT pathway components such as *Apc* and *Ctnnb1* and

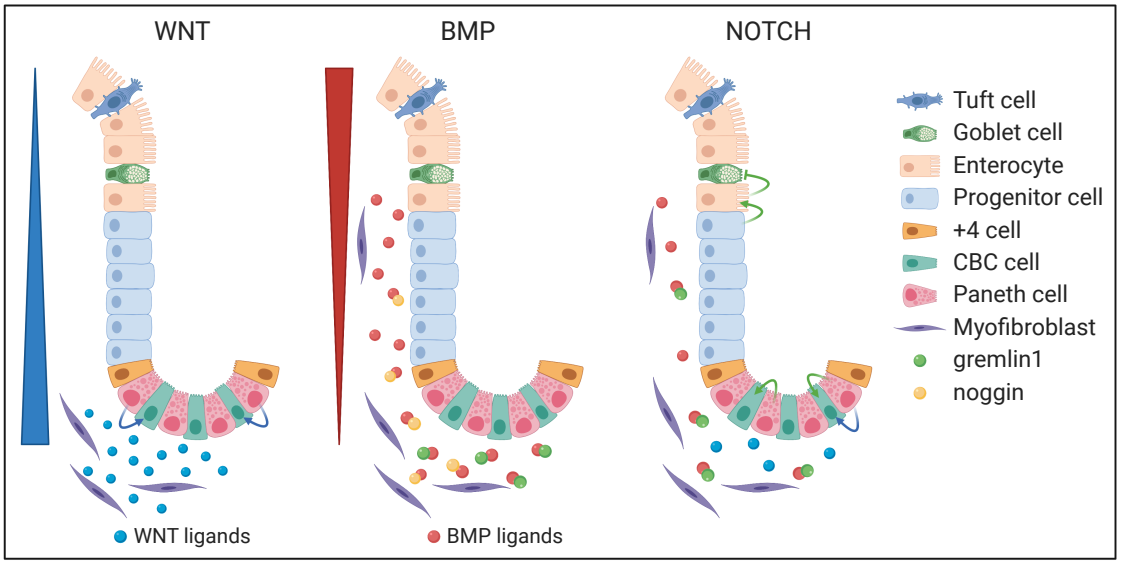


Figure 1.2: Gradients of key signalling pathways within intestinal crypts. High levels of WNT signalling at the crypt base are facilitated by the secretion of WNT ligands by sub-crypt myofibroblasts and stem cell-adjacent Paneth cells. Contrastingly, the secretion of ligand-sequestering BMP antagonists at the crypt base faciltates an inverse gradient of BMP activity. NOTCH signalling in contacting cells acts in concert with Wnt signalling to maintain stemness at the crypt base, while NOTCH-activated enterocytes deprive adjacent cells of NOTCH signals to maintain appropriate levels of differentiation further up the crypt.

overexpression of the WNT activator RSPO1 have been shown to drive intestinal hyperplasia and ISC expansion [25, 26]. Conversely, overexpression of WNT inhibitors blocks WNT signalling activity and results in attrition of the epithelial layer [27]. It is therefore clear that an active, spatially controlled WNT signal is crucial for ISC maintenance and intestinal homeostasis.

1.2.2 Bone Morphogenetic Protein

Bone morphogenetic proteins (BMPs) are a group of signalling molecules within the transforming growth factor-β (TGF-β) superfamily of proteins. BMPs can signal through both SMAD-dependent canonical pathways and SMAD-independent non-canonical pathways. In canonical signalling, BMP ligands bind and activate membrane receptor complexes that in turn phosphorylate receptor-regulated SMAD (R-SMADs). The R-SMADs involved in BMP signalling are SMAD1, SMAD5, and SMAD8 (SMAD1/5/8). Phosphorylation of R-SMADs enables

them to complex with SMAD4 and translocate to the nucleus, where the SMAD complex functions as a transcription factor in order to regulate gene expression. Inhibitory SMADs (I-SMADs), SMAD6 and SMAD7, are involved in feedback inhibition of BMP signalling. Within the intestine, mesenchymal expression of BMP2 and BMP4 activates BMP signalling, which inhibits proliferation in favour of differentiation at the crypt–villus border [20]. This is evidenced by studies that have shown that mice lacking a BMP receptor [28] or those overexpressing BMP inhibitors [29] exhibit hyperproliferation within the intestine. In order to segregate proliferation and differentiation along the crypt-villus axis, BMP ligands secreted by mesenchymal cells are sequestered at the crypt base by BMP antagonists such as gremlin-1, gremlin-2 and noggin [20]. Gremlin1 is a secreted, soluble BMP antagonist that binds BMP2 with the highest affinity, followed by BMP4 and BMP7 [30]. Expression of these ligand sequestering BMP antagonists is restricted to the ISC-underlying mesenchyme at the crypt base [22, 31], resulting in a gradient of BMP activity that is strongest at the top of crypts, where differentiation occurs. Tightly controlled mesenchymal expression of BMPs and their antagonists therefore facilitates the balance between proliferation and differentiation along the crypt-villus axis.

1.2.3 TGF-β

Transforming growth factor-β (TGF-β) is an immune-suppressive cytokine produced by a variety of cells within the intestine. Active TGF-β binds to its membrane receptor complexes and activates intracellular signalling through SMAD-dependent canonical and SMAD-independent non-canonical pathways. In SMAD-dependent canonical signalling, phosphorylated R-SMADs form a complex with SMAD4 and enter the nucleus to regulate the transcription of TGF-β target genes. The R-SMADs involved in TGF-β signalling are SMAD2 and SMAD3. Intact TGF-β signalling maintains intestinal homeostasis through the activation of target genes that govern cell proliferation, apoptosis and immune responses. Higher levels of stromal TGF-β further up the crypt–villus axis therefore result in

the exposure of potential tumour-initiating cells to tumour-suppressive signals [32]. However, loss of TGF-β signalling results in the formation of more aggressive intestinal tumours [33] and dysregulated intestinal TGF-β signalling is observed in inflammatory bowel disease (IBD) patients [34]. The hijacking of dysregulated TGF-β signalling by cancer cells to promote tumour progression further highlights the context-dependent effect of TGF-β pathways [35]. Intact TGF-β signalling therefore plays a tumour-suppressive role in the maintenance of intestinal tissue homeostasis through the regulation of epithelial cell transformation.

1.2.4 Notch

The Notch pathway regulates a variety of cellular processes including proliferation, differentiation and cell fate determination. Notch itself is a transmembrane cell-surface receptor that interacts with ligands from the delta and jagged families. In canonical Notch signalling, a Notch receptor interacts with a canonical Notch ligand on a contacting cell, which initiates cleavage of the receptor and subsequent release of its intracellular domain (ICD). The Notch ICD then translocates to the nucleus where it interacts with various transcription factors to activate transcription of Notch target genes [36]. Notch signalling also shows classical feedback inhibition, in that Notch-activated cells downregulate their own Notch ligand expression and thus deprive adjacent cells of Notch-activating signals [37]. This mechanism enables regulation of cell-fate decisions within neighbouring cells. In the intestine, Notch activity determines lineage decisions between enterocyte and secretory cell differentiation [38]. Inhibition of the Notch pathway results in an increase in the number of goblet cells [39], while increased Notch signalling results in goblet-cell depletion [40]. Notch signalling also plays a role in the crypt, where it acts in concert with WNT signalling to maintain crypt homeostasis. The ISC microenvironment produces Notch-activating signals in order to maintain 'stemness', which is consistent with the observation that Paneth cells at the crypt base express Notch ligands [5]. Notch signalling therefore plays an important role in the maintenance of intestinal homeostasis, where it directs

proliferation within the WNT-high environment of the crypt base and enterocyte differentiation within the WNT-low environment at the top of crypts.

1.2.5 Interleukins

Intestinal immune cells also play an important role in the maintenance of intestinal homeostasis through the production of regulatory signals. Group 3 innate lymphoid cells (ILC3s) are situated throughout the intestine at both mucosal surfaces and within the lamina propria that underlies crypts. These cells lack antigen receptors but respond to intestinal injury through the production of immune signals called interleukins. Activation of ILC3s in response to intestinal stress results in the production of various immune cytokines, of which IL-22 has been well characterised within the intestine. IL-22 is a member of the IL-10 family of cytokines and is expressed by inflammatory cells in a number of tissues. IL-22 supports epithelial regeneration in mice and thus plays an important role in the intestinal inflammatory response, where it can directly stimulate Janus kinase-signal transducer and activators of transcription (JAK-STAT) signalling [41]. Indeed, ISCs stimulated by IL-22 exhibit increased rates of proliferation, a hallmark of the intestinal epithelial response following injury. Interleukins also influence cell fate determination and lineage decisions within the intestine. Tuft cells are 'sentinels' cells situated near the top of crypts that are able to detect infection and stimulate cellular defences [42]. The secretion of interleukins by tuft cells stimulates nearby ILCs that in turn produce interleukins that influence cell lineage decisions to increase differentiation of mucus-secreting goblet cells [43]. This increase in mucus production forms a protective barrier at the top of crypts that serves to expel pathogens and protect the intestinal stem cell niche at the base of crypts. The role of interleukins in the intestinal stress response is further underlined by their regulation by the environmental sensor aryl hydrocarbon receptor (AHR). AHR is activated by a wide variety of intestinal stressors, including physical stress and dietary ligands, leading to its direct regulation of IL-22 expression and ILC3 production [44]. Activation of AHR signalling in

AHR-deficient mice protected the ISC niche through restoration of the mucosal barrier and prevented excessive proliferation within crypts through the regulation of WNT pathway components [45]. Interleukins thus play a key role in intestinal homeostasis through the maintenance of ISC protection and the regulation of intestinal stress responses. Of these, IL-22 is a critical regulator of epithelial homeostasis and has been implicated in tissue regeneration and cell proliferation, pathogen defence and maintenance and protection of the intestinal barrier [46].

1.2.6 Disruption of the intestinal microenvironment

Transient, controlled disruption of intestinal homeostasis is employed in the event of activation of inflammatory or regenerative epithelial responses. Many of the highly evolutionarily conserved pathways that control cell fate determination within the intestine are thus also involved in the physiological response to intestinal damage. Resultant shifts in stem cell plasticity, proliferation, enhanced migration and reduced apoptosis underpin the epithelial response to injury. Following successful epithelial restitution, the homeostatic signalling balance is restored. However, if inflammation is chronic or recurrent, this adaptable cell reprogramming response may be prolonged or hijacked by (epi)genetic mutation, thus increasing the risk of neoplasia. This underlines the importance of the transience of any controlled disruption in intestinal signalling, such that the chronic or recurrent inflammation associated with numerous intestinal diseases, such as inflammatory bowel disease (IBD), Ulcerative Colitis (UC) and Crohn's disease, may predispose to the development of CRC [47]. The role of specific signalling pathways within disrupted intestinal homeostasis is therefore an area of intense research focus; in the next section I focus on signalling pathways shown to play supportive roles in intestinal regeneration and neoplastic transformation.

1.2.6.1 Inflammation: IL-22

Inflammation within the intestine is a stress response invoked and sustained by immune cells. Upon detection of intestinal pathogens or a breach of the

intestinal barrier, inflammatory cells invade the affected region and result in the production of cytokines that direct the inflammatory response [35]. Of these, IL-22 represents a major communication channel between the immune system and intestinal epithelial cells. This is evidenced by IL-22 inhibition or deficiency leading to a reduced frequency of *Lgr5*+ ISCs and defective epithelial regeneration during acute injuries and inflammation [48]. IL-22 is also required for activation of the DNA damage response in the intestinal epithelium. In 2019, Gronke *et al* demonstrated that selective removal of IL-22RA1 from IECs leads to suppression of apoptosis triggered by DNA damage, and, as a consequence, increased tumour formation in an inflammation-driven tumour model [49]. It has also been shown that patients with chronic inflammatory conditions such as inflammatory bowel disease (IBD), Crohn's disease and ulcerative colitis (UC) exhibit increased IL-22 expression within colonic tissue [41, 46]. IL-22 is therefore a key regulator and mediator of intestinal inflammation and failure to restore intestinal epithelial homeostasis is associated with increased levels of IL-22 expression.

1.2.6.2 Disrupted morphogen gradients: HMPS

Disrupted signalling pathways can provoke similar, microenvironmentally-driven changes in cell fate determination as those driven by inflammation-associated immune cells. An example of this in humans is hereditary mixed polyposis syndrome (HMPS), an autosomal dominant condition named for the mixed polyp morphology exhibited by patients. Characteristically, these polyps develop small ectopic crypts (ectopic crypt foci, ECF) that grow orthogonally to the crypt-villus axis [50]. HMPS patients exhibit an increased risk of developing CRC, with a median age of onset of 47 years. HMPS is caused by a 40 kb duplication event upstream of the BMP antagonist *GREM1*, which results in a compartmental expression switch from a restricted mesenchymal gradient to ectopic GREM1 expression throughout the intestinal epithelium [50]. Our lab generated a mouse model of HMPS in which *Grem1* expression is under the control of the intestinal epithelial cell-specific *Villin1* (*Vil1*) promoter. The

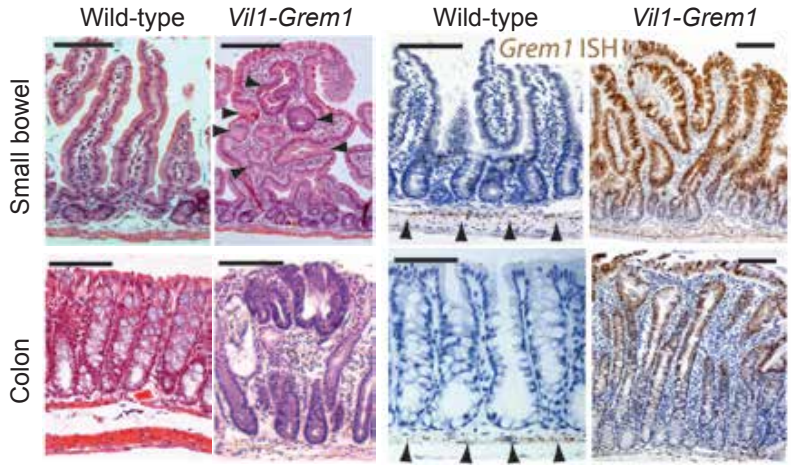


Figure 1.3: *Vil1-Grem1* mouse phenotype. Representative images of wild-type mouse small bowel and colon (leftmost panels) alongside *Vil1-Grem1* small bowel, harbouring widened villi containing intravillus ectopic crypts (black arrowheads), and colon, harbouring luminal surface dysplasia distant from the crypt basal stem cell niche. In situ hybridization (ISH) for mouse *Grem1* (rightmost panels) shows normal *Grem1* expression from subcrypt myofibroblasts (black arrowheads) in wild-type small bowel and colon, and aberrant epithelial Grem1 expression in early small intestinal and colonic *Vil1-Grem1* lesions.

resultant *Vil1-Grem1* (VG) mouse exhibits aberrant epithelial expression of *Grem1* that results in profound intestinal polyposis similar to that observed in human HMPS patients [51]. Polyp formation in VG mice was shown to be driven by an expanded proliferative cell compartment, with ectopic crypts proliferating within the villus compartment of these mice able to acquire mutations that led to intestinal neoplasia [51] (Figure 1.3).

1.2.7 Mouse models of intestinal disease

Mouse models that combine intestinal inflammation with a mutagenic insult thus seek to recapitulate intestinal disease in order to investigate the mechanisms underlying carcinogenesis. In addition to the genetically modified mouse models mentioned to here, a variety of chemical and bacterial agents are widely utilised in order to recapitulate murine intestinal disease phenotypes. Dextran sodium sulfate (DSS) is a water-soluble, negatively charged long chain polysaccharide with a highly variable molecular weight, ranging from 5 to 1400 kDa [52]. DSS dissolved in drinking water is highly toxic to the epithelial lining of the murine

colon, resulting in severe colitis characterized by bloody diarrhea [53]. Multiple cycles of administration of 40-50 kDa DSS in drinking water produces a chronic murine colitis that resembles human UC [54]. The exact mechanism by which DSS induces colitis in mouse colons is unknown but thought to occur via damage to the intestinal epithelial cell barrier increasing its permeability and causing subsequent dissemination of the proinflammatory contents of the gut lumen. Colitis induction can also be achieved via bacterial *Helicobacter hepaticus* (*Hh*) infection. Targeted *Hh* infection can induce colitis, colorectal cancer, and intestinal disease in a variety of immune-deficient mice [55]. *Helicobacter hepaticus* is thus widely utilized as a model system to dissect how intestinal microbiota interact with the host to produce an inflammatory response [56]. Both administration of DSS and infection with *Hh* result in widespread murine colitis that, upon exposure to the mutagen azoxymethane (AOM), can progress to invasive adenocarcinoma [57, 58]. AOM is a potent mutagen that acts as an alkylating agent to induce bulky DNA adducts, misrepair of which induces potentially harmful sequence alterations. AOM is a colonic carcinogen that induces tumours in the distal colon, similar to the localisation of spontaneous human CRC, though metastasis is very rare [58]. Repeated intraperitoneal administration of AOM results in the development of spontaneous tumors within 30 weeks, though in the presence of DSS-driven inflammation this is reduced to 10 weeks [53].

1.3 Homeostatic protection against mutation

1.3.1 Spatial organisation and cell turnover

The balance between proliferation and differentiation within the intestinal epithelium is therefore governed by strictly regulated signalling gradients and immune signals. However, actively proliferating cells along the crypt-villus axis are at risk of acquiring potentially harmful mutations as a result of their exposure to the genotoxic contents of the gut lumen. The spatial organisation of the crypt-villus axis confers some protection here, as the dividing cells of the

stem and TA cell compartments are at the crypt base and thus less exposed to luminal toxins than differentiated surface cells. The rapid turnover of epithelial cells also represents an evolved protective mechanism, as the determined cell fate and limited lifespan of non-stem cells within the crypt protects against the propagation of mutations. Protection is conferred here by intact cell migration, whereby mutated non-stem cells are shed within 3-4 days, a far shorter period than that required for neoplastic transformation to occur [2].

1.3.2 Neutral drift

Proliferative intestinal cells are also at risk of mutations that arise as a result of errors in DNA replication. In this case, TA cells are again protected against mutation accumulation by their rapid turnover. Longer-lived ISCs at the crypt base, however, are instead protected by the mechanism of their division and their positioning within the stem cell niche. It was previously thought that ISCs divide by asymmetric cell division, whereby a single ISC gives rise to one TA cell and one ISC. It has instead been shown that ISC division follows a more random principle of dynamic replacement referred to as 'neutral drift', in which a single ISC can be replaced by any other ISC within the crypt [59, 60]. Stem cells are thus randomly lost from the ISC population and become primed for differentiation, while vacant niches are repopulated by the offspring of neighbouring stem cells. Eventually, due to the stochastic nature of this process, all crypt stem cells will become descendants of the same ISC, resulting in the generation of a clonally related crypt cell population. For a crypt containing 10 ISCs, any single ISC therefore has a 1 in 10 (10%) chance of replacing all other ISCS and populating the entire crypt, referred to as 'fixation'. Live imaging of crypt base *Lgr5*+ cells has also shown that the position of cells within the stem cell niche further impacts upon their likelihood of fixation. Ritsma *et al* found that stem cells occupying 'central' positions at the crypt base exhibit a survival advantage over 'border' stem cells closer to the niche boundary [61]. Functionally, these seminal findings indicate that if one ISC acquires a neutral mutation it has a high chance of being

replaced by another, unmutated ISC within the crypt. However, if an ISC acquires an oncogenic mutation that bestows a selective advantage, this skews neutral drift dynamics to increase the fixation chance of the mutated cell [62]. Vermuelen *et al* demonstrated this in an elegant experiment that utilised lineage tracing to quantify the competitive advantage of oncogenic mutations [63]. Their findings indicate that oncogenic mutations increase the likelihood of a mutated cell replacing its wild-type neighbour and achieving niche fixation. Importantly, however, cells harbouring oncogenic mutations are still subject to stochastic replacement by normal, unmutated ISCs.

Homeostasis within the intestine therefore involves tightly controlled morphogen gradients along the crypt-villus axis that regulate signalling pathways in order to maintain an escalator of cells along the crypt-villus axis. Proliferating cells are protected against the acquisition of mutations by a combination of mechanisms, including the spatial organisation of crypts, high cell turnover of TA cells and neutral drift division employed by ISCs, that reduce the likelihood of mutation fixation throughout a crypt. However, given the wide incidence of intestinal disease and colorectal cancer (CRC), it is evident that mutations are still able to persist in order to give rise to neoplasia within the intestine. In the case of the failure of the above protective mechanisms, cells therefore also employ a comprehensive and sophisticated response to DNA damage, which I will discuss in more detail within the next section. Given that this is a highly complex and evolved field, I will limit my summary of the expansive literature to aspects that are specifically relevant to intestinal epithelial function.

1.4 Mutation control: The DNA damage response

Cells are exposed to a wide variety of genotoxic agents that can cause potentially harmful alterations to the structure and sequence of their DNA. Examples include ultraviolet radiation within sunlight, which can cause up to 1 x 10^5 mutations within a cell in a single day [64], and ionizing radiation, which causes double stranded breaks (DSBs) in the DNA backbone. Endogenous processes

within cells can also impact upon genomic integrity, such as the production of DNA-oxidising reactive oxygen species by normal cellular metabolism and DNA sequence alterations as a result of errors in DNA replication. Within the gut, the impact of dietary carcinogens has been well documented and results in the classification of the contents of the gut lumen as genotoxic [65]. Given the variety and frequency of DNA damage they experience, cells have developed a similarly complex repertoire of mechanisms that counteract these threats to genomic integrity. These mechanisms are referred to collectively as the DNA damage response (DDR) and can be divided into distinct though often functionally overlapping pathways based on the types of DNA lesion they process. Most DDR pathways are comprised of three strictly controlled stages: a 'sensing' step involving the detection of DNA damage; a 'translating' step involving the recruitment of pathway-specific repair components and an 'effecting' stage involving physical repair of the lesion or activation of the appropriate cellular responses as a consequence of irreparable DNA damage.

1.4.1 Excision repair

Subtle changes to DNA such as oxidative lesions, alkylating adducts and single strand breaks (SSBs) are recognised and repaired by excision repair mechanisms. Within these, base excision repair (BER) refers to the removal of damaged DNA bases from the double helix prior to excision and replacement of the damaged region of the backbone with newly synthesised DNA. Central components of BER are the enzymes poly(ADP-ribose) polymerase (PARP) 1 and 2, which act as both sensors and signal transducers [66]. Bulkier single-strand lesions that compromise the structure of DNA, such as those induced by UV radiation, are sensed by the XPC protein for processing by nucleotide excision repair (NER) [67]. NER is initiated upon the detection of bulky adducts during transcription, where they block the progress of polymerases along the DNA strand, or as a result of compromised base pairing as a result of distortion of the DNA helix [68]. Despite differences in lesion detection, these repair pathways process DNA

damage similarly through the excision and replacement of DNA surrounding subtle lesions. PARP enzymes and those involved in excision, such as excision repair cross-complementing protein 1 (ERCC1) are thus central to BER and NER.

1.4.2 Double strand break repair

Double strand breaks (DSBs) form the class of DNA damage with the most serious potential consequences. The utilization of controlled DSB formation as a source of genomic diversity in immune cells, during V(D)J recombination and class switch recombination in B cells, underlines their potential impact upon genomic stability. Ionising radiation (IR) is the most significant external source of DSB induction and, given the widespread use of radiotherapy in cancer treatment, understanding the cellular response to IR in the context of repair pathways has been a long-standing area of intense research. DSB repair is dominated by two pathways: non-homologous end joining (NHEJ) and homologous recombination (HR). NHEJ repair has been further stratified into 'conventional' (c-NHEJ) and 'alternative' (AEJ) pathways. Repair pathway choice in cells harbouring DSBs is dependent upon a variety of factors including cell cycle stage and the availability of key pathway components.

The first step of DSB repair involves the 'sensing' of DNA breaks by nearby histone proteins, which are phosphorylated at the site of damage in order to trigger recruitment of proteins involved in the various repair pathways. Phosphorylation of H2A histone family member X (H2AX) by kinases including ataxia-telangiectasia mutated (ATM) and DNA-dependent protein kinase (DNA-PK) has been well characterised as one of the earliest stages of DDR [69]. Generation of phosphorylated H2AX (γ-H2AX) at sites of DSBs is rapid and thought to peak around 30 minutes following exposure to IR [70]. The pathway employed to repair DSBs is determined by the proteins recruited to the sites of DSBs following sensing of DNA damage.

1.4.2.1 Homologous recombination

HR is the most stringent DSB repair pathway as it involves extensive resection of DNA ends followed by homology-directed sequence generation. In order to achieve this, HR is defined by the recruitment of proteins involved in resection and homology coordination. Upon sensing of a DSB, DNA end resection is initiated by recruitment of the MRE11-RAD50-NBS1 complex [71]. The resultant single stranded DNA (ssDNA) is coated in RPA proteins to prevent inadvertent binding to proximal unrelated ssDNA [72]. RPA proteins are displaced by recombination mediators, such as RAD51 and the BRCA proteins, that coordinate the search for homology and facilitate base pairing to complementary sequences, often on the sister chromatid [73]. The error-free repair associated with this pathway is a result of subsequent homology-directed extension of the resected DNA end by DNA polymerase enzymes prior to ligation to complete the repair process. This search for homology increases the time taken for the completion of HR directed repair so that these pathways are considered the 'slow, stringent' DSB repair choice. In mammalian cells, HR rarely utilises a homologous chromosome as a template for repair but instead employs the sister chromatid present during DNA replication. As a result, HR is largely restricted to the late S and G2 phases of the cell cycle to repair DSBs generated during DNA replication [73].

1.4.2.2 Non-homologous end joining

In contrast to HR directed repair, NHEJ is a rapid, high-capacity pathway that involves the joining of broken DNA ends with minimal reference to homologous sequences of DNA. The lack of requirement for homology enables NHEJ to function across all stages of the cell cycle [74]. Classical NHEJ involves binding of the high-affinity XRCC6 and XRCC5 (Ku70 and Ku80) protein dimer to the broken ends of DSBs. Binding of the Ku dimer drives recruitment of NHEJ repair components to the site of the DSB. Of these, DNA-dependent protein kinase catalytic subunit (DNA-PKcs) activates proteins required for NHEJ in the vicinity of the DSB [75]. Given that most DSBs will be comprised of two

incompatible DNA ends that are unable to be directly ligated together, NHEJ employs minimal end resection in order to generate short 5' and 3' DNA overhangs. End resection is coordinated by a wide variety of exo- and endonucleases that are recruited by DNA-PKcs, such as Artemis, CtBP-interacting protein (CtIP), Werner syndrome ATP-dependent helicase (WRN), flap endonuclease 1 (FEN1), exonuclease 1 (EXO1) and the MRN complex [76]. In NHEJ, minimal end resection exposes short DNA overhangs containing small regions (< 4 nucleotides) of microhomology that facilitate end joining. This is in stark contrast to the extensive end resection (> 20 nucleotides) and homology-directed repair utilised in HR. Subsequent gap filling and ligation by DNA polymerase enzymes and DNA Ligase IV complete the NHEJ repair process [77]. The lack of extensive resection or a search for homology reduces the time taken for completion of NHEJ but increases the potential for inaccurate repair, given the variable level of end resection and gap filling required per DSB.

1.4.2.3 Alternative end joining

Relative to HR and c-NHEJ, alternative end joining pathways are relatively recently characterised and have been suggested to be key drivers of genomic instability. In the absence of the Ku proteins central to c-NHEJ, DNA resection machinery exposes single strand DNA (ssDNA) for repair via either HR or AEJ pathways [78]. AEJ repair processes encompass single strand annealing (SSA) and microhomology-mediated end joining (MMEJ), both of which produce errors in the repaired DNA sequence. SSA involves the annealing of complementary nucleotide sequences within long strands of resected DNA, prior to removal of non-complementary 3' tails, gap filling and ligation [79]. Given that annealing of complementary ssDNA may occur between chromosomes, SSA can result in significant deletion and translocation events. In MMEJ, a lack of c-NHEJ proteins allows for PARP1-driven recruitment of the MRN complex to broken DNA ends [80]. In combination with CtIP, the MRN complex directs extensive resection in order to generate long stretches of ssDNA [81]. The central component of AEJ

is Polymerase θ, a multifunctional protein that possesses helicase, polymerase and transferase abilities [82]. Pol θ competes with HR-associated recombination proteins to displace RPA from stretches of ssDNA and can also directly interact with RAD51 to inhibit HR. The helicase function of Pol θ displaces RPA from ssDNA to reveal regions of microhomology; its polymerase function can stabilize the annealing of 3' ssDNA tails with as little as 2 bp of homology and its terminal transferase activity can add nucleotides to provide or extend existing microhomology [83]. The Pol θ-stablised ssDNA intermediate can then be ligated by DNA ligase I or III, although LIGIII is thought to be the primary ligase in AEJ repair given its preferential association with PARP1 [84]. AEJ pathways therefore function as potential back up repair option in the event of reduced availability of Ku-driven c-NHEJ. In terms of DDR kinetics, it has been suggested that AEJ pathways are slower than NHEJ, given that immunoglobulin class switching in the absence of c-NHEJ-associated DNA ligase IV occurs tenfold more slowly, possibly due to the requirement of extensive resection in AEJ compared to NHEJ [85]. The coupling of extensive resection and microhomology-mediated end joining in AEJ therefore results in frequent, large-scale sequence alterations, rendering the repair of DSBs via AEJ pathways inherently mutagenic.

1.4.3 DDR pathway choice

Of the three DSB repair pathways outlined above, HR is thus generally considered to be error-free due to its use of sister chromatid templates to direct repair, although this restricts its availability to the S and G2 phase of the cell cycle; Ku-dependent c-NHEJ is available at all stages of the cell cycle but can introduce sequence alterations as a consequence of direct end ligation or inaccurate gap filling and AEJ repair operates as an inherently mutagenic pathway in the absence of key c-NHEJ components (Figure 1.4). It is therefore evident that the impact of each pathway on genomic stability varies. The factors that govern pathway choice are therefore the focus of extensive research, with competitive interactions between several components from each of the pathways shedding light on the mechanisms

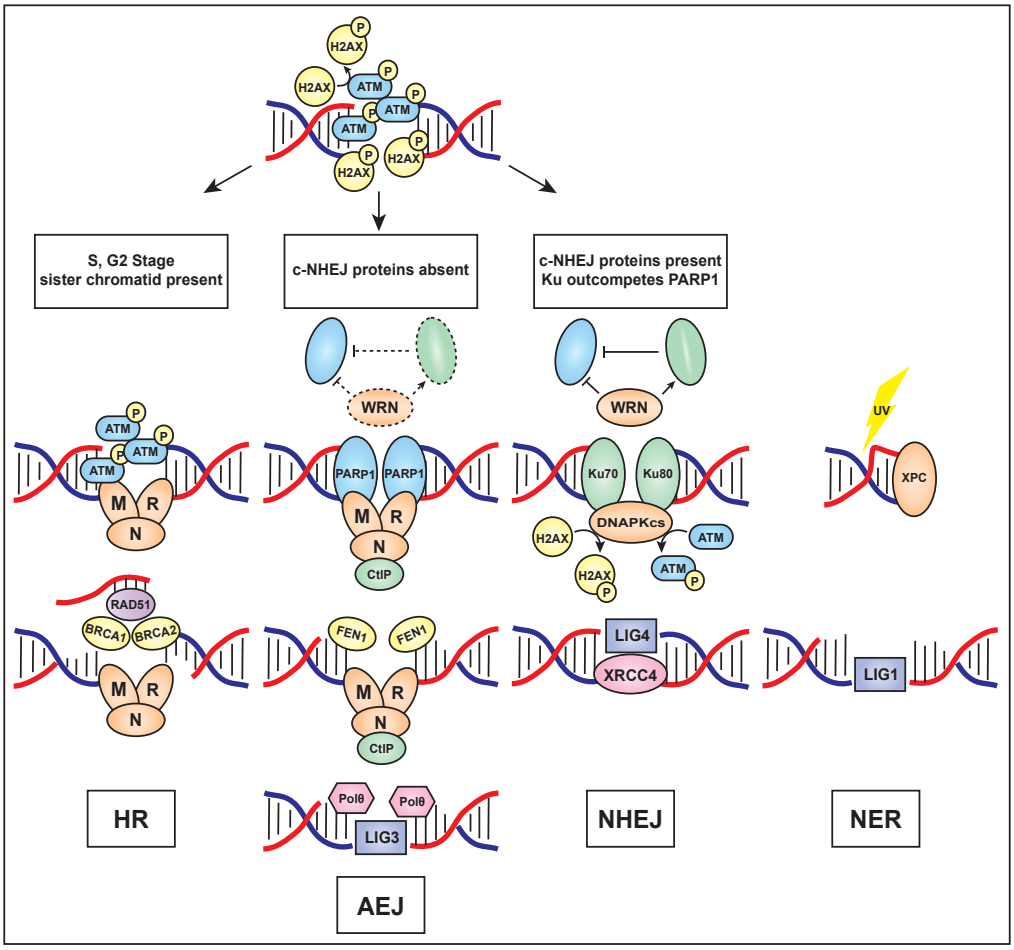


Figure 1.4: Schematic of the repair pathways utilised upon detection of DNA damage. The recruitment and interaction of key components of homologous recombination (HR), alternative end joining (AEJ), non-homologous end joining (NHEJ) and nucleotide excision repair (NER) upon sensing of a double strand break or bulky adduct caused by UV radiation are illustrated. Boxes indicate conditions influencing pathway choice.

that govern DDR pathway choice. This competition seems to be focused upon the earliest stages of DDR, during the recruitment of repair machinery. Prioritisation of c-NHEJ is driven by the high affinity of Ku proteins for DNA ends and their increased abundance compared to the HR and AEJ-associated MRN complex [86]. Additionally, WRN, a multifunctional protein with exonuclease and helicase abilities, is rapidly recruited to DSBs, where its strong association with Ku proteins and inhibition of CtIP and MRE11 binding protects DNA ends from HR- and AEJ-associated resection and promotes Ku-driven c-NHEJ repair [87]. The rapid recruitment of WRN to DSBs thus further reduces the likelihood of MRN end binding when Ku proteins are present. However, long strands of ssDNA bind

with reduced affinity to Ku proteins and so are channelled less efficiently to c-NHEJ [88]. The wide availability and prioritisation of c-NHEJ thus enables it to be employed at all stages of the cell cycle, while stringent HR repair during the S and G2 stages protects the genomic integrity of dividing cells and AEJ pathways compete with HR in the event of reduced c-NHEJ availability.

1.4.4 DDR in the gut

1.4.4.1 DDR along the crypt-villus axis

It has been shown that although all cell populations along the crypt-villus axis accrue rapid formation of γH2AX foci following irradiation, the ensuing repair of DSBs varies according to compartment. Cells at the crypt base exhibit rapid and complete resolution of γH2AX foci in the initial 4-6 h following irradiation, while only minimal resolution is observed in cells within the TA and villus compartments [89]. Contrastingly, TA cells resolve DNA-PKcs foci, indicative of NHEJ repair, only slightly less efficiently than cells at the crypt base, with NHEJ activity absent in villus cells [89]. These findings indicate while proliferative crypt base and TA cells both employ NHEJ effectively, cells at the crypt base utilise HR to resolve DSBs more efficiently (Figure 1.5). Further work supporting this notion showed enrichment of HR-associated BRCA1 and RAD51 expression within crypt base cells [90]. In terms of the 'effecting' stage of DDR pathways, proliferative cells at the crypt base mount a robust and rapid apoptotic response that culminates in peak cleaved caspase-3 levels 6 h following irradiation [89]. Given that apoptosis within the TA compartment peaks later, at 12 h following irradiation, these data indicate that cells at the crypt base make earlier cell fate decisions, following rapid completion of the majority of DNA repair processes. Taken together, these findings indicate that while proliferative cells along the crypt-villus axis employ universally available NHEJ to resolve DSBs, ISCs at the crypt base also employ stringent HR processes to rapidly resolve DSBs in order to preserve genomic integrity within this compartment. The differential response to DNA damage along the crypt-villus axis thus indicates that some cells may be more dispensable

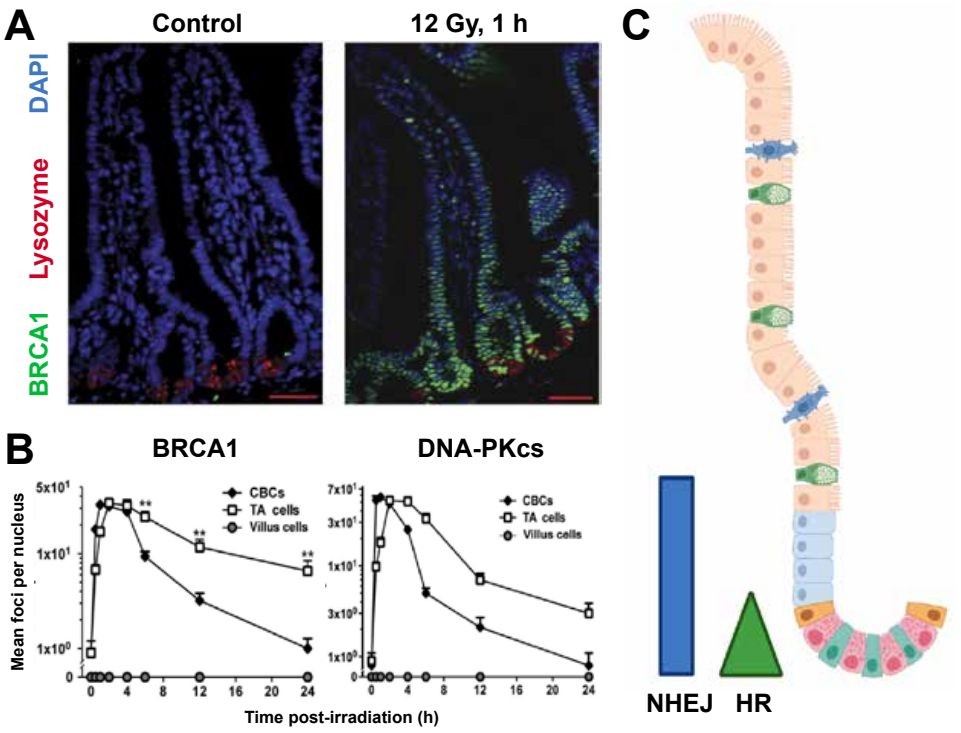


Figure 1.5: DDR along the crypt villus axis. A) Spatial distribution of rapid BRCA1 foci formation within the intestine 1 h post-irradiation. B) Temporal dynamics of homologous recombination repair, represented by formation and resolution of BRCA1 foci, and non-homologous end joining, represented by DNA-PKcs foci, within crypt base, transit amplifying and villus cells following 12 Gy ionizing radiation. Adapted from [89]. C) Model depicting the utilisation of homologous recombination (HR) and non-homologous end joining (NHEJ) along the intestinal crypt-villus axis.

than others, a notion that has been explored extensively in studies investigating the dispensability of stem cells at the crypt base following irradiation.

1.4.4.2 Irradiation-induced stem cell plasticity

Within the intestine, intact DDR pathways therefore contribute to the maintenance of genomic stability during intestinal homeostasis. The consequences of impaired DSB repair have been highlighted in numerous studies that have utilised ionizing radiation to investigate repair dynamics within cell populations along the crypt-villus axis. Within this context, radio-resistance is defined as the capacity of cells to survive and continue functioning following exposure to IR. The dose of IR delivered is clinically relevant here, with the mouse intestinal epithelium capable of regeneration following exposure to IR doses below 14 Gy but doses

above 15 Gy resulting in gastrointestinal acute radiation syndrome (GI-ARS) and mortality [91]. Given that IR targets dividing cells, the response of ISCs and progenitors within crypts has been the focus of extensive research. Seminal work investigating the susceptibility of *Lgr5*+ stem cells to IR found that this population is effectively ablated following exposure to 12 Gy IR, but that this does not impact upon subsequent intestinal regeneration [92]. This indicated the presence of a non-*Lgr5*+ ISC population that is relatively radioresistant and can repopulate crypts in lieu of *Lgr5*+ cells following IR. Elegant lineage tracing studies have since identified a remarkable degree of plasticity within populations of proliferative cells in crypts, with several different cell types capable of regenerating the epithelium following IR. Quiescent cells located at the +4 position of crypts, and identified by their expression of markers including *Krt19*+ [93], *Sox9*+ [94] and *Bmi1*+ [95], are radioresistant cells that survive IR exposure in order to fall back into the stem cell niche and regenerate crypts. While multiple studies have confirmed that *Lgr5*+ ISCs radiosensitive cells undergo apoptosis within 24 h of IR, Hua *et al* showed that high-dose irradiation does not completely eliminate *Lgr5*+ cells and that surviving *Lgr5*+ cells can go on to replenish the intestinal epithelium [96]. Given that the active *Lgr5*+ ISC pool has a relatively slow proliferation rate of around 21.5 h, whereas progenitor cells divide once every 12 h [97], it may be that *Lgr5*+ ISCs that survive irradiation are at phases of the cell cycle that are less sensitive to DNA damage. Another hypothesis is that cell populations marked by *Lgr5* are heterogenous and the subpopulations contained therein harbour differential radio-sensitivity. It is clear then that intestinal crypts harbour heterogenous populations of cell that vary in their radio-sensitivity, with the resulting plasticity enabling repopulation of crypts and subsequent epithelial regeneration following exposure to IR. It thus appears that regulated, temporary disruption of morphogen gradients in inflamed and regenerative intestine is essential to alter cell proliferation migration and proliferation in order to effect epithelial restitution. However, the same characteristics are key attributes of malignant cells, such that failure to restore tight homeostatic control following

intestinal inflammation or pathological disruption of morphogen gradients can result in initiation of neoplasia.

1.4.4.3 The epithelial response to irradiation

The intestinal epithelial response to IR can thus be divided into apoptotic, proliferative and normalisation phases. The apoptotic phase, within the first 2 days following irradiation, involves apoptosis-driven crypt loss, reduction in crypt sizes and villi shortening [91]. The regenerative phase, 2 to 4 days post irradiation, involves the regeneration of entire crypts by those cells that have survived exposure to IR. An approximate doubling in the number of proliferative cells during this phase drives an increase in regenerative crypt size despite overall decreases in crypt number. Finally, the normalisation phase involves homeostatic restoration of crypt and villi size. A key regulator of the epithelial response to DNA damage is the transcription factor p53, a multifunctional protein upregulated upon IR-induced DDR pathway activation [98]. Subsequent nuclear accumulation of p53 induces a variety of downstream signalling pathways that mediate cellular responses to stress. Key p53 target genes include the cell cycle regulator P21, which initiates arrest of the cell cycle in order to enable DSB repair within crypts, and p53 upregulated mediator of apoptosis (PUMA), named for its activation of apoptosis following DNA damage [99]. The importance of p53 in cellular responses to DNA damage is evidenced by mouse models in which suppression of p53 results in reduced levels of IR-induced apoptosis [100]. Epithelial cells within the crypts of p53-null mice are resistant to IR-induced apoptosis and fail to exhibit decreases in proliferation within the first 24 h following irradiation, instead exhibiting defects in cell cycle arrest and continued proliferation [101]. These findings underscore the diverse role of p53 signalling within the epithelial response to IR, whereby activation of p53-mediated signalling can cause cell cycle arrest, to facilitate DNA repair and promote cell survival, or the induction of apoptotic pathways that facilitate cell death. Irradiation therefore causes a complex disruption in

homeostatic signalling that facilitates the cellular reprogramming required for epithelial cell plasticity and subsequent epithelial regeneration.

1.4.5 Signalling pathways and DDR

Given the overlap in outcomes, such as apoptosis and proliferation, between DDR and homeostatic signalling pathways within the intestine, numerous studies have sought to investigate crosstalk between specific signalling pathways and DDR. Here, I will focus on recent work that has suggested links between the TGF-β and BMP signalling pathways and the DNA damage response (summarised in Figure 1.6).

TGF-β As previously discussed, intact TGF-β signalling plays a tumour-suppressive role in the maintenance of intestinal tissue homeostasis through the regulation of epithelial cell transformation. It has also been shown that TGF-β signalling has an impact on genomic stability and DDR pathways: inhibition of TGF-β signalling results in increased genomic instability [102], while DNA damage-induced activation of ATM stabilises TGF-β receptor II (TGFBR2) in order to trigger cell cycle arrest and enable DNA repair [103]. Additionally, head and neck squamous cell carcinoma (HNSCC) cells negative for TGF-β activity exhibit compromised HR-directed DNA repair and an increased reliance upon AEJ [104]. Given the reliance of AEJ pathways upon the PARP1 enzyme, this work provided a mechanistic basis for the use of PARP1 inhibitors in TGF-β-negative HNSCC. Further work examining the relationship between TGF-β activity and DDR pathway components has suggested a role for microRNA-182. MicroRNAs (miRNAs) are a class of short (19-24 bp) non-coding RNAs involved in the regulation of gene expression at a post-transcriptional level, through degradation or translation inhibition of target mRNAs. miR-182 is a member of the miR-183-96-182 cluster, a well characterised oncogenic miRNA family upregulated in a range of cancers [105]. MiRNAs may therefore represent an additional regulatory layer linking intestinal signalling pathways and DDR. The complexity

of the mechanisms underlying microRNA expression, however, is highlighted by conflicting evidence regarding the impact of TGF-β activity upon miR-182 expression. Yu *et al* reported that TGF-β activity induces miR-182 expression and subsequent inhibition of SMAD7 as part of a negative feedback loop within breast and lung cancer cell lines [106], while Barcellos-Hoff *et al* demonstrated that inhibition of TGF-β ligand or signalling in all cells significantly increased the abundance of miR-182 in murine mammary epithelial cells [107].

BMP Despite the inherent crosstalk between TGF-β and BMP signalling as a result of their joint utilization of SMAD4 as a key effector protein, there has been surprisingly little work investigating a link between the BMP pathway and DDR. It has previously been shown that the R-SMAD proteins utilised by BMP signalling (SMAD1, SMAD5, SMAD8) are phosphorylated at C-terminal SXS motifs in order to trigger dimerisation with SMAD4 prior to nuclear translocation. De-phosphorylation of R-SMADs results in their subsequent relocation to the cytoplasm. In 2012, Chau *et al* suggested that ATM-dependent phosphorylation of SMAD1 represented a link between BMP signalling and DDR pathways [108]. Their work demonstrated that upon exposure to IR, ATM phosphorylates SMAD1 at residue S239, disrupting its interaction with its negative regulator PPM1A. This inhibited cytoplasmic relocation of SMAD1, resulting in enhanced levels of nuclear SMAD1. Given that the same study found that SMAD1 interacts with p53 in order to inhibit p53 degradation, these findings suggest that BMP-SMAD1 signalling intersects with DDR pathways via an ATM-p53 axis. Elegant work by Vattulainen *et al* in 2018 has also demonstrated that inactivation of BMPR2 in pulmonary vascular endothelial cells results in a deficit of RAD51 [109]. Restoration of BMPR2 or activation of the BMP signalling pathway rescued RAD51 levels and prevented DNA damage within this model. Interestingly, in parallel with TGF-β signalling, the suggested mechanism linking BMP and DDR pathways involves a member of the miR-182 cluster. miR-96 directly targets the coding region of RAD51 mRNA

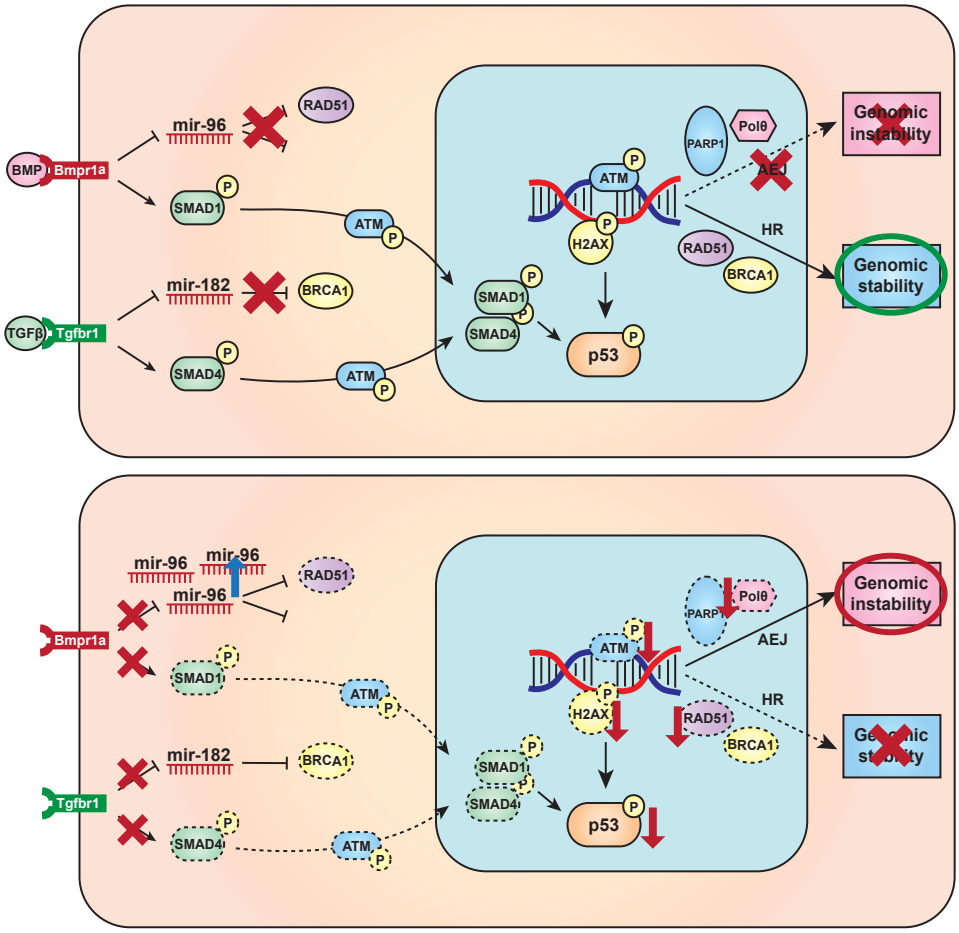


Figure 1.6: Model depicting the proposed links between TGF-β and BMP signalling and DDR outcomes. Represented as DDR pathway activation in the presence (top) and absence (bottom) of TGF-β and BMP signalling.

[110] and miR-96 expression has been shown to be repressed upon stimulation of BMP signalling in pulmonary artery smooth muscle cells [111]. Furthermore, mirroring the impact of miR-182 upon genomic instability, overexpression of miR-96 decreased the efficiency of HR repair and enhanced sensitivity to PARP1 inhibitors in human cancer cells, while miR-96-overexpressing breast cancer cells exhibited increased sensitivity to the DNA damaging agent cisplatin in a mouse xenograft model [110].

Rapid renewal of the intestinal epithelium is thus achieved through the unique spatial organisation and stringently controlled signalling pathways that govern cell populations along the crypt-villus axis. During intestinal homeostasis, protection

against the acquisition and accumulation of mutations is achieved through high epithelial cell turnover and the employment of DDR pathways. However, these evolved protective mechanisms against the propagation of mutated cells are not infallible and colorectal cancer has long been considered to arise as a consequence of the accumulation of sequential epithelial mutation.

1.5 Colorectal Cancer

1.5.1 The adenoma-carcinoma sequence

Seminal work by Fearon and Vogelstein in 1990 established the prevailing model of CRC development, whereby an accumulation of mutations drives tumour formation and progression [112]. Within this model, each mutational 'hit' represents a step along the adenoma-carcinoma sequence and drives neoplastic transformation of normal tissue. Activating mutations within the Wnt pathway are thought to represent the most likely first step, with loss of the *APC* gene comprising the most common Wnt mutation within sporadic human CRC [113]. Progression to adenoma and carcinoma formation is then driven by further accumulation of mutations within key KRAS, TGF-β, PI3K and p53 signalling pathway components [114] (Figure 1.7). These include activating mutations of *KRAS* and *BRAF* and inactivating mutations of *p53*, *SMAD4* and *PTEN* [115]. Enhanced genetic predisposition to CRC is exhibited by patients of familial syndromes such as famililal adenomatous polyposis (FAP) and hereditary mixed polyposis syndrome (HMPS), which are caused by mutations within Wnt (*APC*) and BMP (*GREM1*) pathway components, respectively. Given the widespread incidence of CRC as the 4th most common cancer worldwide, with the 2nd highest mortality rate [116], revealing the mechanisms that underlie CRC disease pathogenesis is an area of considerable research interest. The conventional adenoma-carcinoma sequence described initially has thus been updated to reflect the discovery of other key factors influencing tumourigenesis within the colon.

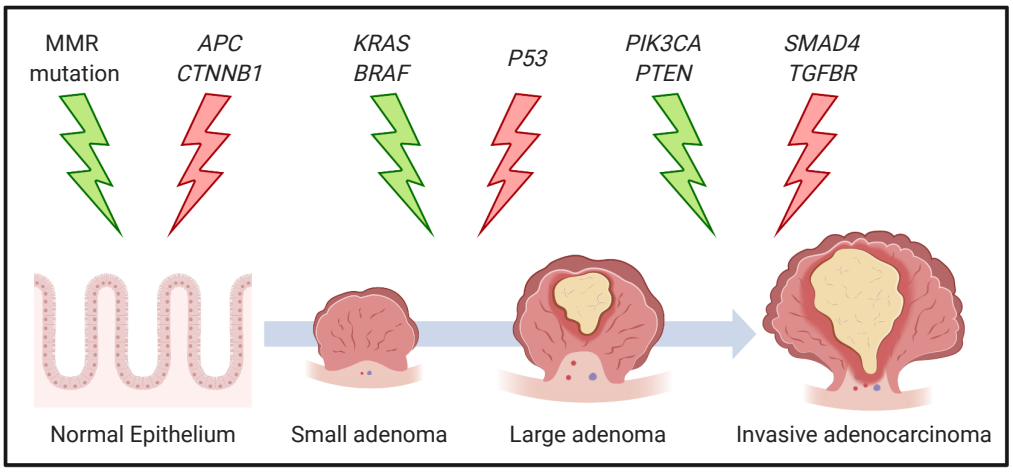


Figure 1.7: The adenoma-carcinoma sequence. Carcinogenesis within the colon is thought to occur via the step-wise accumulation of activating mutations in tumour-promoting oncogenes (green) and inactivating mutations in tumour-suppressor genes (red). The acquisition of mutations in Wnt pathway components (*APC* or *CTNNB1*) or genes involved in mismatch-repair (MMR) represent the earliest stages of adenoma formation within normal colonic epithelium. Subsequent accumulation of mutations in key driver genes is associated with further progression to an advanced adenoma harbouring high-grade dysplasia and then to invasive carcinoma.

1.5.2 Added complexity: cell of origin and the microenvironment

Given the ability of basal ISCs to repopulate entire crypts, the initial dogma regarding the cell of origin of CRC involved a 'bottom-up' model whereby ISCs at the crypt base mutate to give rise to upward neoplasia. The discovery of expression markers specific to ISCs enabled functional evaluation of this notion, with numerous studies since establishing that specific induction of *Apc* deletion within *Lgr*5+ [117] and *Lrig1*+ cells [13], and activating exon-3 *Ctnnb1* mutations within *Bmi1*+ [15] and *CD133* cells, [118] at the crypt base can induce adenoma formation. It thus appeared that Wnt activation in basal ISCs is sufficient for murine adenoma formation. However, elegant work has since demonstrated the potential of cells outside of the crypt base niche to drive neoplasia upon the acquisition of key CRC driver gene mutations, especially when combined with microenvironmental disruption. Here, loss of *Apc* alone within differentiated *Dclk1*+ tuft cells was unable to induce tumour formation [119], but combining APC-deficient *Dclk1*+ cells with a mouse model of colitis resulted in tumour formation

[42]. Similarly, targeting of an activating β-catenin mutation to differentiated *Xbps1*+ cells was insufficient for epithelilal transformation, but the same mutation coupled with increased inflammatory NF-Kβ signalling activity resulted in tumour formation [120]. These results demonstrate the tumourigenic potential of non-stem cells harbouring CRC driver mutations when combined with extra events, such as inflammation. Importantly, it has also been shown that changes in the intestinal microenvironment alone can also initiate tumour formation in non-stem cells [51]. Deregulation of morphogenetic pathways can thus play a key role in cancer development, with both spontaneous and constitutively disrupted signalling pathways associated with neoplastic transformation.

1.5.3 Polyp pathways

The significance of the cell of origin to intestinal disease is illustrated in polyps, the benign precursor lesions to every established CRC. Conventional (tubular and tubulovillous) adenomas comprise the majority of those detected during endoscopic surveillance. These polyps are thought to occur via an initiating *APC* or *CTNNB1* mutation in an *LGR5*+ stem cell at the crypt base that induces an endogenous, constitutive Wnt drive, such that these mutations alone are sufficient to induce dysplasia. Contrastingly, traditional serrated adenomas (TSAs) are a distinct entity of sessile serrated lesions [121] and are rare (<1% of resected polyps) [122]. TSAs account for 5% of all serrated polyps and have been associated with a significant cancer risk, suggested to be twice that of conventional polyps [123]. Our lab has demonstrated that some TSAs arise through aberrant epithelial expression of *GREM1*, which alters cell fate determination along the crypt-villus axis and causes an expansion of progenitor cells [51]. In TSAs, the cell of origin is therefore the ectopic crypt progenitor cell, with *KRAS* or *BRAF* the earliest detectable somatic mutations. TSAs also preferentially acquire mutations that potentiate and amplify endogenous, ligand-dependent Wnt signalling [124], as opposed to the constitutive, ligand-independent Wnt activation caused by *APC* and *CTNNB1* mutations. A strong, constitutive Wnt drive in an *LGR5*+ cell of origin might be

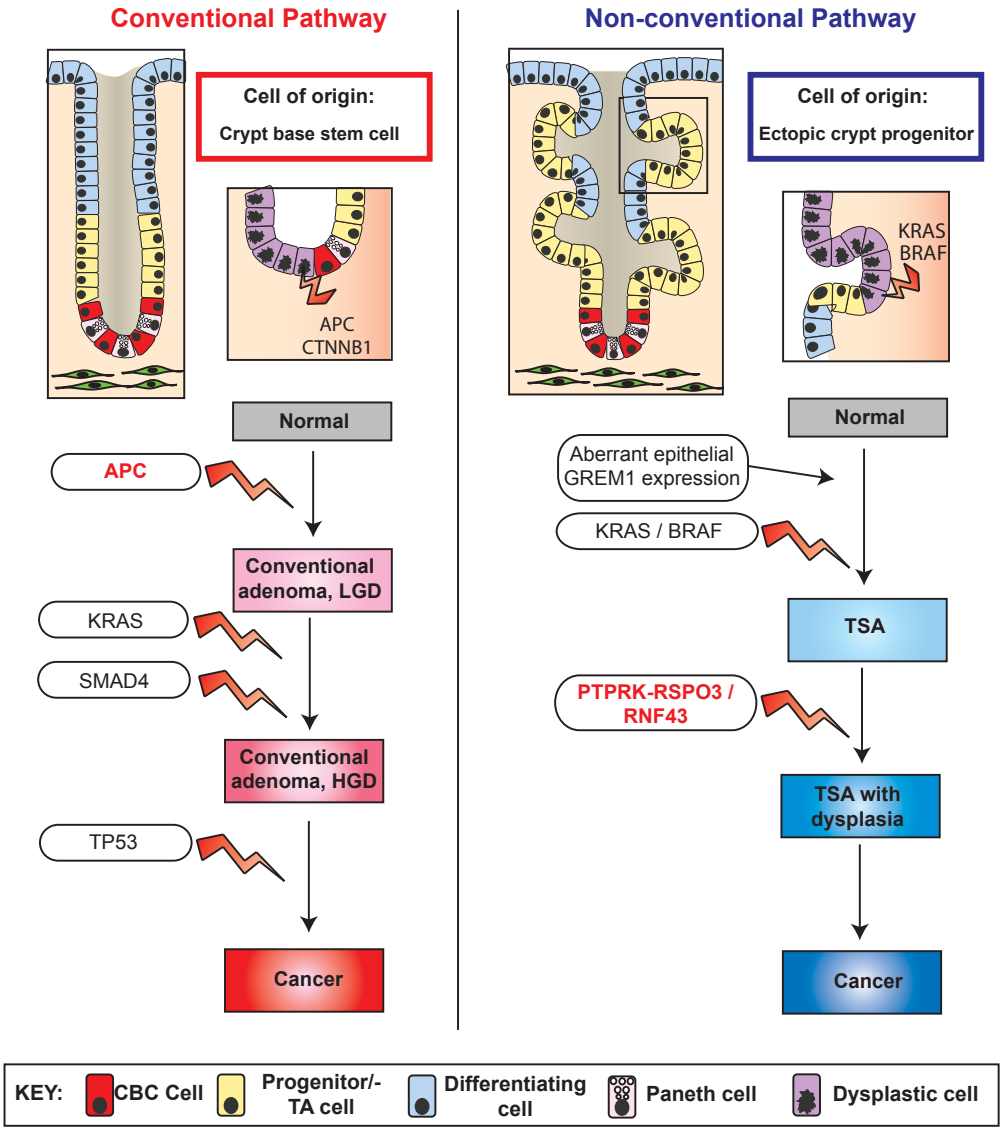


Figure 1.8: Model depicting 'conventional' and 'non-conventional' polyp pathways. Red text indicates activting Wnt pathway mutations; LGD: low grade dysplasia, HGD: high grade dysplasia

thus required to induce 'conventional' adenoma dysplasia, whereas different Wnt disrupting mutations provide a more graduated, cumulative selective advantage in 'non-conventional', serrated lesions that originate in a non-*LGR*5+ cell of origin (illustrated in Figure 1.8). Given that an emerging cancer frequently outgrows and destroys the initial polyp, understanding molecular pathogenesis in this way can thus shed light on the earliest stages of CRC evolution.

1.5.4 Molecular characterisation of human CRC

Advances in sequencing technology have enabled a rapid broadening of our understanding of the genetic, phenotypic and transcriptomic factors that underpin disease pathogenesis, given that thousands of somatic mutations can now be identified within a single cancer sample. In 2014, this expansion in scale enabled Alexandrov *et al* to carry out mutational signature analysis of more than 7000 primary cancers, thus revealing the vast diversity of the mutational processes active during carcinogenesis [125]. In 2015, an international consortium analysed sample data from over 4000 CRC patients, resulting in the stratification of CRC into four consensus molecular subtypes [126]. Despite contributing enormously to our understanding of CRC, the generation of these subtypes therefore also underlined the vast heterogeneity present within established cancers. However, the increased availability of high-depth next-generation sequencing technology has also enabled analysis of normal, non-dysplastic tissues. In the same year, Martincorena *et al* carried out high depth sequencing of skin cancer genes in biopsies of normal skin to reveal that positively selected driver mutations were found in a remarkable 18 - 32% of normal skin cells [127] (Figure 1.9). This translated to a greater number of *NOTCH1* mutations in 5 cm^2 of aged, sun-exposed skin than in more than 5000 sequenced cancers within The Cancer Genome Atlas. This seminal work, since recapitulated in normal oesophagus and colon, thus shed light on the earliest stages of cancer development and resulted in this memorable riposte from the authors: "Studying tumour evolution by sequencing established cancers is akin to inferring the rules of a musical talent quest by identifying similarities across the show's annual winners. Successful aspirants undoubtedly have common properties that identify necessary criteria for victory, but there is no substitute for directly observing the competition in its raw, early, local heats."

1.5.5 The genotype-phenotype axis in intestinal disease

It is therefore clear that neoplastic transformation of normal intestine to CRC involves a vast, complex network of interacting factors that influence disease

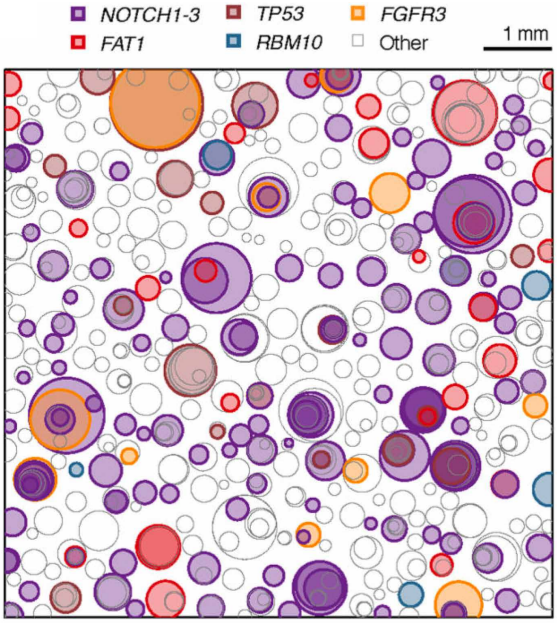


Figure 1.9: Schematic representation of mutant clones detected by Martincorena *et al* in an average 1 cm^2 of normal eyelid skin. Adapted from [127].

progression to varying degrees in a context-dependent manner. Our lab has therefore proposed a model of intestinal disease pathogenesis based on the evolutionary maxim that phenotype is a product of both genotype and environment (Figure 1.10). Within this model, genotype is divided into genetic predisposition and somatic mutation, and environment is divided into cell-extrinsic factors, such as the signalling environment, and cell-intrinsic factors, such as the cell of origin [128]. The resultant disease phenotype is therefore determined by variable and interacting contributions from each factor, the relative importance of which will vary according to disease. The inherent heterogeneity present within CRC is reflected in the endlessly variable interactions between each factor. As an example, the FAP disease phenotype is produced by a germline APC mutation that drives somatic mutation accumulation, resulting in tumourigenesis without the need for cell-intrinsic or extrinsic factors. Conversely, HMPS lesions (and TSAs) are driven by a disrupted cell-extrinsic environment that results in tumourigenesis originating from outside of the crypt base. Given that somatic mutation is but one variable influence upon disease landscape, I hypothesised that the impact of mutation on disease phenotype and outcome will vary according to

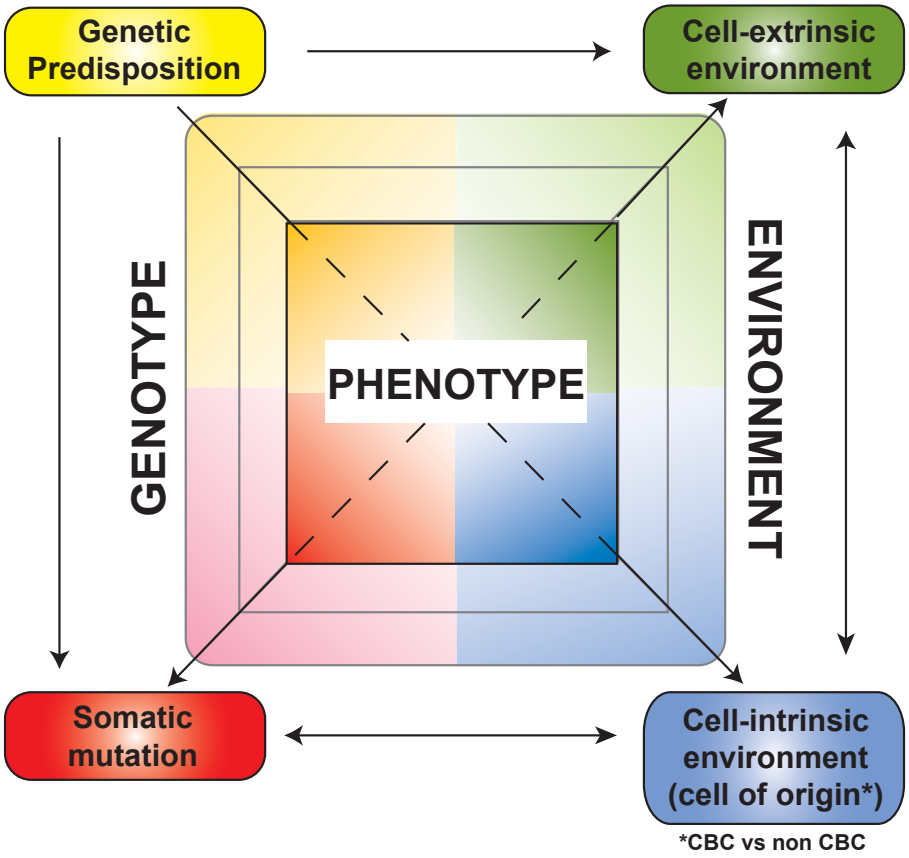


Figure 1.10: Model depicting the proposed genotype-phenotype axis in intestinal disease

epithelial cell-intrinsic and cell-extrinsic factors. The work presented in this thesis thus explored somatic mutation acquisition and the effect of microenvironmental influence using human and animal models of colorectal tumourigenesis.

1.6 Outline of thesis

– **Somatic mutation is detectable in non-dysplastic gut and will vary according to disease state**

In this section I investigated the acquisition of somatic mutation in normal tissue and used animal models to assess whether mutation impact varied according to disease state-induced microenvironmental differences. I first sought to assess CRC driver gene mutations within healthy and inflamed human colon using a novel sequencing method. As a further proof of principle experiment, I recapitulated varying mouse models of intestinal disease in order to investigate

the impact of disease state upon somatic mutation acquisition following uniform exposure to the mutagen AOM.

- **The mechanism of somatic mutation acquisition within a disrupted morphogen signalling context**

Earlier in this chapter I described the various homeostatic mechanisms that protect against mutation accumulation within the intestine, including intact signalling and DDR pathways and the spatial organisation of the intestinal stem cell niche. Our lab has generated a mouse model of disrupted intestinal homeostasis in which aberrant epithelial expression of a key signalling pathway component results in the formation of ectopic crypts. These crypts harbour progenitor cells that are actively proliferating outside of the protective environment of the crypt base. Given previous evidence linking key intestinal signalling pathways to DDR, such as TGF-β and IL-22 signalling, work presented in this section explored my hypothesis that aberrant epithelial *Grem1* expression in VG mice impacts upon intestinal DDR pathways via a novel mechanism linking BMP signalling and DNA repair.

- **Therapy-driven restoration of intestinal cell fate determination will re-instate homeostatic protection against mutation**

Work presented in my first two results chapters centres upon the investigation of somatic mutation during early stages of disease pathogenesis and within a disrupted morphogen signalling context. To date, however, there has been very little research into the impact of restoration of cell fate determination on somatic mutation acquisition. Given the well-documented impact of dysregulated signalling on mutation acquisition, I hypothesised that cells harbouring mutations acquired as a consequence of inflammation or intestinal regeneration would either be 'arrested' by therapy and so remain, in a 'dormant' state, within treated tissue or instead 'shed' through therapy-driven remodelling of dysplastic epithelium. Given that the outcomes of both hypotheses are detectable, as either a maintenance or reduction in somatic mutation burden, work presented in the final two chapters

explored the impact of therapeutic restoration of cell fate determination upon somatic mutation in mouse models of intestinal disease.

2
Materials and Methods

Contents

2.1 Human Samples . . . 39
2.1.1 Consent and ethics for human samples . . . 39
2.1.2 Single and pooled crypt dissection of fresh biopsies . . . 39
2.2 Mouse procedures . . . 40
2.2.1 Mouse strains . . . 40
2.2.2 Ear snips . . . 41
2.2.3 Breeding . . . 41
2.2.4 Administration of substances . . . 41
2.2.4.1 Azoxymethane . . . 41
2.2.4.2 Dextran Sodium Sulfate . . . 41
2.2.4.3 *Helicobacter hepaticus* . . . 42
2.2.4.4 *In vivo* antibody treatment . . . 42
2.2.5 Irradiation . . . 42
2.2.6 Gut preparation . . . 42
2.2.6.1 Collection of small intestine crypt and villi . . . 43
2.2.6.2 Collection of colon samples . . . 43
2.3 Histology . . . 43
2.3.1 Adenoma scoring . . . 43
2.3.2 Processing . . . 44
2.3.3 Embedding . . . 44
2.3.4 Sectioning and staining . . . 44
2.3.4.1 Haematoxylin and Eosin (H and E) staining . . . 44
2.3.5 Pathology scoring . . . 45
2.4 DNA . . . 45
2.4.1 DNA extraction . . . 45
2.4.1.1 DNA extraction from tissue . . . 45
2.4.1.2 DNA extraction from blood . . . 46
2.4.2 Nucleic acid quality control . . . 46

2.4.2.1 Nanodrop spectrophotometer 46
2.4.2.2 Qubit fluorometer 46
2.4.3 DNA sequencing 47
2.4.3.1 Molecular Inversion Probe Sequencing 47
2.4.3.2 Whole Exome Sequencing 51
2.4.3.3 Pre-processing 51
2.4.3.4 Merging multiplex sequences 51
2.4.3.5 Single Nucleotide Variant and Indel calling . . 52
2.4.3.6 SNV/Indel post-processing 53
2.4.3.7 Driver mutation identification 54
2.4.3.8 Mutational signature analysis 54
2.5 RNA / cDNA 55
2.5.1 RNA extraction 55
2.5.2 RNA sequencing 56
2.5.3 Reverse transcription of RNA into cDNA 56
2.5.3.1 Pre-amplification of cDNA 56
2.5.4 Quantitative real-time PCR 57
2.5.4.1 Fluidigm 57
2.5.4.2 TaqMan micro-RNA assay 57
2.5.5 In-situ micro-RNA hybridisation 58
2.6 Protein 59
2.6.1 Immunohistochemistry 59
2.6.2 Digital pathology 60
2.6.2.1 HALO Image Analysis 60
2.6.2.2 Aperio ImageScope 61
2.6.2.3 QuPath 61

2.1 Human Samples

2.1.1 Consent and ethics for human samples

All samples were collected in the endoscopy department at the Oxford Radcliffe NHS Foundation Trust. Patients were consented using the GI illness biobank or IBD biobank. Ethical approval for the collection and use of samples had been sought for these biobanks from independent ethical review committees (REC 11/YH/0020 for GI biobank and REC 09/H1204/30 for IBD biobank cohort).

2.1.2 Single and pooled crypt dissection of fresh biopsies

Prospective tissue collection was carried out on patients undergoing routine colonoscopy or flexible sigmoidoscopy. Following informed consent, three

pairwise 5 mm forceps biopsies from the ascending, transverse and descending colon were taken from patients deemed healthy by a clear scope with no polyps. From patients exhibiting active inflammation as symptomatic of UC, PSC, Crohn's Disease or IBD, a single pairwise 5 mm forceps biopsy was taken from the site of inflammation only. Blood samples were taken from all patients for assessment of germline mutations. Fresh biopsies were placed in 5 ml dissociation media (30 mM EDTA in DMEM without Ca^{2+} and Mg^{2+}, 0.5 mM DTT, 2% RNAlater (Qiagen)). The digested tissue was then transferred to PBS and shaken vigorously for 10 - 30 seconds to release individual crypts and villi. 5 single crypts per healthy pair of biopsies and 10 single crypts per inflamed pair of biopsies were individually selected under a dissection light microscope (2x) within 60 minutes of collection, with the remaining crypts collected as pooled samples per pair of biopsies. Crypt samples were then transferred to buffer ATL containing 10% (v/v) proteinase K and digested overnight at 56°C. DNA was isolated using the QIAamp DNA Micro kit (Qiagen, as per manufacturer's instructions), quantified using NanoDrop-100 spectrophotometer quantification (Thermo Fisher Scientific) and stored at -20°C.

2.2 Mouse procedures

All procedures were carried out in accordance with Home Office UK regulations and the Animals Scientific Procedures Act 1986 (Project Licenses PB802BC1E; P0B63BC4D). All C57BL/6J mice were housed in individually ventilated cages at the Functional Genomics Facility, Wellcome Centre for Human Genetics, University of Oxford. 129SvEv.*Rag*$^{-/-}$ were housed in individually ventilated cages in an accredited specific pathogen-free facility at the Kennedy Institute of Rheumatology, University of Oxford. My personal license number is I89FB8ACF.

2.2.1 Mouse strains

WT and *Vil1-Grem1* mice used in this study were maintained on C57BL/6J background for ≥6 generations: *Vil1-Grem1* mice were established in our lab and have been described previously [51]; 129SvEv.*Rag*$^{-/-}$ mice were established in

the Powrie lab, routinely screened for the absence of pathogens and *Helicobacter spp* and have been described previously [57].

2.2.2 Ear snips

At three weeks of age, post-weaning, ear punches were taken from mice both for identification purposes and to provide material for genotyping. An identification system whereby different numbers of ear punches on either or both ears was carried out, to ensure each mouse per cage had a different identification mark. DNA was extracted from these ear snips (section 2.4.1) and used for genotyping.

2.2.3 Breeding

Breeding pairs were put together from the age of sexual maturity (6 weeks) onwards. Trios with two females to one stud male were also used to enhance the number of offspring per generation of breeding as required.

2.2.4 Administration of substances

2.2.4.1 Azoxymethane

Administration of AOM was carried out according to the protocol described by Neufert *et al* [53]. Briefly, C57BL/6J wild type, *Vil1-Grem1* or 129SvEv.*Rag*$^{-/-}$ mice were weighed and given a single intraperitoneal injection of azoxymethane (AOM; Sigma; 10 mg/kg) or vehicle (saline) once a week for the treatment period and then maintained until the experimental endpoint.

2.2.4.2 Dextran Sodium Sulfate

Animals were given 2% Dextran Sodium Sulfate (DSS; 40 kDa, MP Biochemical) dissolved in distilled water for three cycles of 5 days each, separated by a 2-week recovery period where normal drinking water was given. Weights were recorded daily and mice were culled at the experimental endpoint. The threshold for immediate sacrifice was loss of ≥20% body weight.

2.2.4.3 *Helicobacter hepaticus*

Helicobacter hepaticus NCI-Frederick isolate 1A (*Hh*; strain 51449) was grown as described previously [129]. 129SvEv.*Rag*$^{-/-}$ mice were fed with *Hh* 1A (1.0 × 108 CFU) by oral gavage for three consecutive days and injected with AOM (Sigma-Aldrich) intraperitoneally once a week at a dose of 10 mg/kg for 5 weeks, starting 6 weeks after *Hh* infection [130].

2.2.4.4 *In vivo* antibody treatment

Therapeutic administration of antibodies was started at week 12 of *Hh*+AOM treatment in 129SvEv.*Rag*$^{-/-}$ mice, when inflammation was already fully developed, and was continued until the end of the experiment at week 20. To specifically block the cytokine IL-22 *in vivo*, a blocking mAb against IL-22 (Genentech, clone 8E11) or isotype control (GP120 10E7.1D2) was administered intraperitoneally three times per week at 0.15 mg. Treatment groups were mixed in cages to minimize cage effects. To block the BMP antagonist Grem1 in *Vil1-Grem1* mice, an anti-Grem1 antibody (UCB Pharmaceuticals) or isotype control (mIgG1) antibody was administered subcutaneously once a week for 9 weeks at a dose of 30 mg/kg, initiated at approximately 3 months of age.

2.2.5 Irradiation

C57BL/6J wild type or *Vil1-Grem1* mice underwent a single whole-body dose of 10 Gy irradiation, performed using an IBL-637 irradiator with a Caesium-137 source. Irradiation was delivered over 652 seconds at 0.92 Gy min^{-1}. Animals were regularly monitored until collection at the appropriate experimental timepoint (2, 5, 8 or 24 hours) following irradiation.

2.2.6 Gut preparation

Upon observation of symptoms of intestinal polyps (anaemia, hunching) or at a relevant experimental timepoint, mice were killed by cervical dislocation. The intestinal tract was removed immediately and divided in to three regions of

small intestine (proximal/SB1, middle/SB2 and distal/SB3) and the whole colon. Sections were opened longitudinally using a gut preparation apparatus described by Rudling *et al* [131], washed in PBS and fixed overnight in 10 % neutral buffered formalin (NBF) before being transferred to 70% ethanol for temporary storage prior to polyp counting (section 2.2.1) or processing for embedding (section 2.2.2).

2.2.6.1 Collection of small intestine crypt and villi

Respective SB1, SB2 and SB3 intestinal regions were washed with PBS to remove faecal matter before being gently scraped with a coverslip to collect the villi. Following three gentle repeats, the villi tissue was placed into 2 ml RNAlater (Qiagen) and the remaining tissue region placed in 5 ml dissociation media (30 mM EDTA in DMEM without Ca^{2+} and Mg^{2+}, 0.5 mM DTT, 2% RNAlater (Qiagen). The digested tissue was then transferred to PBS and shaken vigorously for 10 - 30 seconds to release individual crypts. The intestinal tissue region was then discarded and the remaining PBS spun down to collect the dissociated crypts from that region, which were then placed into 2 ml RNAlater (Qiagen) prior to processing for RNA extraction.

2.2.6.2 Collection of colon samples

Histologically normal colon samples were collected via 3 mm punch biopsies (PFM Medical). Polyps were identified by eye and carefully collected using a scalpel to avoid sampling of the underlying stromal and muscularis tissue. Colonic normal and polyp samples were immediately snap frozen in liquid nitrogen prior to being stored at -80°C.

2.3 Histology

2.3.1 Adenoma scoring

After fixation but before processing, gut preparations were stained with 0.2% methylene blue for 10 s and washed in PBS for 20 min. Small intestine or colon adenomas were counted with a dissecting microscope at 3x magnification and

categorised according to size. Care was taken to avoid mistaking Peyer's patches (aggregate lymphoid follicles) as tumours.

2.3.2 Processing

Prior to processing, fixed gut sections were rolled into swiss rolls, pinned together and placed in a histology cassette. Specimens were processed using a Histomaster machine (Bavimed). Gut rolls were processed with the following protocol: 70% ethanol for 45 min, 95% ethanol for 30 min, 95% ethanol for 45 min, 100% ethanol for 30 min, 100% ethanol for 45 min, 100% ethanol for 45 min, Histoclear for 30 min, Histoclear for 30 min, Histoclear for 30 min, wax for 30 min, wax for 45 min, wax for 45 min.

2.3.3 Embedding

Processed samples were embedded in paraffin wax using a paraffin embedding station (EG1150H, Leica).

2.3.4 Sectioning and staining

Tissues embedded in paraffin blocks were sectioned as 4 μm slices using a microtome (RM2255, Leica).

2.3.4.1 Haematoxylin and Eosin (H and E) staining

Formalin-fixed, paraffin embedded tissue sections (4 μm) were dewaxed in xylene for 5 min and rehydrated through graded ethanols (100%, 90%, 70%, 5 min each) followed by 2 min in tap H_2O. Sections were then stained with Harris Haematoxylin (30 sec) and washed in running H_2O until it ran clear. Slides were then briefly dipped in acid alcohol and left in H_2O for 5 min prior to being stained with Eosin (3 min), quickly dipped in H_2O, dehydrated through degraded alcohols for 2-5 min each and mounted with coverslips using DPX (ThermoFisher).

2.3.5 Pathology scoring

Pathology scoring was carried out by Adj Assoc Prof Wang Lai Mun, a senior consultant pathologist within the Department of Laboratory Medicine at Changi General Hospital, Singapore. H and E stained intestinal sections were blinded prior to analysis of technical duplicates. Intestinal sections were scored for crypt architecture, degree of inflammatory cell infiltration and muscle thickening (0 = absent, 1 = mild, 2 = moderate, 3 = severe); goblet cell depletion and presence of crypt abscesses (0 = absent, 1 = present). The number of cancers, number of dysplastic lesions, percentage of dysplasia and dysplasia grade (0 = absent, 1 = low grade, 2 = high grade) were also assessed. The average of both technical repeats per metric was used for subsequent analysis. Composite pathology scores were calculated as the sum of all pathology scores, excluding the number of cancers, dysplastic lesions and percentage dysplasia, and averaged across technical replicates prior to use in downstream analyses.

2.4 DNA

2.4.1 DNA extraction

2.4.1.1 DNA extraction from tissue

Samples were incubated overnight at 56°C in 200 μl buffer ATL containing 10% (v/v) proteinase K. DNA was isolated using the QIAamp DNA Micro kit (Qiagen) according to manufacturer's instructions. Briefly, upon complete digestion of the tissue, 200 μl Buffer AL was added for sample lysis by pulse-vortexing. 100% ethanol was added prior to the entire sample being passed through a QIAamp MinElute column to bind any preciptated DNA to the membrane. Following consecutive washing steps with Buffers AW1 and AW2, the membrane was dried by centrifugation prior to elution of bound DNA with 50 μl TE buffer. For DNA extraction from tissues weighing ≥10 mg, the QIAamp DNA Mini kit (Qiagen) was used as above, with the addition of a tissue homogenisation step where up to 25 mg tissue in 80 μl PBS was homogenised with a TissueRupter prior to

addition of 100 μl Buffer ATL and 20 μl Proteinase K, with all other volumes and buffers the same and elution volume adjusted to 200 μl.

2.4.1.2 DNA extraction from blood

Up to 200 μl blood was added to 20 μl QIAGEN Protease prior to the addition of 200 μl Buffer AL for sample lysis by pulse-vortexing. Samples were incubated at 56°C for 10 min prior to processing with the QIAamp DNA Blood and Mini kit according to manufacturer's instructions, described briefly in 2.4.1.1. DNA was eluted in 100 μl TE buffer prior to storage at -20°C.

2.4.2 Nucleic acid quality control

2.4.2.1 Nanodrop spectrophotometer

Determinaton of nucleic acid quality was performed using a micro-volume UV-visible spectrophotometer (Nanodrop One, ThermoScientific). In brief, 1 μl sample was loaded onto an optical pedestal and the absorbed light from the sample measured at a wavelength of 260 nm. The ratio of readings at 260 nm and 280 nm provided an estimate of the purity of DNA (target $\simeq$ 1.8) and RNA (target $\simeq$ 2.0).

2.4.2.2 Qubit fluorometer

Determination of nucleic acid concentration was performed using the Qubit dsDNA Broad Range Assay kit with a Qubit 2.0 Fluorometer according to manufacturer's instructions. Briefly, 1 μl sample was added to 199 μl Qubit working solution (Qubit dsDNA HS reagent diluted 1:200 in Qubit dsDNA HS buffer), mixed by vortexing and incubated at room temperature for 2 min. Sample concentration was then determined by a curve-fitting algorithm based on the relationship between Qubit standards #1 and #2 used in calibration of the fluorometer.

2.4.3 DNA sequencing

2.4.3.1 Molecular Inversion Probe Sequencing

MIP sequencing facilitates highly accurate base calling, with substitution error rates below 1 in 10000, in order to achieve sensitive and precise detection of low-frequency somatic mutation [132]. Primer design, sample processing and sequencing were carried out by Dr Claire Palles and Laura Chegwidden of the Palles Group (Institute of Cancer and Genomic Sciences, University of Birmingham). Primers were designed against the coding regions of genes of interest, with each individual primer containing a unique single molecule tag. Sequencing reads originating from the same primer therefore shared a molecular tag and were able to be grouped together. Variants were then called on a molecule by molecule basis.

Primer panels We utilised a panel of MIP primers targeting genes that are frequently mutated in CRC and exhibit signals of positive selection. TCGA stratification of CRC revealed two classes, hypermutated (≥675 exome somatic mutations, 16% of CRC) and non-hypermutated (84% of CRC) [133]. Each class harbours different driver genes and were considered separately for primer design. The panel was sequenced at high depth (x1000) and contained approximately 2000 MIPs against 30 driver genes of hypermutated and non-hypermutated CRC (Table 2.1), thus corresponding to tumours with an exceptional burden of single nucleotide variants (SNVs) and a strong tendency to microsatellite stability.

Preparation of probes for capture Aliquots of each probe were pooled at equimolar ratios and phosphorylated using 50 units of T4 Polynucleotide Kinase (NEB) and 1 x T4 DNA ligase buffer for 45 min at 37°C, followed by 20 min at 80°C to inactivate the kinase. Captures were performed as described previously [134]. Briefly, a total of up to 100 ng of DNA, 330 femtomoles of probe mixture, and 1 μl of Ampligase DNA ligase buffer (Epicentre) were added to dH_2O for a total of 10 μl. For the probe hybridization phase, these mixtures were incubated in a

thermocycler (Bio-Rad) with a heated lid at 98°C for 3 min, 85°C for 30 min, 60°C for 60 min, and 56°C for 120 min. For the gap-fill and ligation phase, 300 picomoles each dNTPs (NEB), 7.5 micromoles betaine (Sigma), 20 nanomoles NAD+ (NEB), 1 μl of 10X Ampligase buffer, 5 units of Ampligase DNA ligase (Epicentre), 3.2 units of Phusion DNA polymerase (NEB) were added to up to 10 μl dH_2O for a total reaction volume of 20 μl. The gap-fill and ligation phase was carried out at 56°C for 60 min and 72°C for 20 min. Reactions were then cooled to 37°C, and 20 units of Exonuclease I (NEB) and 100 units of Exonuclease III (NEB) added to each reaction to degrade uncircularized probe and genomic DNA. The digestion was incubated at 37°C for 45 min, heated to 80°C, and incubated for 20 min to inactivate the exonucleases. Samples were then cooled on ice and, optionally, stored at -20°C.

Captures and library construction For each capture reaction, two PCR reactions were prepared, each with Phusion HF buffer (Fermentas), forward primer and indexed reverse PCR primers to 500 nM, SYBR green (Invitrogen) to 0.5X, dNTPs to 200 mm each (NEB), 2 units of Phusion Hot-Start II polymerase, 10 μl of capture reaction, and nuclease-free water to 50 μl. PCR cycling conditions were an initial denaturation step for 2 min at 95°C, followed by 26 cycles of: 15 sec at 98°C, 15 sec at 65°C, and 45 sec at 72°C.

Library purification and pooling PCR products were purified individually using Ampure XP beads (Agencourt) according to the manufacturer's instructions. Purified PCR products were then pooled naively (i.e., equal volumes) for initial quality control. To remove a low molecular-weight artifact, the PCR product pool was split across 8 wells of a 10-well pre-cast 6% polyacrylamide TBE gel (Invitrogen), run at 140 V for 50 min, and stained with 5 μl of SYBR Gold (Invitrogen). The capture product band at 280 bp was excised, crushed, soaked in 800 μl of Tris-EDTA (pH 8.0), and recovered from the supernatant using 100 μl of Ampure XP beads and 700 μl of home-made Ampure buffer (20% PEG 8000, 2.5M NaCl) according to the manufacturer's instructions. Pool concentration was assessed using Qubit (Invitrogen).

Sequencing and primary analysis Samples were sequenced using the HiSeq 2000 (Illumina) platform according to the manufacturer's instructions. Two 101 nt reads were collected to determine the sequence of the gap-fill and the molecular tag and one 8 nt read to determine the sequence of the sample index. Read-pairs were assigned to samples requiring an exact match to the expected 8 nt sample index sequence and the first 12 nt of the reverse read (corresponding to the molecular tag sequence) were stripped out and placed in the header. Read-pairs with molecular tags with homopolymers longer than 4 nt were discarded. Overlapping regions of read-pairs were then reconciled to form single "fr-reads" using a custom Smith-Waterman-based strategy. For positions where the read-pairs did not overlap, quality scores from the individual reads were retained. For positions where the read-pairs did overlap, quality scores for the resulting consensus calls were estimated as below for smc- reads. Only successfully overlapped fr-reads were retained for downstream analysis. Fr-reads were aligned to the human reference genome (hg19) using the bwasw alignment mode of the aligner bwa [135](v0.5.9) with non-default parameters "-r 1". Based on expected alignment positions according to the probe design, fr-reads were then assigned to individual probes, allowing 1 nt of tolerance in each direction for the beginning of the read, which primarily acommodates insertion and deletion mutations during probe synthesis. Then, for each sample and each probe, fr-reads were grouped by molecular tag sequence to form tag-defined read groups (TDRGs).

Single molecule consensus read calling Alignments to the reference genome were used to call a consensus sequence (i.e. a single molecule consensus read or "smc-read") for each TDRG. Positions expected to be derived from probe targeting arms (and therefore synthetic DNA) were excluded from consideration at this step. A likelihood ratio framework was adopted to incorporate both the abundance and associated quality-scores of fr-read base-calls supporting each possible nucleotide at each position. Briefly, we calculated the log-likelihood L_x of consensus nucleotide x as the difference of the log-likelihood of a model in which

Non-hypermutated CRC	Hypermutated CRC	Both	Custom
APC	ACVR2A	KRAS	GNAS codon 200
CTNNB1	ARID1A	PIK3CA	MLH1
ELF3	B2M	RNF43	MSH2
FBXW7	BCL9L	TP53	MSH6
NRAS	BMPR2		POLD1 (ED)
SMAD4	BRAF		POLE (ED)
SOX9	PTEN		SMAD2
TCF7L2	RPL22		ATM
TGIF1			
ZFP36L2			

Table 2.1: Genes targeted by MIP sequencing primers

a given nucleotide x was underlying none of the calls and the log-likelihood of a model in which x was underlying all of the calls. This can be represented as in the equation below, where a given observation is associated with a nucleotide call o and a "*phred*-like" quality score Q_o.

$$L_x = \sum_{(o|o\neq x)\in calls} \left(\log_{10}\left(1 - \frac{10^{\frac{-Q_o}{10}}}{3}\right) - \log_{10}\left(\frac{10^{\frac{-Q_o}{10}}}{3}\right)\right) + \sum_{(o|o\neq x)\in calls} \left(\log_{10}\left(10^{\frac{-Q_o}{10}}\right) - \log_{10}\left(1 - 10^{\frac{-Q_o}{10}}\right)\right)$$

This value (L_x) was computed for each possible consensus call (A, C, G, T, N, and deletion) at each position in the alignment based on the set of base-calls and associated quality scores in the TDRG at that position in the alignment. The consensus call was then determined as the call with the minimum (i.e. most negative) log-likelihood value, as this indicates the consensus nucleotide where the model assuming that nucleotide did not underlie any of the fr- readcalls was the least likely relative to the model assuming that nucleotide underlay all of the fr- read calls. The final *phred*-like quality score is calculated as the integer casting of 10^*L_x.

2.4.3.2 Whole Exome Sequencing

Library preparation (SureSelect, Agilent and Nextera, Illumina), DNA-sequencing (HiSeq4000, Illumina) and mapping (BWA) were performed by the High-Throughput Genomics Group (Wellcome Centre for Human Genetics, Oxford).

2.4.3.3 Pre-processing

Binary alignment/map (BAM) files were prepared from the fastq generated files provided by the sequencing facility. These were paired-end reads from multiplexed samples on the Illumina platform; the provided BAM files were mapped to the mm10 reference reference build (with the chr prefix). For purposes of reliability and reproducibility (with extended versioning and release/patch information), we only utilized the GRCm38 ENSEMBL reference genome. The first stage of BAM processing involved mapping the reads using BWA (Burrows-Wheeler transformation alignments) with the BWA-MEM algorithm (maximal exact matches) from the fastq_1 and _2 files; this algorithm exhibits higher performance for 70-100 bp reads generated by the Illumina platform [135]. Read groups (tags to identify information within each sample) were assigned as given by the sequencing core. Once alignments were completed, a Sequence alignment/map (SAM) file for each of the multiplexed sequencing runs (multiple runs aligned separately from the same sample) was generated. Due to the storage requirements of a large number of SAM files, the output from the alignment program was piped to to samtools [136] to perform an on the fly SAM to BAM conversion while maintaining the header information. The BAM files were then sorted by coordinate using GATK (v4.0.2.0) [137].

2.4.3.4 Merging multiplex sequences

GATK best practices [138] were followed, as modified by the latest GATK toolset (namely local rearrangements around haplotypes as a pre-processing step) (v4.0.2.0). First, the multiplexed sequences were merged in their respective BAM files while marking duplicate reads simultaneously. Marking duplicates

flags reads with the same start position to reduce error introduced by PCR amplification steps during the library prepartion process. This allows for more robust statistical testing depending on the downstream applications. For variant calling using MuTect2 this is a required step. Marking duplicates is performed for each individual library (thus the importance of ensuring adequete read group information on each BAM). This is accomplished while preserving the sort order of the BAM, maintaining a compressed output BAM, and generating a BAM index file. This process results in a marked duplicate and merged multiplexed sample BAM. Once multiplexed BAM files were merged into a single sample BAM file, base quality score recalibration (BQSR) tables were generated to apply to the BAM. This process utilizes commonly observed sites for both SNVs and small insertions and deletions (indels) along with global base quality metrics to recalibrate base quality scores for downstream error models. Once the BQSR tables are generated, base recalibrations were then applied to the BAM files.

2.4.3.5 Single Nucleotide Variant and Indel calling

There are a number of useful single nucleotide variant (SNV) calling tools, but the underlying type of algorithm and the implementation of that algorithm can vary. Haplotype analysis is a method suitable for calling both insertions and deletions (indels) and SNVs. A haplotype is simply the contiguous genes or alleles that are inherited together and are generally on the same chromosomal segment. The reason this method is commonly used for indel calling is that it uses local de novo re-alignments around a region of interest after reads have been mapped to a reference genome; likewise, this is also useful in determing whether a variant exists within a particular position. There are a number of variant calling tools that employ this method, one of the most widely used is MuTect (now MuTect2) from the Broad Institute. MuTect2 is similar to HaplotypeCaller, but it does not make the assumption that the sample is diploid, allowing for CNVs to be present. The PON (panel of normal) VCF (variant call format) file for variant calling in GATK4 helps to reduce false positives by filtering out

systemic errors caused by sequencing and target capture methods. The number of 'normals' used in the PON file can vary, but the premise is that they need to be sequenced and prepared using the same method as test samples to be called. For our sample sets, because we are interested in 'normal' colon tissue variants, we used blood samples to inform the PON files.

2.4.3.6 SNV/Indel post-processing

A program 'Variant Filter' was created to perform variant filtering and gene set annotations dependant on a number of parameters and the variant caller used. This step is the most important when filtering for variants of interest as the callset of variants from most of these variant calling algorithms have false positives. We therefore sought to balance false positives with the detection of rare and low frequency mutations. The program calls and default settings for each argument are outlined in Table 2.3. The unique aspect of 'Variant Filter' is to allow for joint filtering if two samples from the same mouse are taken and exhibit the variant in some capacity. In these scenarios, the -uaf and -uvr options can be set, which allow for more lenient filtering if the exact same variant passes in at least one of the related samples. The program works as follows 1) Run through all variants independantly and assign flags, 2) if a joint filtering method is employed then run back through and adjust those flags at the shared mutation site, and 3) output a filtered VCF. If a user selects gene set annotations, snpEff or ANNOVAR [139] were used to label variants. Within snpEff annotations, only the canonical transcripts were used. The program then builds a final table using the -t flag, which extracts the tri-nucleotide context where that mutation occurs. This is useful when exploring signatures and looking for loss of DNA-repair and other signatures indicative of various mutational processes. The results table contains all samples of interest and all variants for further analysis and exploration in R.

Filter	Value
Minimum reads in normal	10
Minimum variant reads	2
Total reads at variant site sum	10
Minimum variant allele frequency	0.05
Minimum 'multisample' variant allele frequency	0.005
Maximum number of events allowed at a loci	1
Flanking region to omit clustered events	5
Reference genome	GRCm38

Table 2.2: 'Variant Filter' program arguments

2.4.3.7 Driver mutation identification

Driver mutations were identified using the 'Common Drivers in CRC' list of 113 genes, outlined in Table 2.3. The list combines 95 genes from the Intogen Colorectal Cancer (COREAD) Driver Gene list, identified from TCGA and ICGC analyses of COREAD data, with 18 additional signalling and mismatch repair genes of interest to the Leedham lab.

2.4.3.8 Mutational signature analysis

Mutation signatures, incorporating base substitutions and information on the sequence context of each mutation, were extracted using the R packages 'MutationalPatterns' [140] and 'deconstructSigs' [141]. Since there are six classes of base substitution, C.A, C.G, C.T, T.A, T.C, T.G (all substitutions are referred to by the pyrimidine of the mutated Watson-Crick base pair), and as we incorporated information on the bases immediately 5' and 3' to each mutated base, there are 96 possible mutations in this classification [142]. This 96-substitution classification is particularly useful for distinguishing mutational signatures that cause the same substitutions but in different sequence contexts. The contribution of each detected signature to total mutations was calculated by the whichSignatures function using the COSMIC database as a reference [143].

Common Drivers in CRC				
Aco1	*Chd9*	*Lgr4*	*Polb*	*Smad2*
Acsl6	*Clspn*	*Lgr5*	*Pold1*	*Smad4*
Acvr1b	*Cnot1*	*Lgr6*	*Pold2*	*Smc1a*
Akap9	*Crebbp*	*Map2k1*	*Pold3*	*Sos2*
Apc	*Ctcf*	*Map3k4*	*Pold4*	*Sox9*
Arid1a	*Ctnnb1*	*Mecom*	*Poldip2*	*Srgap3*
Arntl	*Cul1*	*Med12*	*Poldip3*	*Stag2*
Aspm	*Dis3*	*Med24*	*Pole*	*Syncrip*
Atm	*Dnmt3a*	*Mga*	*Pole2*	*Taf1*
Atrx	*Egfr*	*Msr1*	*Pole3*	*Tbx3*
Axin2	*Elf3*	*Nf1*	*Pole4*	*Tcf12*
Bcor	*Fbxw7*	*Notch1*	*Polr2b*	*Tcf7l2*
Bmpr2	*Fn1*	*Nr2f2*	*Ppp2r1a*	*Tgfbr2*
Bptf	*Foxp1*	*Nr4a2*	*Pten*	*Trio*
Braf	*Fxr1*	*Nras*	*Ptgs1*	*Trp53*
Brwd1	*Gata3*	*Ntn4*	*Ptpn11*	*Trp53bp1*
Cad	*Gm49384*	*Nup107*	*Ptpru*	*Wipf1*
Casp8	*Gnas*	*Nup98*	*Rad21*	*Wt1*
Cdc73	*Golga5*	*Pcbp1*	*Rbm10*	*Zc3h11a*
Cdk12	*Gsk3b*	*Pik3ca*	*Rnf43*	*Zc3h11a*
Cdkn1b	*Idh2*	*Pik3r1*	*Rtn4*	*Znrf3*
Cep290	*Itsn1*	*Pola1*	*Runx1*	
Chd4	*Kras*	*Pola2*	*Sf3b1*	

Table 2.3: List of genes within the 'Common Drivers in CRC' gene list

2.5 RNA / cDNA

2.5.1 RNA extraction

RNA was isolated from cells with the RNeasy Mini kit (Qiagen) according to manufacturer's instructions. Purified RNA was treated with DNase I (as per manufacturer's instructions; Thermo Fisher Scientific) to degrade residual DNA. RNA sample quality was measured with the NanoDrop-2000 spectrophotometer (ThermoFisher Scientific) as described in 2.4.2.1.

2.5.2 RNA sequencing

Library preparation, RNA-sequencing (HiSeq4000, Illumina), and sample Quality Control were performed by the High-Throughput Genomics Group (Wellcome Centre for Human Genetics, Oxford). Gene mapping was performed using R software (bowtie, mm10 (Version 10) and featureCounts).

2.5.3 Reverse transcription of RNA into cDNA

Complementary DNA was reverse transcribed *in vitro* from extracted RNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) as per manufacturer's instructions. Sample concentration and quality was measured using a NanoDrop-2000 spectrophotometer (Thermo Fisher Scientific). For miRNA expression analysis, miRNA-specific reverse transcription of cDNA was carried out using the TaqMan microRNA Reverse Transcription Kit (Applied Biosystems) according to manufacturers instructions.

2.5.3.1 Pre-amplification of cDNA

For high throughput Fluidigm gene expression analysis, mRNA-specific pre-amplification of cDNA was necessary prior to qRT-PCR. A pooled mix of TaqMan probes was generated by adding 2 μl of each 20X probe and adjusting to a final probe concentration of 0.2X μl with TE. Pre-amplification reaction mixes were prepared as follows: 1 μl Preamp Master Mix (Fluidigm), 1.25 μl pooled assay mix, 1.5 μl dH_2O and 1.25 μl cDNA for a total reaction volume of 5 μl. The pre-amplification reaction was run on a thermo-cycler with the following programme: initial enzyme activation step at 95°C for 2 min, followed by 14 cycles of 95°C for 15 s and 60°C for 4 min. Upon completion the reaction was immediately placed on ice. The pre-amplification product was diluted 1:5 in dH_2O for immediate use or stored at -20°C.

2.5.4 Quantitative real-time PCR

2.5.4.1 Fluidigm

High throughput absolute quantification qRT-PCR was performed using a Biomark HD (Fluidigm) cycler for each sample (in triplicate), and data were analysed using standard protocols to calculate expression as ΔCt [144] between experimental condition, normalised to endogenous controls. TaqMan Gene Expression Assays were used for each gene of interest and are shown in Table 2.4. These pre-prepared assays contain primer pairs and a probe that have already been tested for optimal PCR efficiency by Applied Biosystems. A reporter dye (FAM dye in all cases in this thesis) is linked to the 5' end of the probe and a non fluorescent quencher is at the 3' end of the probe. A typical sample inlet contained 2.7 μl of preamplified cDNA, 3 μl of TaqMan FAST Universal Mastermix and 0.3 μl 20X GE Sample Loading Reagent (Fluidigm, PN 100-7610). A typical assay inlet contained 3 μl Taqman probe and 3 μl 2X Assay Loading Reagent (Fluidigm, PN 100-7611) for total inlet volumes of 6 μl. Following priming of the Integrated Fluidic Circuit (IFC, Fluidigm) on the appropriate IFC Controller (Fluidigm, HX or RX), samples and assays were loaded into the appropriate inlets prior to running on the Biomark HD using the 'GE Fast v1.pcl' protocol.

2.5.4.2 TaqMan micro-RNA assay

Quantification of mmu-miR-96 (Taqman; 000186) expression was performed using a QuantStudio6 (Applied Biosystems) cycler for each sample (in triplicate) and data were analysed using standard protocols to calculate relative expression as ΔCt [144] between experimental condition, with U6, sno202 and sno234 (Taqman; 001973, 001232, 001234 respectively) serving as endogenous controls. A typical reaction contained 0.67 μl of miRNA-specific cDNA template described in 2.5.3, 0.5 μl of 20X TaqMan microRNA assay probe, 5 μl of TaqMan FAST Universal Mastermix and 3.84 μl dH_2O for a total reaction volume of 10 μl. Reactions were run on the QuantStudio6 with the following conditions: enzyme activation at

Gene	Assay ID
Atm	Mm01177457_m1
Actb	Mm02619580_g1
Bmpr1a	Mm00477650_m1
Bmpr2	Mm00432134_m1
Brca1	Mm00515386_m1
Brca2	Mm01218747_m1
Cdkn1a	Mm00432448_m1
Ercc1	Mm00468337_m1
Fen1	Mm01700195_m1
Foxo1	Mm00490671_m1
Foxo3	Mm01185722_m1
Gapdh	Mm99999915_g1
Lig1	Mm00495331_m1
Lig3	Mm01303107_m1
Mbd4	Mm01184338_m1
Mlh1	Mm00503449_m1
Mre11a	Mm00450600_m1
Msh2	Mm00500563_m1
Myc	Mm00487804_m1
Parp1	Mm01321084_m1
Polq	Mm00712819_m1
Prkdc	Mm01342967_m1
Rad51	Mm00487905_m1
Raf1	Mm00466513_m1
Smad7	Mm00484742_m1
Tgfbr1	Mm00436964_m1
Tgfbr2	Mm03024091_m1
Wrn	Mm01197916_m1
Xpc	Mm01183434_m1
Xrcc1	Mm00494222_m1
Xrcc5	Mm00550142_m1
Xrcc6	Mm00487458_m1

Table 2.4: TaqMan probes

95°C for 20 seconds followed by 40 cycles of denaturation at 95°C for 1 second and simultaneous annealing and extension at 60°C for 20 seconds.

2.5.5 In-situ micro-RNA hybridisation

In order to detect the expression of specific miRNA in FFPE mouse intestinal sections, the miRCURY LNA miRNA ISH Buffer and Controls kit (339459, QIAGEN) was used according to manufacturers instructions. Briefly, formalin-

fixed, paraffin embedded tissue sections (4 μm) were de-waxed in xylene for 3 x 5 min and rehydrated through graded alcohols (3 x 100%, 2 x 96%, 2 x 70%, 5 min each) followed by 2 min in PBS. miRNAs were then demasked using 0.2X Proteinase K solution for 10 min at 37°C and the sections washed twice in PBS. Hybridisation mix, containing 40 nm Mmu-mir-96-5p ISH probe (5p double dig labeled, YD00615907-BCG, QIAGEN) or 1 nm U6 LNA probe (5' DIG) as a positive control, was then applied and the sections incubated at 5°C for 1 h in a hybridisation oven (HybEZ, Advanced Cell Diagnostics). Slides were then washed in graded SSC buffer (5X, 2 x 1X, 2 x 0.2X, 5 min each) at 55°C followed by 0.2X SSC buffer for 5 min at room temperature, with all further steps carried out using a humidifying chamber. A blocking solution (PBS, 0.1% Tween, 2% sheep serum, 1% BSA) was then applied for 15 min at room temperature prior to slides being washed in PBS (3 x 3 min). To recognise the digoxigenins, sheep anti-DIG-AP antibody (11093274910, Sigma-Aldrich) was diluted 1:800 in blocking solution and applied to the slides for 60 min at room temperature. AP substrate solution was then applied at 30°C for 2 h, for conversion of the soluble substrates 4-nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro- 3'-indolylphosphate (BCIP) to an insoluble, dark blue precipitate. To stop the reaction, slides were incubated in KTBT buffer for 2 x 5 min prior to being washed twice in H_2O. Sections were then counterstained with 200 μl Nucelar Fast Red (H-3403, Vector labs) for 1 min, dehydrated through graded alcohols (2 x 70%, 2 x 96%, 2 x 100%, 2 min each) and counterstained using DPX (ThermoFisher).

2.6 Protein

2.6.1 Immunohistochemistry

Formalin-fixed, paraffin embedded tissue sections (4 μm) were de-waxed in xylene for 5 min and rehydrated through graded alcohols (100%, 90%, 70%, 3-5 min each) followed by 2 min in tap H_2O water. Endogenous peroxidase was blocked using 1.6% H_2O_2 for 20 min. Sections were pressure cooked in 10

Primary Antibody	Manufacturer	Dilution
Bag1 (Rabbit)	AB32109, Abcam	1:500
β-catenin (Mouse)	610154, BD	1:50
cleaved caspase-3 (Rabbit)	AF835, Abcam	1:800
Ki67 (Rabbit)	DB35, Cell Signaling	1:500
p-ATM (Ser1981) (Rabbit)	MA5-32751, Invitrogen	1:100
p-p53 (Ser15) (Rabbit)	12571S, Cell Signaling	1:500
Sox9 (Rabbit)	AB5355, Abcam	1:1000
Rad51 (Rabbit)	MA5-32789, Invitrogen	1:100
γH2AX (Ser139) (Rabbit)	CST2577, Cell Signaling	1:500

Table 2.5: Primary antibodies used for immunohistochemistry

mmol/L citrate buffer (pH 6.0) for 5 min, blocked with 10% serum (of secondary antibody species) for 30 min and incubated with primary antibody for at least 1 h (for antibody dilutions see Table 2.4). Slides were washed (5 min, PBS), incubated with secondary antibody (against primary antibody species) for 30 min and washed again prior to incubation with ABC (containing avidin and biotinylated horseradish peroxidase, Vector labs) for 30 min. Slides were then washed for a final time prior to application of DAB solution for 2-5 min, with development of the colour reaction monitored microscopically. Slides were counterstained with Haematoxylin, dehydrated, cleared and then mounted with a coverslip using DPX (ThermoFisher).

2.6.2 Digital pathology

2.6.2.1 HALO Image Analysis

Anaysis of IHC staining intensity using the HALO image analysis platform was carried out by Dr Viktor Koelzer. Villus and crypt regions were analysed separately using manual annotations to define the topographic areas and a machine learning step to separately analyse epithelial and stromal expression of the IHC marker. Briefly, all slides were first oriented so that the colon sample was at 11:00, the first roll of small intestine clockwise at 1:00, the second at 4:00, the third at 8:00. Within each of the 3 rolls of small intestine, cell counts from 4 separate areas of crypt and villus regions were then sampled and counted to generate six data categories

per case (Crypt 1:00, Crypt 4:00, Crypt 8:00; Villus 1:00, Villus 4:00, Villus 8:00). Where ectopic crypts were present, a minimum of 16 to a maximum of 78 ectopic crypt regions were sampled and the cell counts and expression levels averaged (data point 'ectopic crypts'). A machine learning step was then run in each region to separate epithelial and stromal expression. Following tissue segmentation, cell quantification was run within each compartment and region. Cell quantification was set to produce absolute cell counts on a four-point scale (negative, weakly positive, moderate, strong), the average optical density (OD) of the reaction product within each region and additional cell data such as nuclear size and cell area. I then carried out downstream analysis and visualisation using R.

2.6.2.2 Aperio ImageScope

For measurement of intestinal and crypt length, H&E stained slides were scanned at 20x magnification on a Leica Aperio CS2 and stored in Scanned Virtual Slide (.svs) format. Digital image analysis was performed using the Aperio ImageScope (v12.3.3) program. Briefly, for intestinal length the entire cross section of the relevant intestinal region was manually annotated and the length recorded. For crypt length, the ruler tool was used to measure at least 10 crypts per intestinal region per mouse. For colon crypt length analysis at least 30 crypts per colon were measured. Lengths were recorded and the average crypt length per mouse calculated.

2.6.2.3 QuPath

Proteins were detected by immunohistochemistry as described in 2.6.1. All slides were scanned at 40x magnification on a Leica Aperio CS2 and stored in Scanned Virtual Slide (.svs) format. Digital image analysis was performed on blinded slide files using QuPath (v.0.1.2) image analysis software. Briefly, for slides containing sections from the entire intestinal tract, an average of 10 individual crypts, villi or ectopic crypts were identified and manually annotated per intestinal region (SB1, SB2, SB3 and colon). For slides containing only colon

Parameter	Value
Detection image	Optical Density Sum
Requested pixel size	0.0 (full image resolution)
Background radius	20 μm
Median filter radius	0 μm
Sigma	1.5 μm
Minimum area	5 μm^2
Maximum area	150 μm^2
Threshold	0.15
Max background intensity	2.0
Split by shape	TRUE
Exclude DAB	FALSE
Cell expansion	4 μm
Include cell nucleus	TRUE
Smooth boundaries	TRUE
Make measurements	TRUE

Table 2.6: QuPath 'positive cell detection' algorithm settings

sections, an average of 30 crypts were annotated. For analysis of polyp regions, the entire epithelial cell area of the polyp was annotated, taking care to omit stromal content. Next, the 'positive cell detection' algorithm was used to identify marker-positive cells. The H&E nuclear stain was set as R=0.651, G=0.701, B=0.29, and the DAB positive stain as R=0.268, G=0.568, B=0.778. Algorithm settings are provided in Table 2.6. Total counts and percentage of marker-positive 1+, 2+ or 3+ cells per compartment were recorded and the average percentage of positive cells per compartment per mouse calculated.

"Choose your mutations carefully." - Mark Mothersbaugh

3

Underlying disease state impacts upon mutation acquisition

Contents

3.1 Introduction 64
3.1.1 Chapter aims 65
3.2 Evidence of selective pressures in normal and inflamed human colon 65
3.2.1 Mutation burden analysis 68
3.2.2 Characterisation of mutation types 69
3.2.3 Mutational signature analysis 72
3.2.4 Characterisation of driver gene mutations 75
3.2.5 Summary of findings 79
3.3 Evidence of selective pressures in mouse colon from varying disease states following uniform exposure to a mutagen 81
3.3.1 Survival analysis 82
3.3.2 Characterisation of intestinal phenotype 83
3.3.2.1 Macroscopic appearance 83
3.3.2.2 Immunohistochemical analysis 86
3.3.3 Genomic analysis 89
3.3.3.1 Analysis of mutation burden 90
3.3.3.2 Characterisation of mutation types 92
3.3.3.3 Mutational signature analysis 95
3.3.3.4 Characterisation of driver gene mutations . . . 101
3.3.4 Summary of findings 105
3.4 Discussion 106
3.4.1 Evidence of selective pressures in human colonic epithelium 106
3.4.2 Evidence of selective pressures in mouse colonic epithelium 109
3.4.3 Conclusion 111

3.1 Introduction

Cancer has long been considered to be a disease driven by mutation. Carcinogenesis thus relies upon the occurrence of somatic mutations: some of these mutations fall within key driver genes and bestow a selective advantage that promotes subsequent clonal expansion [63]. This premise suggests that stochastic mutations are present in normal cells prior to dysplastic transformation and are therefore detectable in normal tissue, a notion that has recently been proven in human skin. The Campbell lab utilised ultra-deep sequencing to reveal that mutations in skin cancer driver genes are under strong positive selection in 18 – 32% of histologically normal tissue [127]. These findings provided an initial glimpse into the capacity of normal tissue to acquire and tolerate somatic mutation. There are obvious parallels between the skin and the gut - both tissues are comprised of rapidly renewing epithelium under constant exposure to genotoxic stress in the form of UV radiation or the toxic contents of the gut lumen [128] respectively. It has also recently been demonstrated that different cancers exhibit distinct mutational signatures that reflect the various processes underlying their respective pathogeneses [142]. I therefore hypothesised that somatic mutation occurs frequently and is detectable in non-dysplastic human colonic epithelium. The mutational processes active within intestinal tissue from different disease states (such as those caused by familial syndromes or inflammation) will thus vary and reflect variable disease pathogenesis. In the first part of this chapter, I utilise a novel sequencing method to investigate this hypothesis within human colonic epithelium. I then go on to explore the impact of underlying disease state upon somatic mutation acquisition in mouse models of intestinal disease.

3.1.1 Chapter aims

Specific aims of this chapter were

- To investigate whether mutational signatures in histologically normal human colonic epithelium vary with underlying disease state, thus reflecting variable pathogenesis
- To utilise animal models to assess whether underlying disease state impacts upon the effect of uniform exposure to a standard mutagen

3.2 Evidence of selective pressures in normal and inflamed human colon

In order to characterise somatic mutation within human colonic epithelium, intestinal biopsies were collected from three patient groups (Figure 3.1). The 'healthy' group contained patients with no diagnosis of intestinal disease and who exhibited a 'clear' endoscopic assessment with no signs of inflammation or dysplasia. The age of patients within this group ranged from 45 to 55 years old, with biopsies taken from the ascending (A), transverse (T) and sigmoid (S) regions of the large colon. 'Inflamed' biopsies were taken from confirmed sites of inflammation from patients diagnosed with inflammatory conditions, such as Crohn's disease, Ulcerative Colitis (UC), Primary Sclerosing Cholangitis (PSC) or Inflammatory Bowel Disease (IBD). The age of patients within the inflamed group ranged from 24 to 54 years old. The third experimental group was comprised of two familial syndrome patients, harbouring confirmed germline mutations for HMPS and Lynch syndrome respectively, and one patient harbouring a confirmed homozygous loss of function mutation in the base excision repair (BER) gene *MBD4*. Intestinal biopsies from the A, T and S regions were collected from the HMPS and Lynch patients during routine scheduled endoscopies, with an additional aberrant crypt foci (ACF) biopsy taken from the Lynch patient. The BER patient harboured 70 - 100 colonic polyps and thus underwent a total colectomy,

Patient	Age	Intestinal Diagnosis	Sampled Regions
1	49	Normal	A, T, S
2	50	Normal	A, T, S
3	50	Normal	A, T, S
4	55	Normal	A, T, S
5	45	Normal	S
6	50	Normal	A, T, S
7	49	Normal	A, T, S
8	54	Crohns, UC	Transverse Ileum
9	36	Inflamed	Rectum
10	24	Crohns	Transverse Ileum, Rectum
11	30	Crohns	Transverse Ileum
12	53	UC	Rectum
13	25	Crohns, PSC	Transverse Ileum
14	35	Possible PSC	S
15	48	UC	Rectum
16	52	HMPS	A, T, S
17	50	Lynch (MSH2)	A, T, S, ACF
18	46	BER	Intervening normal A, T, S

Table 3.1: Patient information

with 'intervening normal' A, T and S biopsies sampled from the colectomy specimen at surgery. All other biopsies, and blood samples for germline mutation analysis, were collected during routine endoscopic procedures by a consultant gastroenterologist at the John Radcliffe Hospital, Oxford (Section 2.1.1). A summary of patient information is outlined in Table 3.1.

Following collection of 5 mm^2 intestinal biopsies, crypts were isolated by shaking (Section 2.1.2). DNA was extracted from pooled crypt (Section 2.4.1.1) and blood (Section 2.4.1.2) samples prior to molecular inversion probe (MIP) sequencing analysis.

MIP sequencing is a novel method that seeks to combine the sensitivity and coverage of multiplex targeted capture with the accuracy of single molecule tagging [132]. For this experiment, we utilised a panel of ~2000 MIP primers targeting 30 genes that are frequently mutated in CRC, at an average depth of 255x (Appendix Figure 2). Primer design, sample processing and sequencing (Section 2.4.3.1) were carried out by Dr Claire Palles and Laura Chegwidden

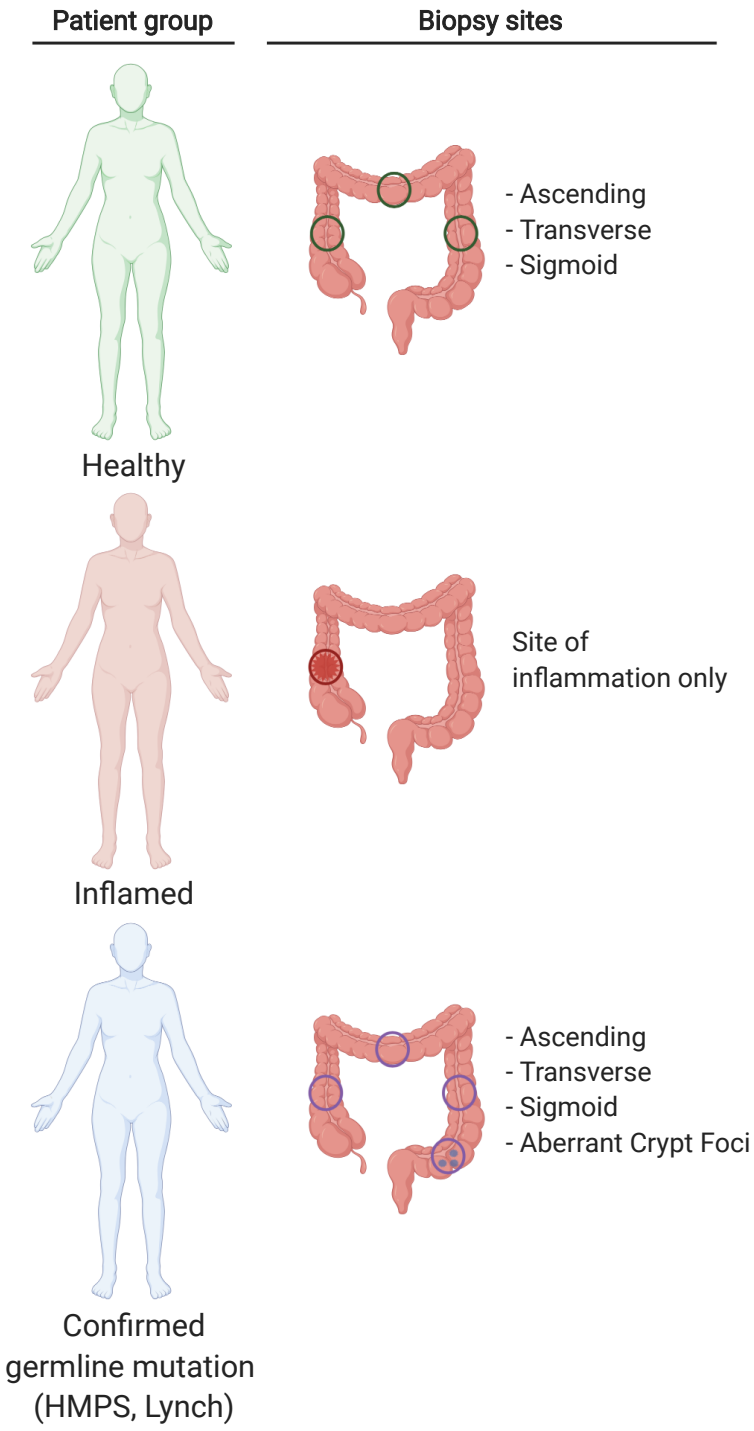


Figure 3.1: Colon biopsy sites for each patient group.

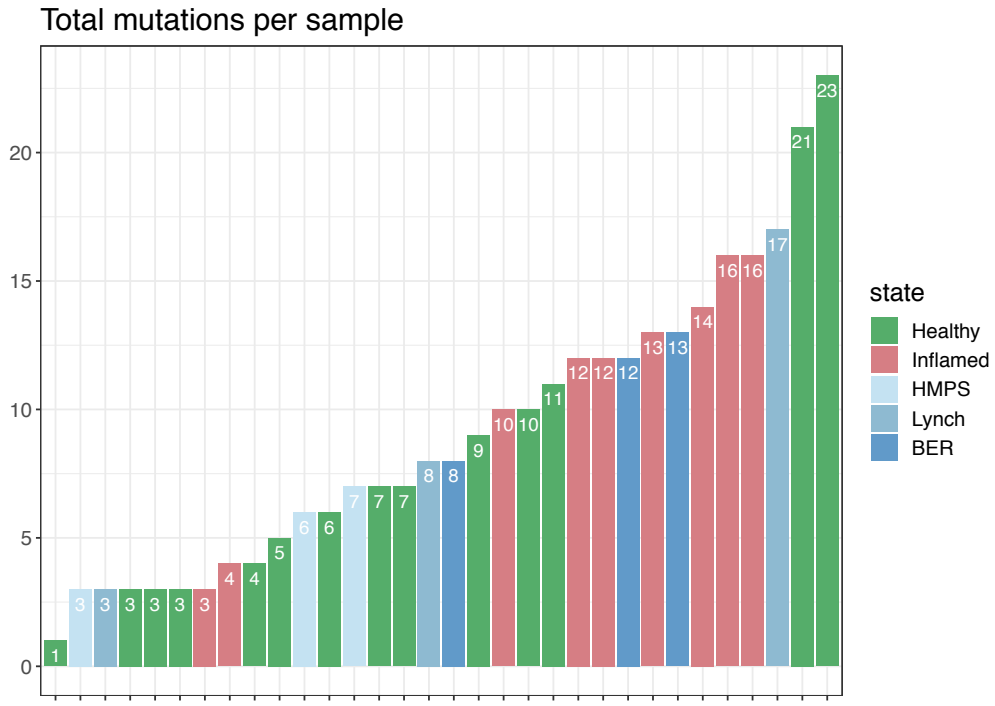


Figure 3.2: Total number of mutations detected per sample according to patient group

of the Palles Group (Institute of Cancer and Genomic Sciences, University of Birmingham). Variant calling (Section 2.4.3.5) and post-processing (Section 2.4.3.6) were carried out by Ryan Schenck, a bioinformatician within the Leedham group. I carried out all subsequent analysis using R and various publicly available resources, cited hereafter.

3.2.1 Mutation burden analysis

I first sought to analyse the total mutation burden within each sample. 14/19 biopsies from 6/7 healthy patients harboured detectable mutations within driver genes, ranging from 1 to 23 mutations per sample (Figure 3.2). Mutations were detected in all 9 inflamed biopsies, with a range of 3 to 16 mutations per sample; samples from the HMPS patient contained 3, 6 and 7 mutations, the Lynch patient 3, 8 and 17 (the transverse region biopsy failed to meet the read depth threshold and was omitted from analysis). Samples from the BER patient contained 8, 12 and 13 driver mutations. The age range of patients was too narrow to carry out meaningful analysis of age-associated mutation

accumulation, given that only inflamed samples harbouring increased numbers of mutations were from younger patients.

Analysis of the non-synonymous mutation burden per sample (Fig 3.3; A) showed a statistically significant increase in average non-synonymous mutation burden in samples of inflammation compared to those from healthy patients. This indicates that inflammation results in an increase in somatic mutations within driver genes compared to that observed within non-dysplastic intestine. Assuming an even distribution of mutations among crypts and that 10 mm^2 of colonic epithelium will harbour ~1000 crypts [145–147], we can estimate that 1.12% of crypts from non-dysplastic epithelium and 2% of crypts from inflamed epithelium will harbour a non-synonymous mutation in a driver gene. This doubling of the mutation rate from normal to inflamed tissue is in line with the findings of Lee-Six *et al* and Olaffson *et al*, in which whole genome sequencing of normal [146] and IBD-associated [148] colonic crypts revealed a low rate of mutations in normal crypts that doubles upon onset of IBD. There are, however, caveats to this comparison that will be discussed further later in this section.

3.2.2 Characterisation of mutation types

I next sought to investigate whether underlying disease state results in a shift in the types of mutation acquired. Analysis of the contribution of mutation types within each sample (Figure 3.4, A) revealed a high degree of variation in the percentage contribution of each mutation type within each sample. Only healthy and inflamed samples harboured indels, while UTR variants were present in inflamed, Lynch and BER samples but not healthy and HMPS samples. The percentage of missense variants was similar across healthy, inflamed, HMPS and Lynch samples, while BER samples exhibited an increase in the contribution of all mutation types (barring an absence of nonsense mutations) and a reduction in the contribution of missense variants relative to all other groups, presumably as a result of a global increase in genomic instability in these samples. These

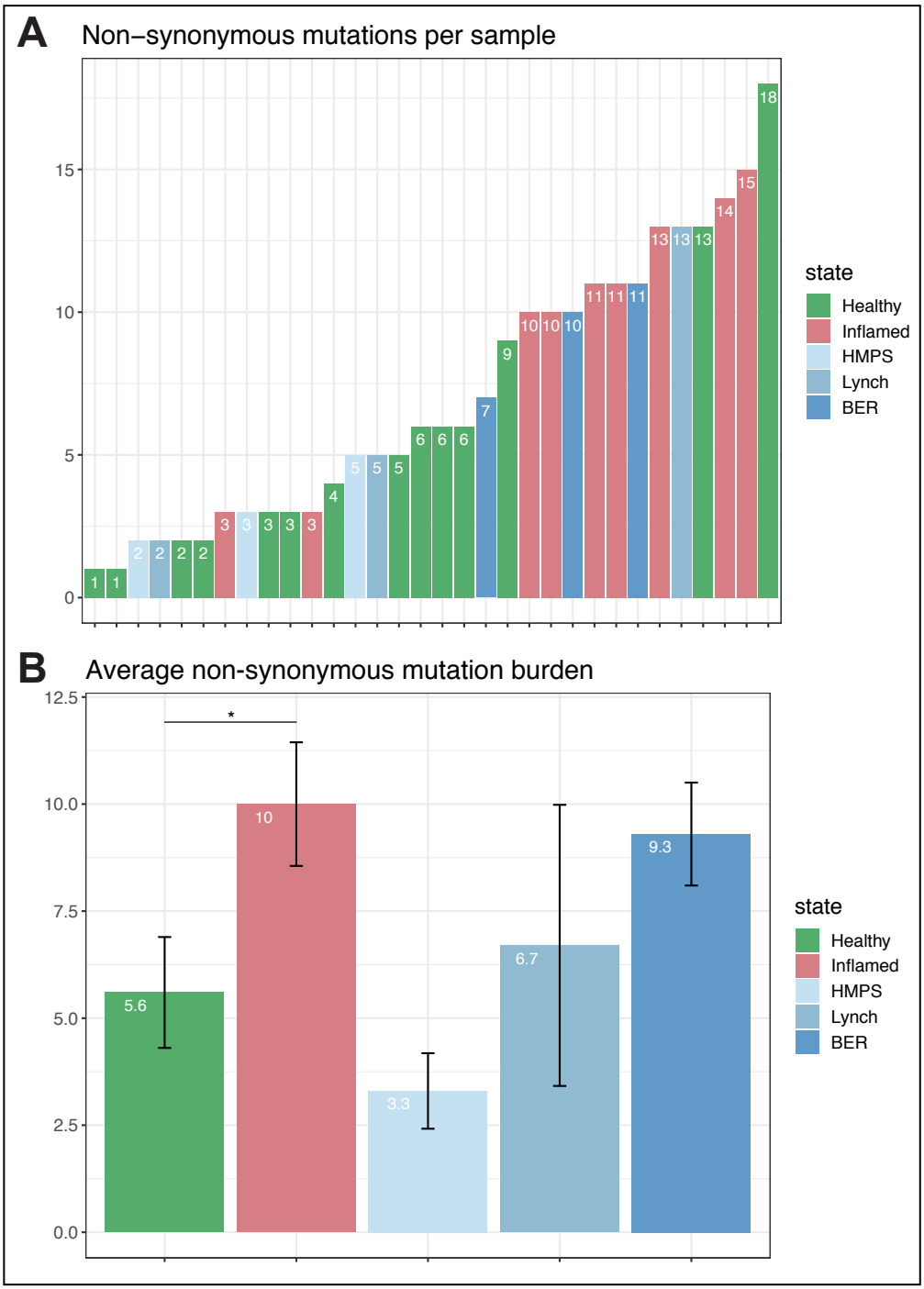


Figure 3.3: Non-synonymous mutations per sample (A) and patient group (B). Error bars represent SEM. *p < 0.05 as calculated by one-way ANOVA with Tukey post-test.

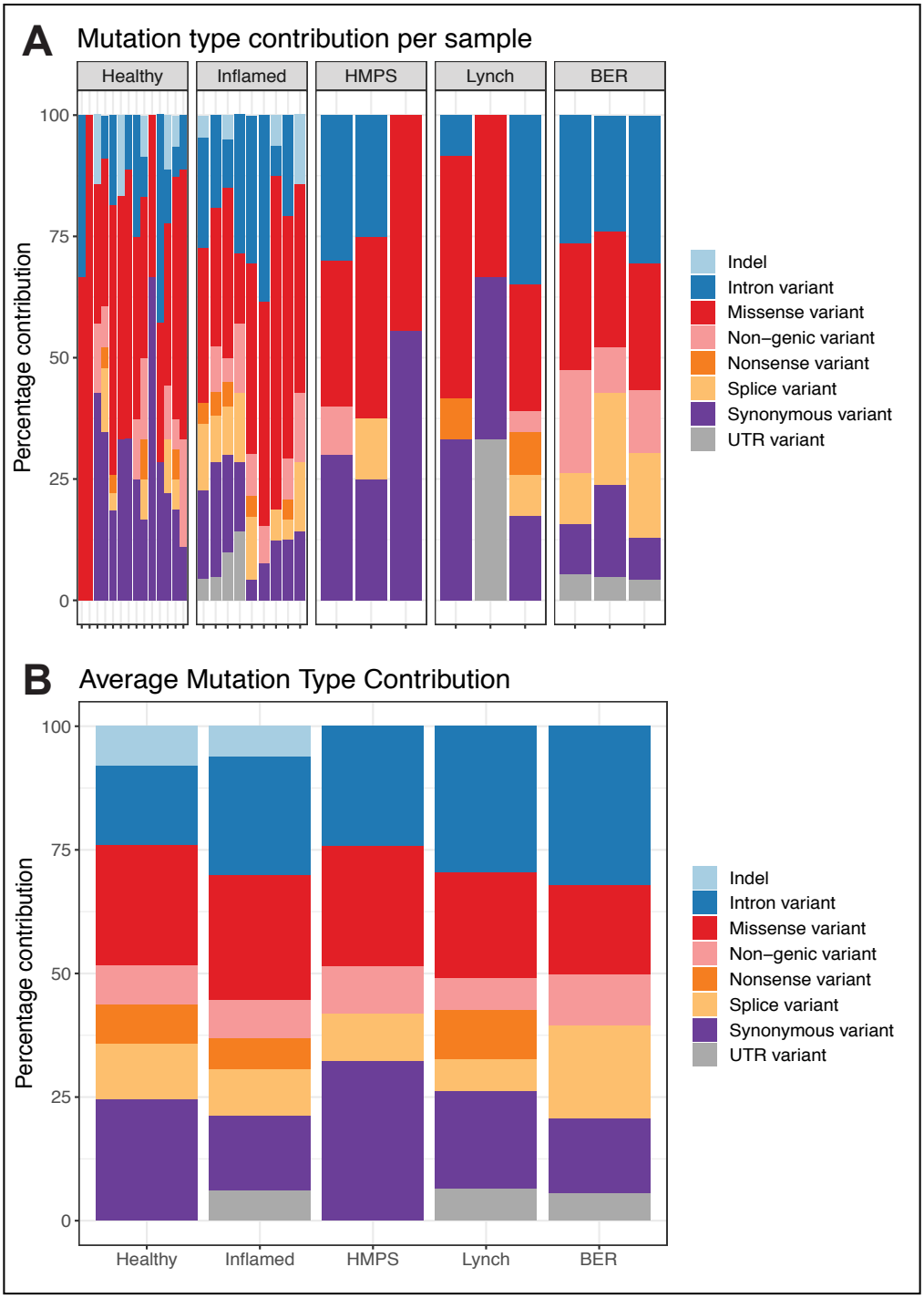


Figure 3.4: Characterisation of mutation types in samples from each patient group. Represented as percentage contribution of each mutation type per sample (A) and average percentage contribution of each mutation type per patient (B).

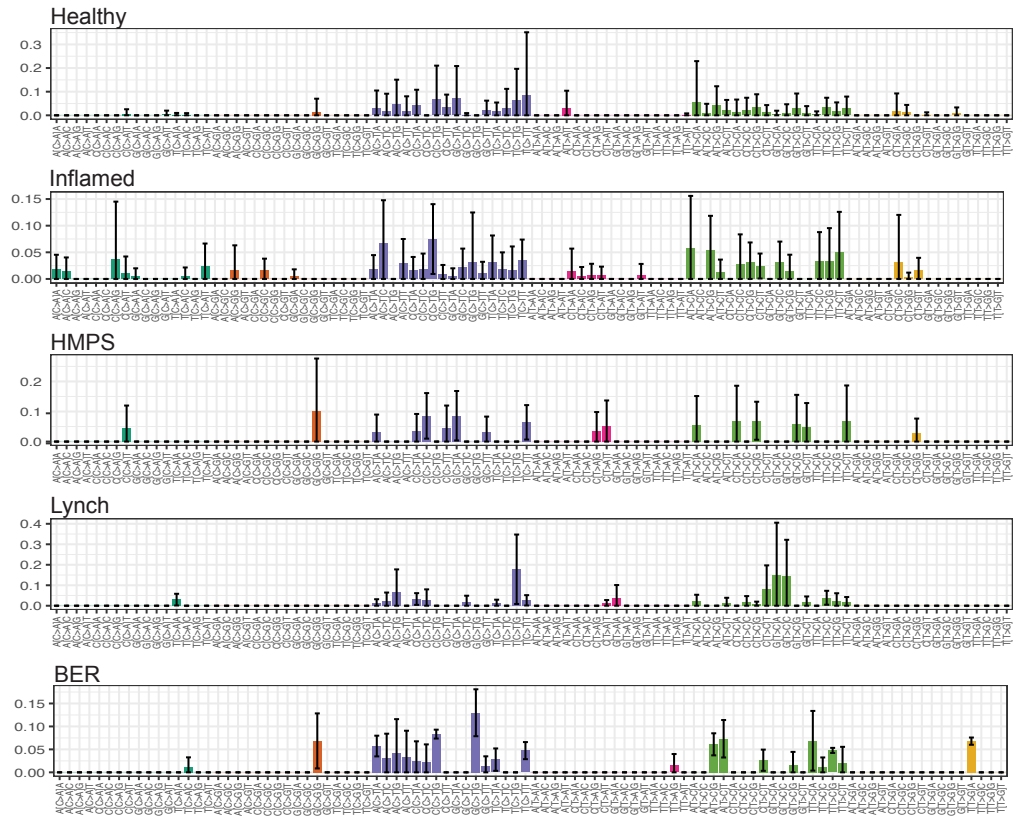


Figure 3.5: Characterisation of mutational spectrum of samples within each patient group. Represented as the average 96-substitution classification within samples from each patient group. Error bars represent SEM.

data thus indicate that there is a high degree of variation in mutation types that does not appear to be skewed according to intestinal condition.

3.2.3 Mutational signature analysis

I next sought to characterise the mutational processes active within samples from each patient group. Extraction of the average 96-class mutation signature for each patient group revealed the presence of distinct mutational 'fingerprints' (Figure 3.5) within the samples. Analysis of these revealed the contribution of 25 validated signatures associated with various mutational processes and documented within the COSMIC database.

The majority of mutations detected within samples from the Lynch patient were contributed by the clock-like SBS1 and MMR-associated SBS21 (Figure 3.6). Similarly, the majority of mutations detected in samples from the BER

patient were also contributed by SBS1, with the remainder contributed by a greater variety of MMR, NER and MSI-associated processes. This indicates that the majority of mutations in non-dysplastic tissue harbouring confirmed germline mutations in genes involved in the maintenance of genomic instability are contributed by aging-associated mutation accumulation alongside a variable number of processes associated with defects in genome surveillance mechanisms. The fact that the Lynch patient exhibited a lesser degree of variation in these non-aging associated mutational processes compared to the BER patient may reflect the specificity of the processes impacted by their respective mutations, with the Lynch patient harbouring a mutation in the MMR-specific *MSH2* gene while the BER mutation results in a loss of function (LOF) in the more versatile MBD4 protein. Samples from healthy patients harboured a broad variety of active mutational processes, including those associated with defects in MMR, NER and alkylation-induced repair, but clock-like mutation accumulation remained detectable as contributing ~40% of all mutations.

Contrastingly, the clock-like signature SBS1 was absent in samples from the HMPS patient, with mutations instead contributed by a variety of processes associated with defects in NER, MMR and alkylation-induced repair. A similar dilution of the clock-like mutational process signal was detected in inflamed samples, in which the majority of accumulations were contributed by SBS15 and 16, associated with defects in MMR and NER respectively, alongside the greatest variety of other mutational processes detected across all patient groups. This perhaps indicates that the accumulation of mutations in non-dysplastic tissue exposed to constitutive microenvironmental disruption occurs via mutational processes that overtake clock-like mutation accumulation as the primary mutagenic driver; whereas defective genome surveillance mechanisms contribute to mutation acquisition alongside age-associated accumulation. This is illustrated by the difference in the average contribution of SBS1 in samples from healthy, Lynch and BER patients (42%, 58% and 61% respectively) compared to samples from the inflamed patient group (12%) and the single HMPS patient (absent). Taken together,

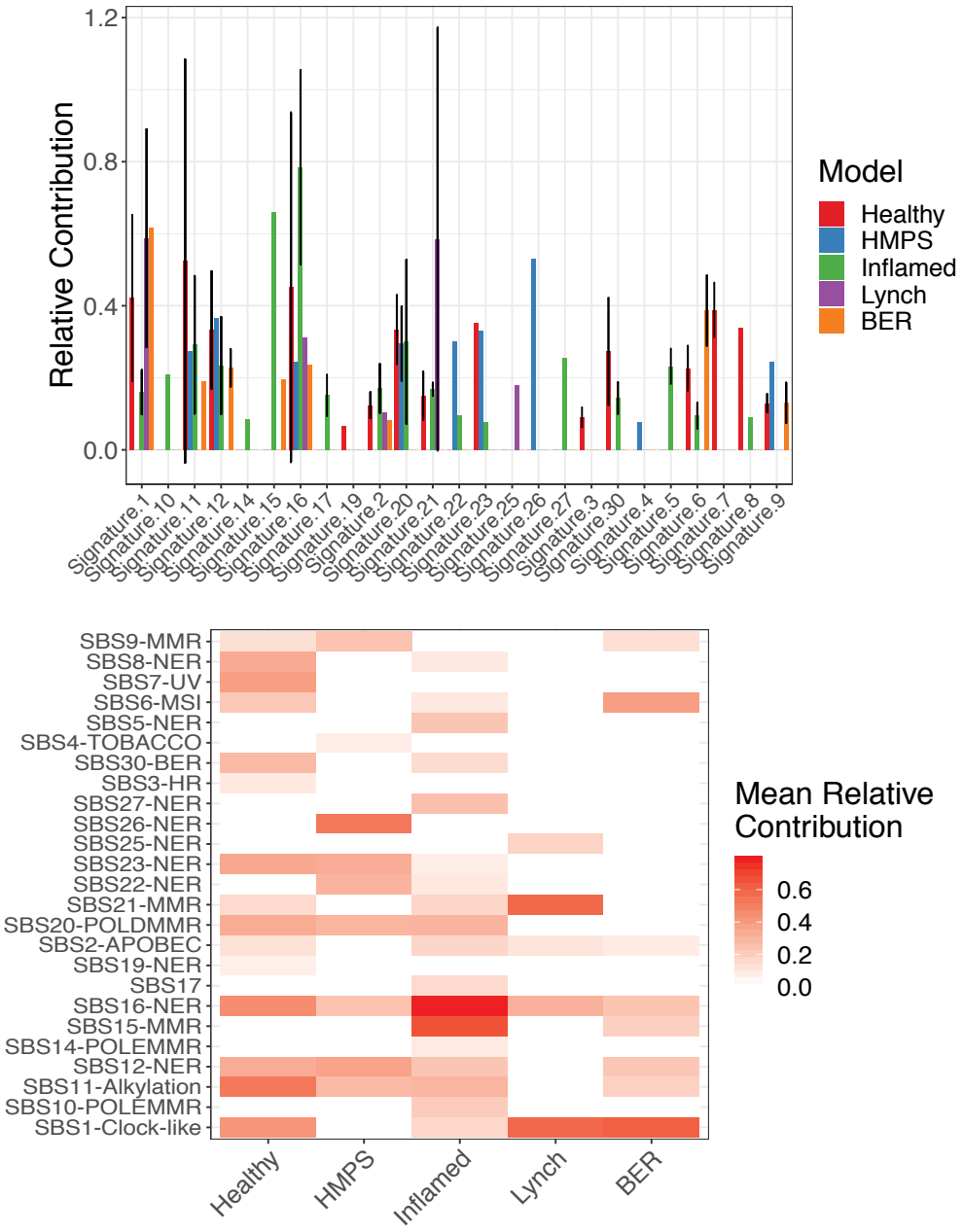


Figure 3.6: Mutational process contribution in samples from each patient group. Represented as the relative contribution of each detected mutational signature to all mutations within samples from each patient group (top as bar plot, bottom as heat map). Error bars represent SEM.

these findings indicate that differences in intestinal disease state are reflected in the variety of active mutational processes. Interestingly, samples taken from patients harbouring microenvironmental disruption, as a result of inflammation or HMPS, exhibit a variety of mutational processes that appear to overtake clock-like mutation accumulation. Conversely, samples from healthy patients and those harbouring mutations in MMR-associated genes exhibited largely clock-like mutational accumulation alongside additional mutational processes associated with defective genome surveillance mechanisms. These data thus shed light on the impact of microenvironmentally driven disease states upon mutation acquisition compared to those underpinned by genomic instability as a result of mutations in the MMR genes *MSH2* and *MBD4*, respectively. This notion raises the intriguing prospect that cell-extrinsic microenvironmental disruption can also impact upon the acquisition of somatic mutation in a way that is distinct from germline mutations in genes specifically involved in the maintenance of genomic integrity.

3.2.4 Characterisation of driver gene mutations

I next sought to specifically investigate the genes mutated within samples from each patient group. Characterisation of the genes harbouring non-synonymous mutations per sample revealed significant variety in the genes mutated per patient group (Figure 3.7).

Visualisation of the variant allele frequency (VAF) of missense mutations across all samples showed the majority of missense mutatations are present at low frequencies ranging from 1-10% (Figure 3.8). However, a subset of mutations occurred at higher VAFs, indicative of a loss of heterozygosity. The majority of mutations detected at these higher VAFs occurred in inflamed samples and, respectively, indicated a loss of heterozygosity of *MLH1* and *RNF43* (P9) and *ATM*, *TGIF1* and *RNF43* (P10).

However, all samples from the BER patient and one sample from a healthy patient (P6) also harboured high VAF missense *RNF43* mutations. These data thus indicate that non-dysplastic intestinal epithelium harbours missense variants

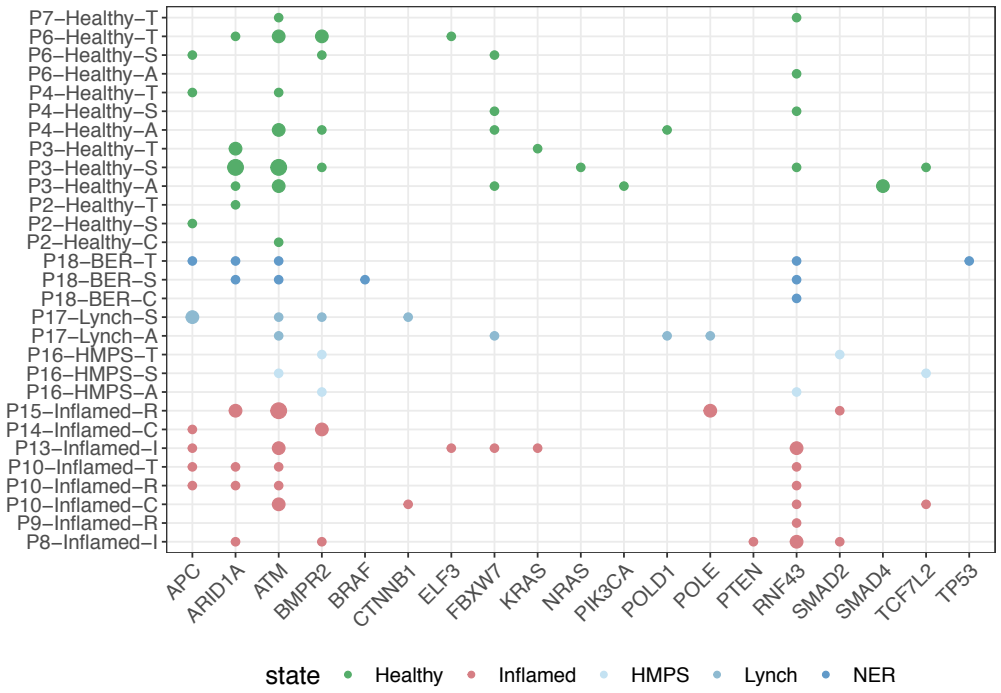


Figure 3.7: Distribution of non-synonymous driver gene mutations in samples from all patient groups. Samples (coloured according to patient group) are indicated on the y-axis, driver genes on the x-axis. Dot size represents the number of mutations (range 1-3)

that largely occur at low VAFs, with variants detected at higher VAFs indicative of potential loss of heterozygosity detected primarily in inflamed samples. It may also be that intestinal inflammation results in preferential acquisition of *RNF43* mutations, discussed later in this section.

In order to further characterise missense variants, I next analysed the number of variants per target gene and their predicted impact per patient group. Analysis of the number of variants per gene per patient group shows that the majority of missense variants were detected in healthy and inflamed samples (Figure 3.9).

Despite this, the majority of missense single nucleotide variants (SNVs) detected per patient group (except for HMPS) were predicted to be deleterious variants by PredictSNP2, a tool that utilises four SNV classifiers to generate a consensus prediction of the effect of an SNV [149]. Furthermore, the majority of protein changes caused by the missense SNVs within each patient group (except for Lynch) were predicted as 'damaging' by PROVEAN, a a software tool which predicts whether an amino acid substitution has an impact on the biological

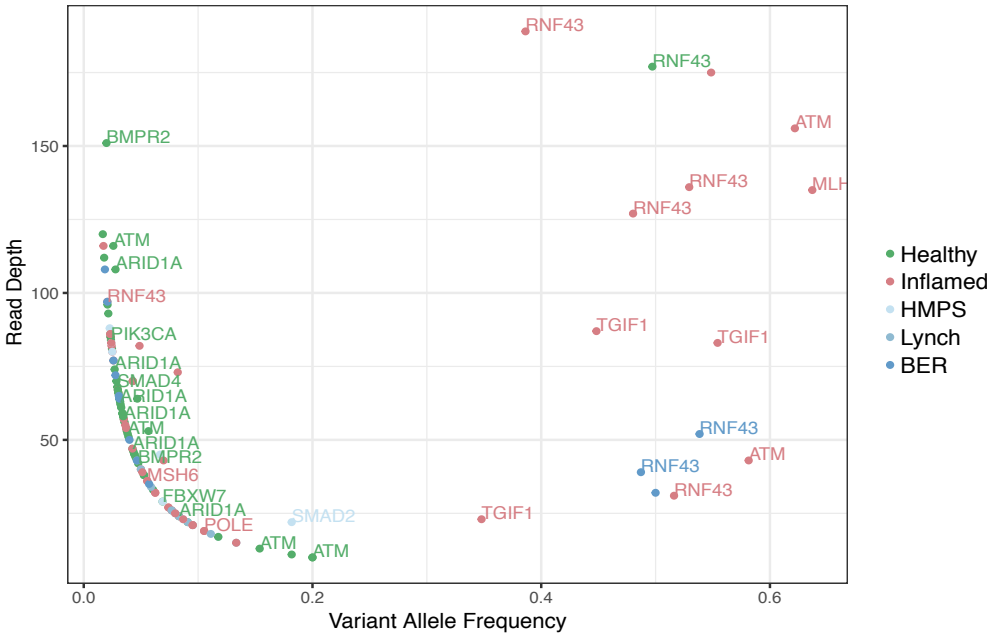


Figure 3.8: Visualisation of the read depth and variant allele frequency of missense mutations, coloured by patient group.

function of a protein [150]. These data thus indicate that potentially damaging missense mutations are present within samples from all patient groups, with the highest number of these variants detected in healthy and inflamed samples.

I next sought to investigate the most recurrently mutated target genes within each patient group. As illustrated in Figure 3.10, *ATM* was the most frequently mutated gene in this analysis, with all patient groups harbouring at least one *ATM* mutation. There were no preferentially mutated genes within each patient group, with variants within the most frequently mutated genes evenly distributed among all samples. However, *SMAD4* mutations were absent in healthy patients and all three samples analysed from the BER patient harboured *PIK3CA* mutations.

Given the lack of preferentially mutated target genes within patient groups, I next sought to contextualise my findings with recent data generated by Lee-Six *et al*, who utilised whole-genome sequencing to analyse the landscape of somatic mutation in normal colonic crypts [146]. Figure 3.11 shows the proportion of driver gene mutations they detected in normal colon ('Normal Drivers'), alongside those detected in TCGA colorectal cancer samples ('Cancer Drivers'). Comparison

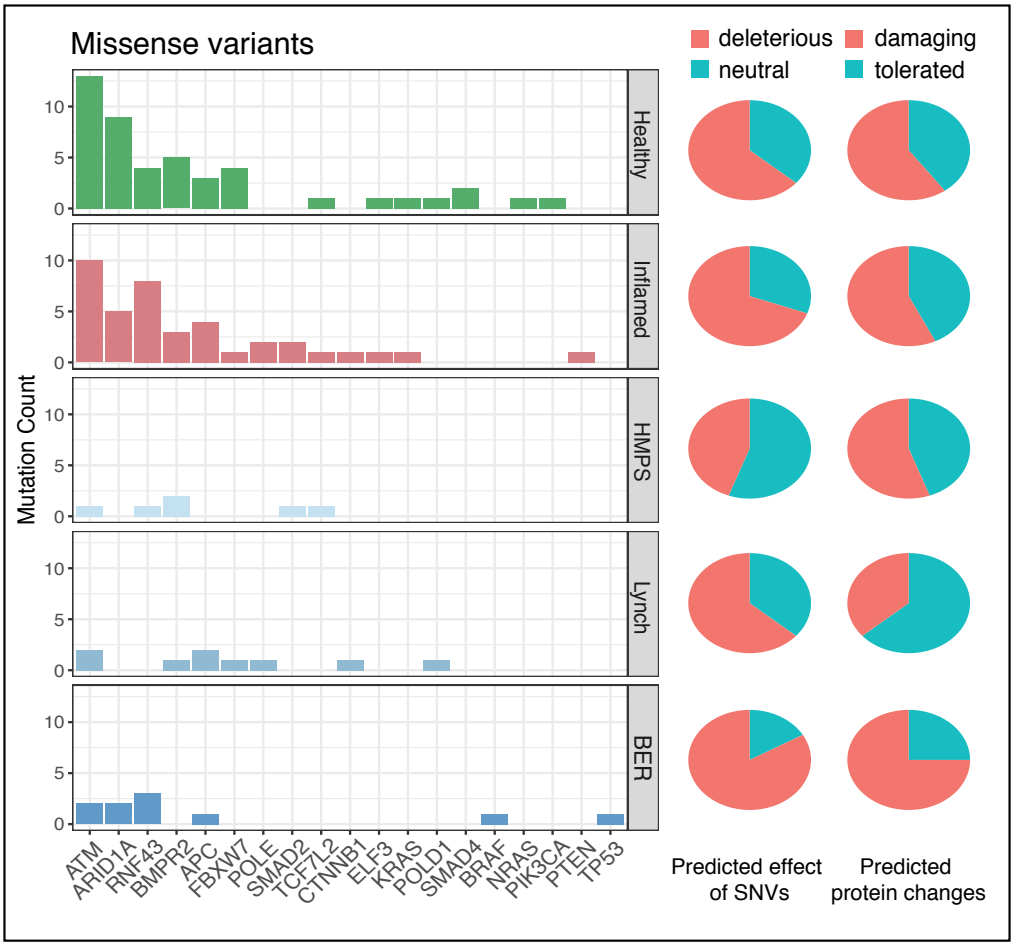


Figure 3.9: Characterisation of missense variants within samples from each patient group. Represented as the number of missense mutations per gene within each patient group (left panel), the proportion of missense mutations predicated to be deleterious or neutral by the PredictSNP2 tool (left pie charts) and the proportion of missense mutations predicted to result in damaging or tolerated protein changes by the PROVEAN tool (right pie charts).

of these data highlights the greater variety of mutated genes detected in our analysis compared to that of Lee-Six *et al*, with only 3 of the most frequently mutated drivers in CRC also detected in normal colon by Lee Six *et al* [146]. In our samples, however, 8 cancer drivers are present in the most frequently mutated genes in both healthy and inflamed samples.

Taken together, these findings indicate that potentially damaging missense mutations are occurring in CRC driver genes within inflamed intestinal epithelium but, strikingly, also within non-dysplastic tissue from healthy patients. This indicates a level of tolerance within non-dysplastic epithelium for missense driver

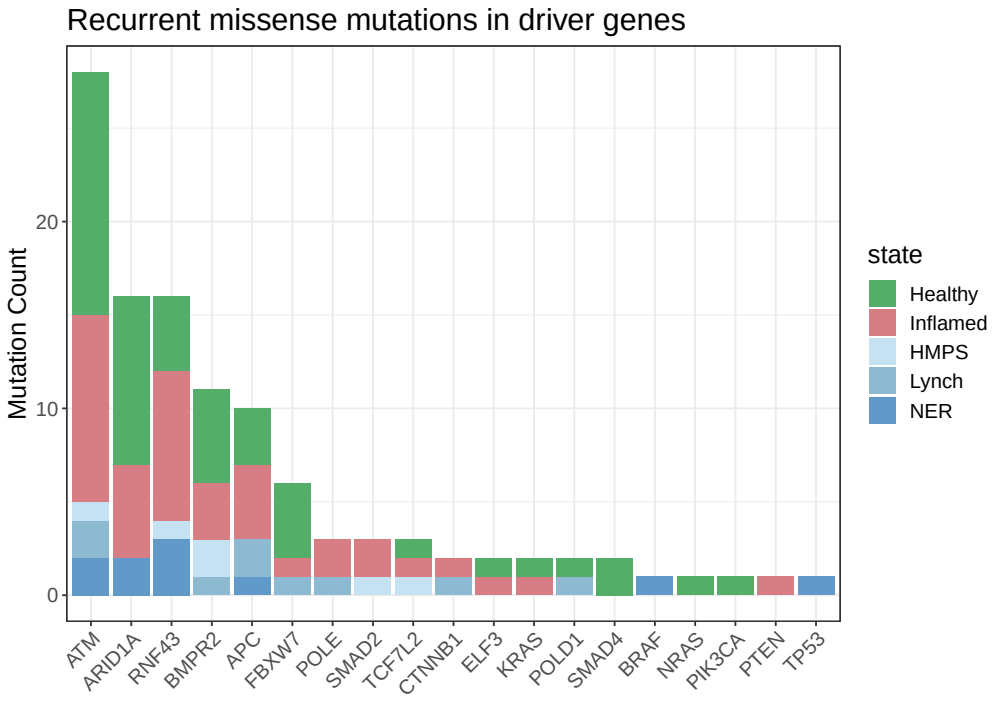


Figure 3.10: Characterisation of genes harbouring the most missense mutations per patient group.

gene mutations occurring at low VAFs. However, inflamed tissue is more likely to contain potentially damaging mutations, especially within *RNF43*, occurring at higher VAFs that indicate potential loss of heterozygosity within these samples.

3.2.5 Summary of findings

Key findings of the work presented thus far are outlined below:

- Mutations in CRC driver genes are present even in non-dysplastic human colonic epithelium
- Differences in intestinal disease state do not appear to produce shifts in the types of mutations present within healthy and inflamed samples
- Differences in intestinal disease state appear to be reflected by the mutational processes active within colonic epithelium, with the majority of mutations in healthy, non-dysplastic tissue contributed by clock-like processes that are

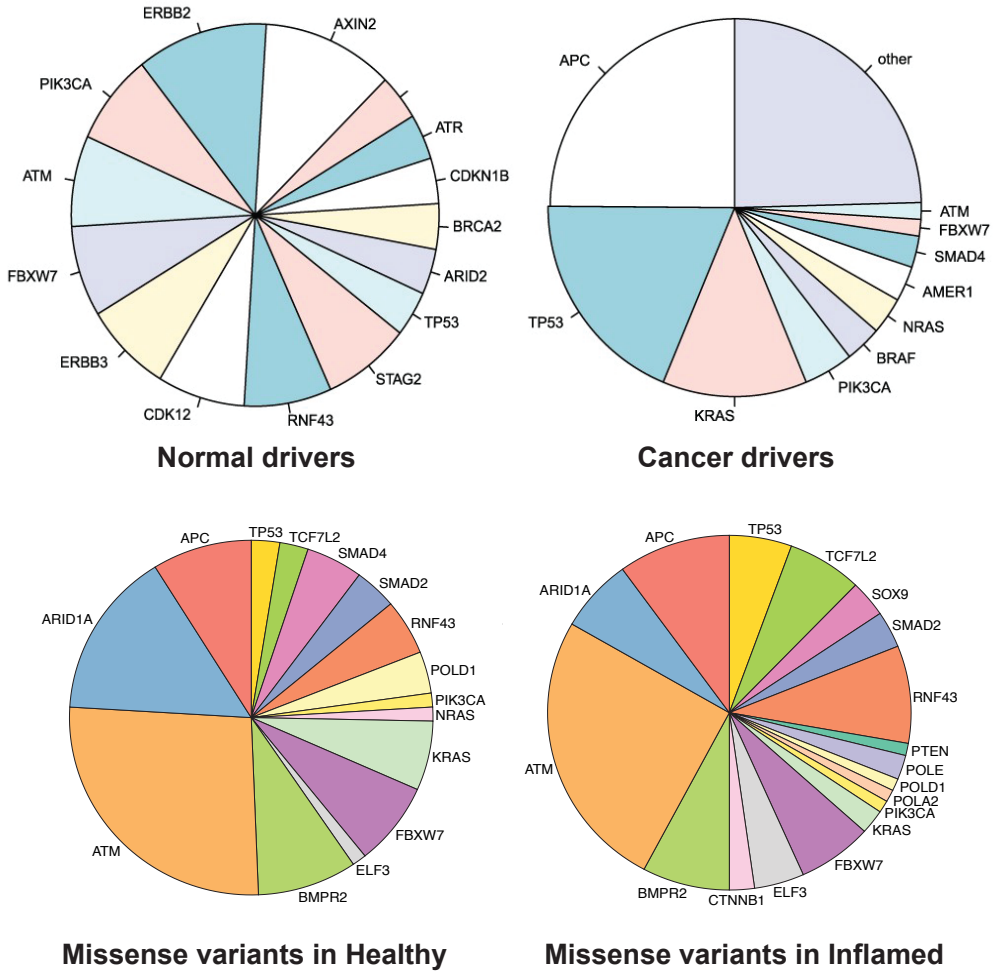


Figure 3.11: Comparison of missense mutations detected in healthy and inflamed samples. Represented as the proportion of missense variants per gene within healthy (bottom left) and inflamed (bottom right) samples. The proportion of mutations in driver genes of CRC detected in normal colonic epithelium (top left) and TCGA colorectal adenomas (top right) by Lee-Six *et al* [146] is shown for comparison.

overtaken or absent in tissue exposed to an inflammatory or constitutively deranged microenvironment

– Missense mutations in CRC driver genes are as prevalant in non-dysplastic tissue from healthy patients as in biopsies of inflamed epithelium. A majority of these variants are predicted to be deleterious SNVs that result in damaging protein changes.

– In non-dysplastic epithelium, these potentially functionally damaging missense mutations occur at low VAFs, ranging from 1-10%. Inflamed

tissue, however, harbours an increase in high-VAF missense mutations that may indicate inflammation-associated loss of heterozygosity in these samples.

In view of my original hypothesis, the results presented here indicate that differences in intestinal disease state correspond to differences in both underlying mutational processes and the acquisition of potentially damaging mutations within CRC driver genes. These findings are discussed further at the end of this chapter.

3.3 Evidence of selective pressures in mouse colon from varying disease states following uniform exposure to a mutagen

Given that results presented thus far have shown that mutations in CRC driver genes are prevalent even in healthy, non-dysplastic human colon, and more so in biopsies of inflamed colon, I next sought to further explore the effect of underlying disease state on somatic mutation impact. I hypothesised that microenvironmental and cell of origin-driven differences in disease state would correspond to detectable differences in mutation acquisition following uniform mutagen exposure. In order to explore my hypothesis, I exposed mice with varied underlying intestinal disease processes to the known carcinogen AOM (Section 2.2.4.1-2). Mice within each experimental group were from either wild-type (WT) or *Vil1-Grem1* (VG) C57BL/6J genetic backgrounds, aged three months to allow for the development of ectopic crypts in VG mice. The treatment protocol for each of the six experimental groups utilised is outlined in Figure 3.12. Following the treatment period, mice were monitored daily and euthanised upon exhibition of symptoms of intestinal disease, such as anaemia, hunching or diarrhoea. The intestinal tract was collected prior to sampling of colonic polyps and adjacent normal tissue for DNA extraction (Section 2.2.6.2). The small intestine, divided into three regions (SB1, SB2 and SB3), and colon were then swiss rolled and

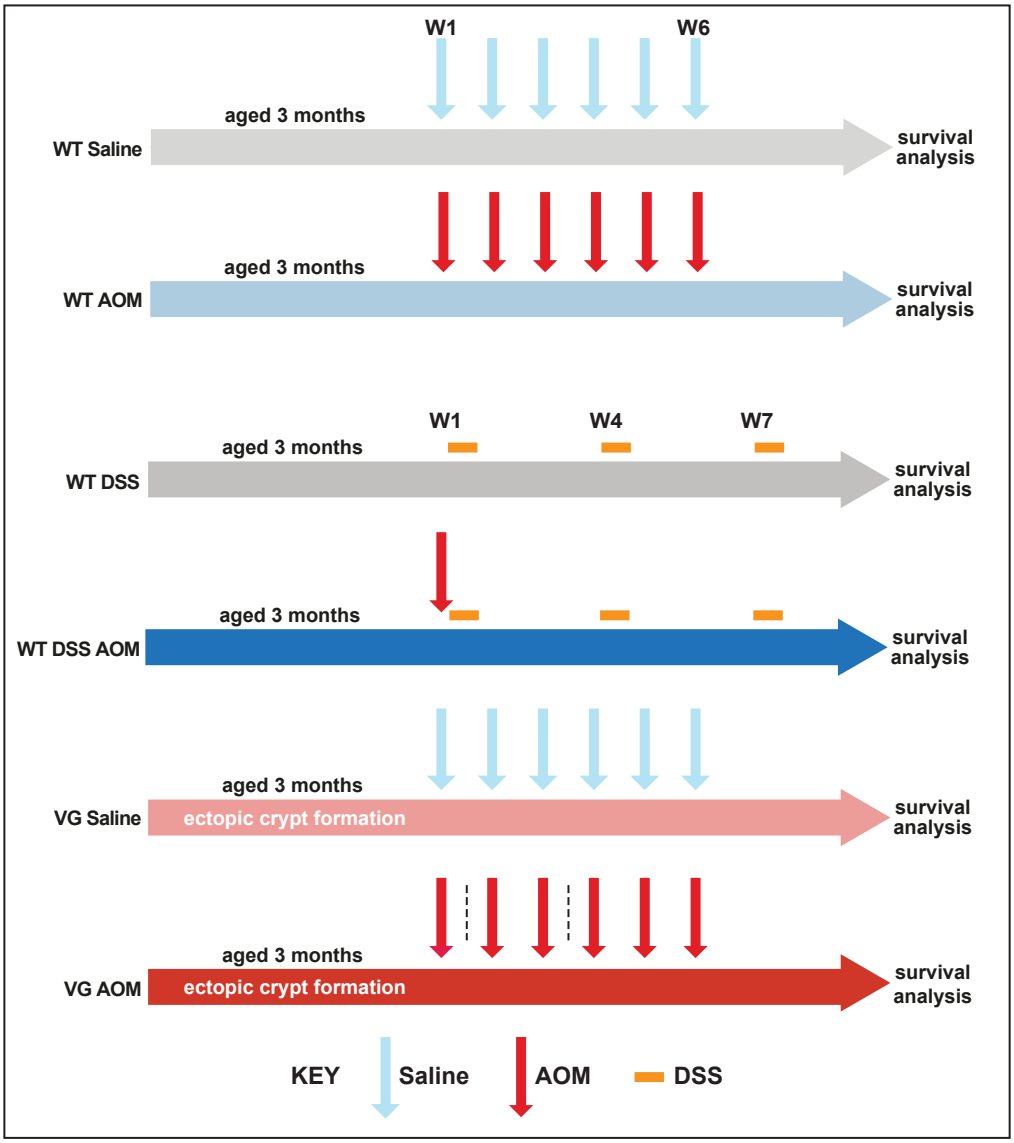


Figure 3.12: Treatment scheme for control and experimental groups of wild-type and *Vil1-Grem1* mice.

processed for immunohistochemical analysis of intestinal epithelial markers (Section 2.2.6) and pathology scoring (Section 2.3.5).

3.3.1 Survival analysis of wild-type and *Vil1-Grem1* mice exposed to DSS and AOM

I first sought to characterise differences in survival between experimental groups. As shown in Figure 3.13, WT mice administered saline or DSS alone did not exhibit decreases in survival, with none of the mice in these groups culled due to

reaching experimental endpoints. Given that AOM had never been administered to VG mice before, one, three and six doses were given in order to establish a dose-response relationship in this model. Three and six doses of AOM produced a small, significant ($p = <0.0001$) reduction in survival compared to a single dose, with a median survival of 175 days for VG AOM 3 mice and 181 days for VG AOM 6 mice compared to 215 days for VG AOM 1 mice (Figure 3.13; A). Overall, the median survival within WT experimental groups was 413 days for WT AOM and 264 days for WT DSS AOM, while median survival within VG experimental groups was 249 days for VG Saline and 182 days for all VG AOM mice (Figure 3.13; B). These findings demonstrate that AOM mutagenesis is associated with a decrease in survival. Interestingly, these data also indicate that the effect of mutagenesis upon underlying disease state, driven either by inflammation or aberrant *Grem1* expression, is cumulative, with 'mutagen+inflammation' WT DSS AOM and 'constitutively deranged' VG experimental groups exhibiting significant decreases in median survival (Figure 3.13; C). These results therefore provide evidence of the differential impact of mutagenesis on different underlying disease states.

3.3.2 Characterisation of the intestinal phenotype of wild-type and *Vil1-Grem1* mice exposed to DSS and AOM

3.3.2.1 Macroscopic appearance

Having demonstrated the impact of underlying disease states on mouse survival following standard mutagenesis, I next sought to characterise the impact of AOM-driven mutagenesis on intestinal phenotype within each experimental group (Figure 3.14; A). As previously demonstrated by Davis *et al* [51], VG mice exhibited an increase in average SB and villus length (Appendix Figure 3), and exposure to AOM did not impact upon small intestinal polyp histology in these mice (Appendix Figure 4). Given the colonotropic nature of AOM-induced mutagenesis, I thus focused on characterisation of colon pathology within mice from each experimental group.

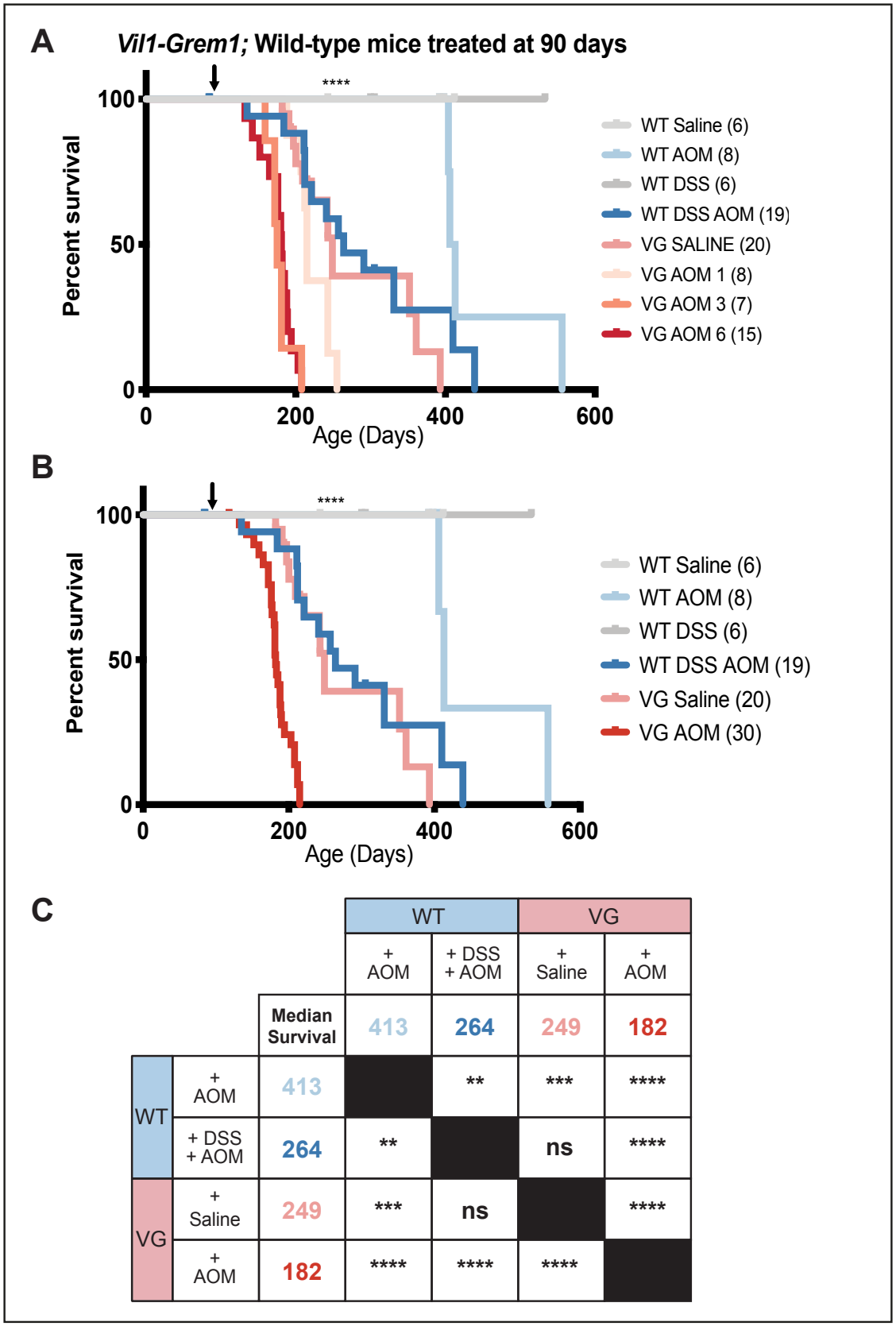


Figure 3.13: Kaplan-Meier survival curves (top) and analysis (bottom) of WT and VG mice exposed to AOM and DSS, showing individual (A) and combined (B) curves for VG mice exposed to varying doses of AOM. Arrow indicates start of treatment scheme, N as indicated. **p < 0.01, ***p < 0.001, ****p < 0.0001 as calculated by log-rank Mantel-Cox test.

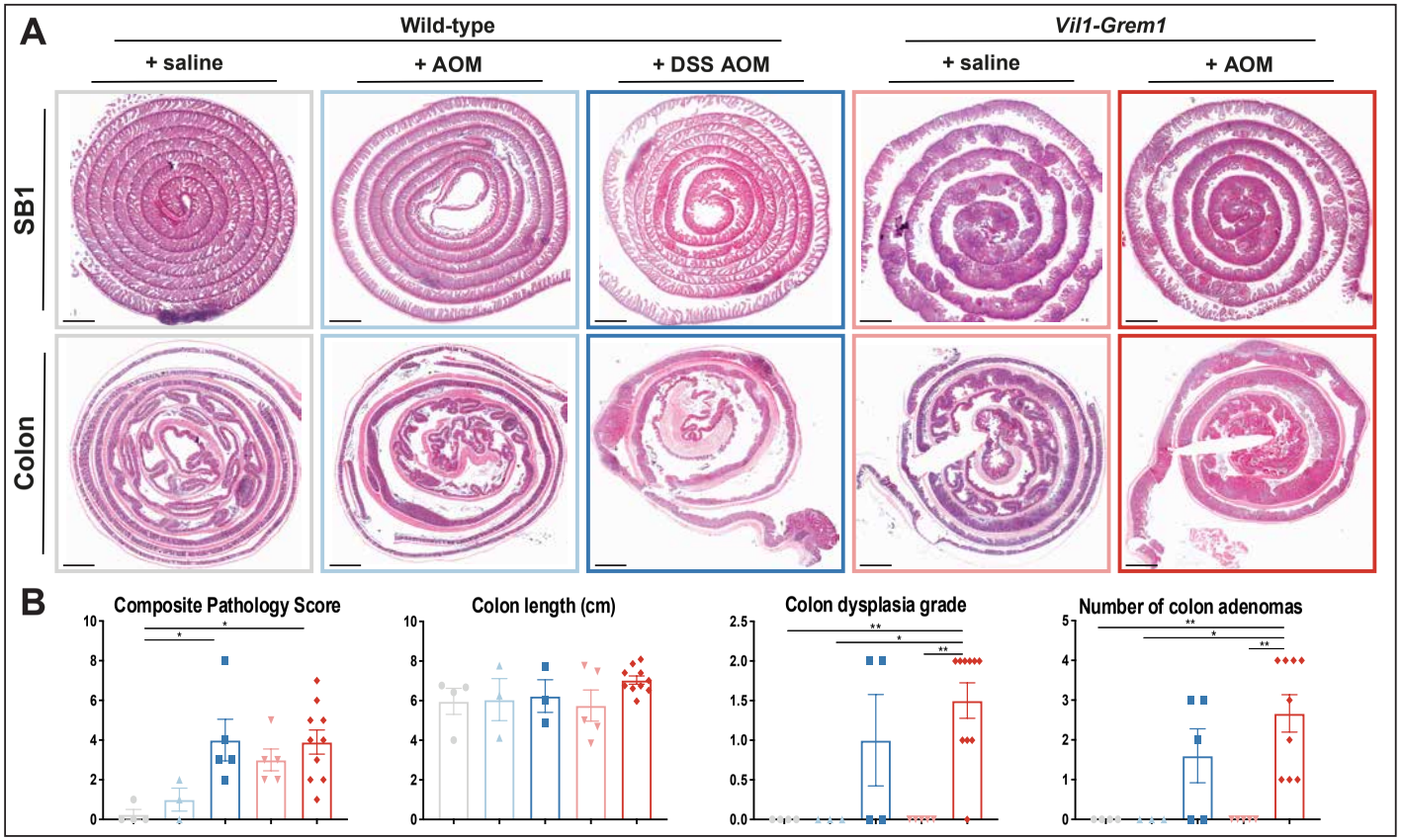


Figure 3.14: Macroscopic intestinal phenotype of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. A) Representative images of hematoxylin and eosin staining of small intestine section 1 (SB1) and colon of experimental mice. Scale bars represent 1 mm. B) Quantification of whole bowel composite pathology scores and colon length, dysplasia grade and adenoma number of mice within each experimental group. Points represent individual mice, bars and error bars represent mean±SEM. *$p < 0.05$, **$p < 0.01$ as calculated by one-way ANOVA with Tukey post-test.

3. Underlying disease state impacts upon mutation acquisition

There was no significant difference in colon length between experimental groups, although mice in the VG AOM group had the longest colons on average. Composite pathology scores (Section 2.3.5) relating to the whole intestine varied according to disease state, with intestinal sections from WT DSS AOM and VG AOM mice exhibiting the highest average composite scores (3.9 and 4.0, respectively) (Figure 3.14; B), further confirming the well documented impact of AOM mutagenesis in the context of underlying DSS-induced inflammation.

Within the colon, the average dysplasia grades for WT DSS AOM and VG AOM sections were 1 and 1.5, respectively, where a score of 1 indicates low grade dysplasia and 2 indicates high grade dysplasia. Furthermore, colons from WT DSS AOM mice contained an average of 1.6 adenomas, compared to 2.7 for sections from VG AOM mice. Taken together, these findings indicate that WT DSS AOM and VG AOM mice develop the most severe colon phenotype of all experimental groups, with colons from VG AOM mice harbouring higher grade dysplasia and more adenomas on average.

3.3.2.2 Immunohistochemical analysis

I next sought to investigate whether the observed differences in survival and intestinal pathology were driven by changes in colon histology. The average length of colon crypts was significantly increased in VG AOM mice compared to WT Saline and WT AOM mice, and approximately double that of all other groups. Administration of DSS resulted in a decrease in the proportion of proliferative and progenitor cells per colon crypt compared to WT Saline and WT AOM mice, demonstrating the impact of DSS-driven regeneration on colon histology. In VG mice, the increase in colon crypt length observed in VG AOM mice appears to be driven by an expansion of the proliferative cell compartment, compared to that of VG Saline mice, that is not accompanied by a concurrent increase in the percentage of progenitor cells per crypt. This is evidenced in Figure 3.15 as a statistically significiant increase in the percentage of $Ki67^{+}$ cells in VG AOM

colon crypts compared to VG Saline colon crypts, with no such difference in the percentage of $Sox9^+$ cells per crypt.

I next characterised the percentage of $Ki67^+$ and $Sox9^+$ cells within colon polyps from WT DSS AOM and VG AOM mice. There was no significant difference in polyp size or the percentage of $Ki67^+$ and $Sox9^+$ cells within colon polyps from WT DSS AOM mice compared to VG AOM mice (Figure 3.15), indicating that disease state does not result in a change in the size or proportion of proliferative and progenitor cells within colon polyps following exposure to AOM.

Given the observed variety in colon polyp sizes, I next analysed the relationship between colon polyp size and survival and found a significant negative correlation within samples from the VG AOM experimental group only ($R = -0.71$, $p = 0.046$; Figure 3.16).

This indicates that VG AOM mice harbouring larger colon polyps met the experimental endpoint earlier than those with smaller polyps. Importantly, given that the survival of VG mice is usually determined by SB phenotype, which is unimpacted by exposure to AOM, this result establishes a link between colon phenotype and survival in VG mice exposed to AOM, strengthened by the dose-survival relationship described earlier. Taken together, these findings indicate that the differences in survival and colon pathology between WT DSS AOM and VG AOM mice described earlier are associated with distinct AOM-induced changes in colon histology: WT DSS AOM mice exhibit a reduction in proliferative and progenitor cells per colon crypt, while increased colon crypt length in VG AOM mice is driven by an increase in mutationally susceptible proliferative cells, which may explain the increase in colon polyp number observed in VG AOM mice. These differences do not, however, result in detectable differences in colon polyp size or histology.

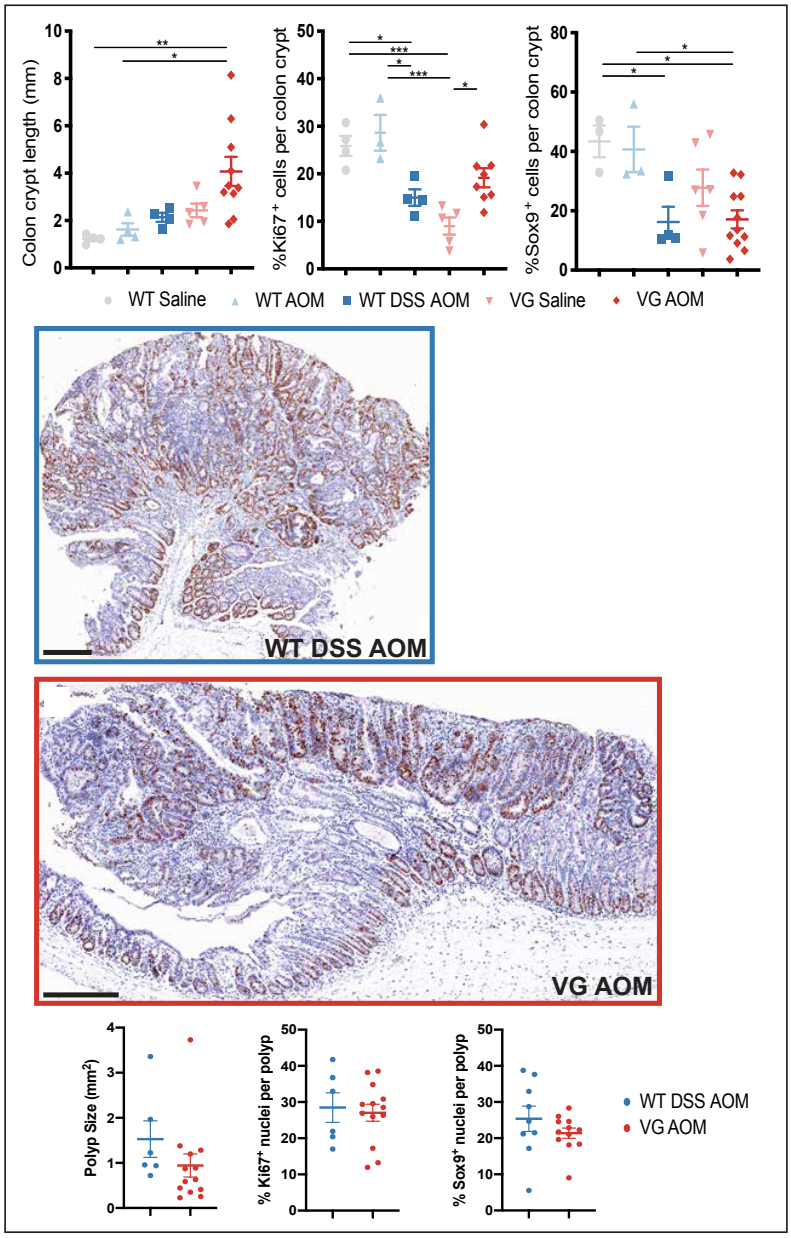


Figure 3.15: Characterisation of colon histology in wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Quantification of colon crypt length and the percentage of Ki67$^+$ and Sox9$^+$ cells per colon crypt (top panels). Points represent average per mouse, error bars represent SEM, *p < 0.05, **p < 0.01, ***p < 0.001 as calculated by one-way ANOVA with Tukey post-test. ; Representative images of Ki67 IHC staining of a WT DSS AOM (blue) and VG AOM (red) colon polyp (middle panels). Scale bars represent 250 um.; Analysis of polyp size and the percentage of Ki67$^+$ and Sox9$^+$ nuclei per colon polyp per experimental condition (bottom panels). Points represent polyps. Error bars represent SEM.

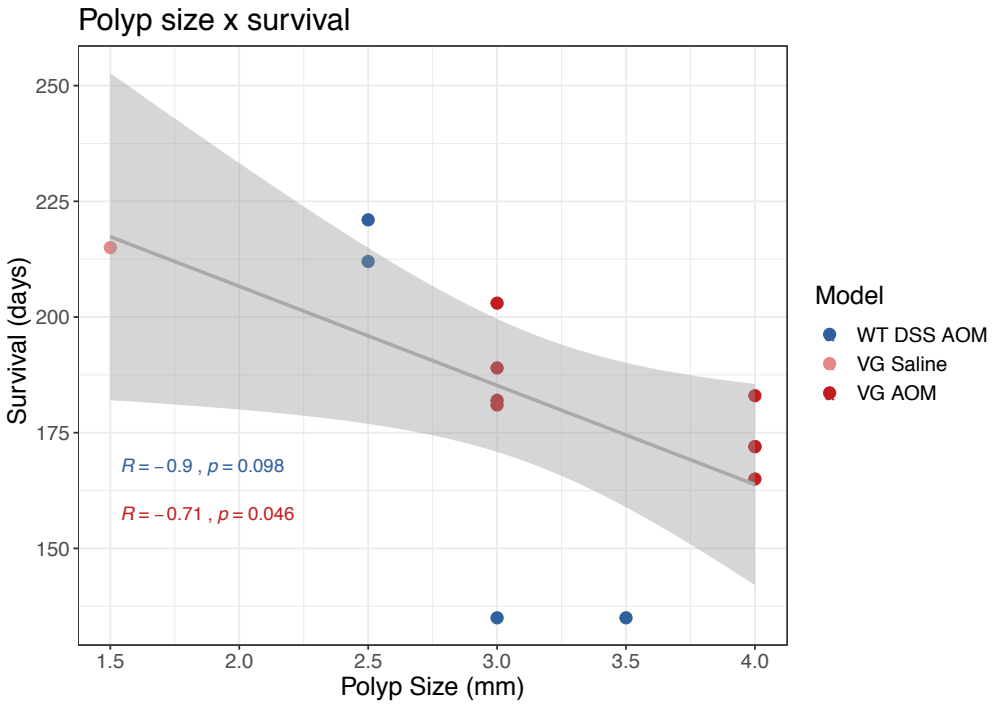


Figure 3.16: Correlation analysis of colon polyp size and survival of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Each point represents a colon polyp sampled for genomic analysis, with the age of the mouse at collection represented on the y-axis. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

3.3.3 Genomic analysis of normal and polyp samples from wild-type and *Vil1-Grem1* mice exposed to DSS and AOM

Having characterised the difference in survival and intestinal phenotype between experimental groups, and established the effect of underlying disease state on the impact of AOM within the colon, I next sought to investigate the genomic landscape of normal and polyp samples from the colons of mice within each experimental group. DNA was extracted from each sample and submitted to the High-Throughput Genomics Group at the Wellcome Centre for Human Genetics, University of Oxford (Section 2.4.3.2) for whole exome sequencing. Pre-processing, variant calling and filtering of the resultant sequencing data were carried out by Ryan Schenck, a bioinformatician within the Leedham lab (Section 2.4.3.2-6). I carried out all subsequent variant analysis using R and various publicly available resources cited hereafter.

3.3.3.1 Analysis of mutation burden

In order to determine the effect of disease state on mutation acquisition, I assessed the non-synonymous mutation burden in normal and polyp samples from mice within each experimental group (Figure 3.17). Normal samples from all experimental groups harboured relatively low non-synonymous mutation burdens, ranging from 18 to 86 mutations per sample. Interestingly, normal colon from WT Saline mice harboured an average of 64 mutations while normal colon from VG AOM mice contained an average of 20 mutations, although WT normal samples were taken from older mice as a consequence of the shortened survival of VG AOM mice. The reduction in non-synonymous mutation burden in non-dysplastic VG AOM colon samples compared to WT Saline samples thus likely reflects survival as opposed to a higher degree of stabilising selection, confirmation of which which would require sampling of normal VG AOM tissue across a greater age range. The average non-synonymous mutation burden in polyps, however, was significantly higher than in normal samples, with colon polyps from WT DSS AOM mice harbouring an average burden of 331 mutations and those from VG AOM mice harbouring a significantly increased average of 833 mutations. This demonstrates that differences in disease state correlate with differences in mutation acquisition, with AOM exposure in VG mice resulting in higher polyp mutation burdens compared to WT DSS AOM mice.

To further characterise mutation burden in colon polyps, I next analysed the relationship between polyp size and non-synonymous mutation burden and found a weakly positive correlation that did not reach statistical significance, indicating that increased mutation acquisition does not simply reflect larger polyp size (Figure 3.18).

Interestingly, eliminating disease group as a factor results in a significant positive correlation ($R = 0.5$, $p = 0.03$) between the number of non-synonymous mutations and age in normal samples that is not evident in polyp samples ($R = 0.1$, $p = 0.74$; Figure 3.19). This indicates that the number of non-synonymous mutations in normal, non-dysplastic tissue increases with age across all disease

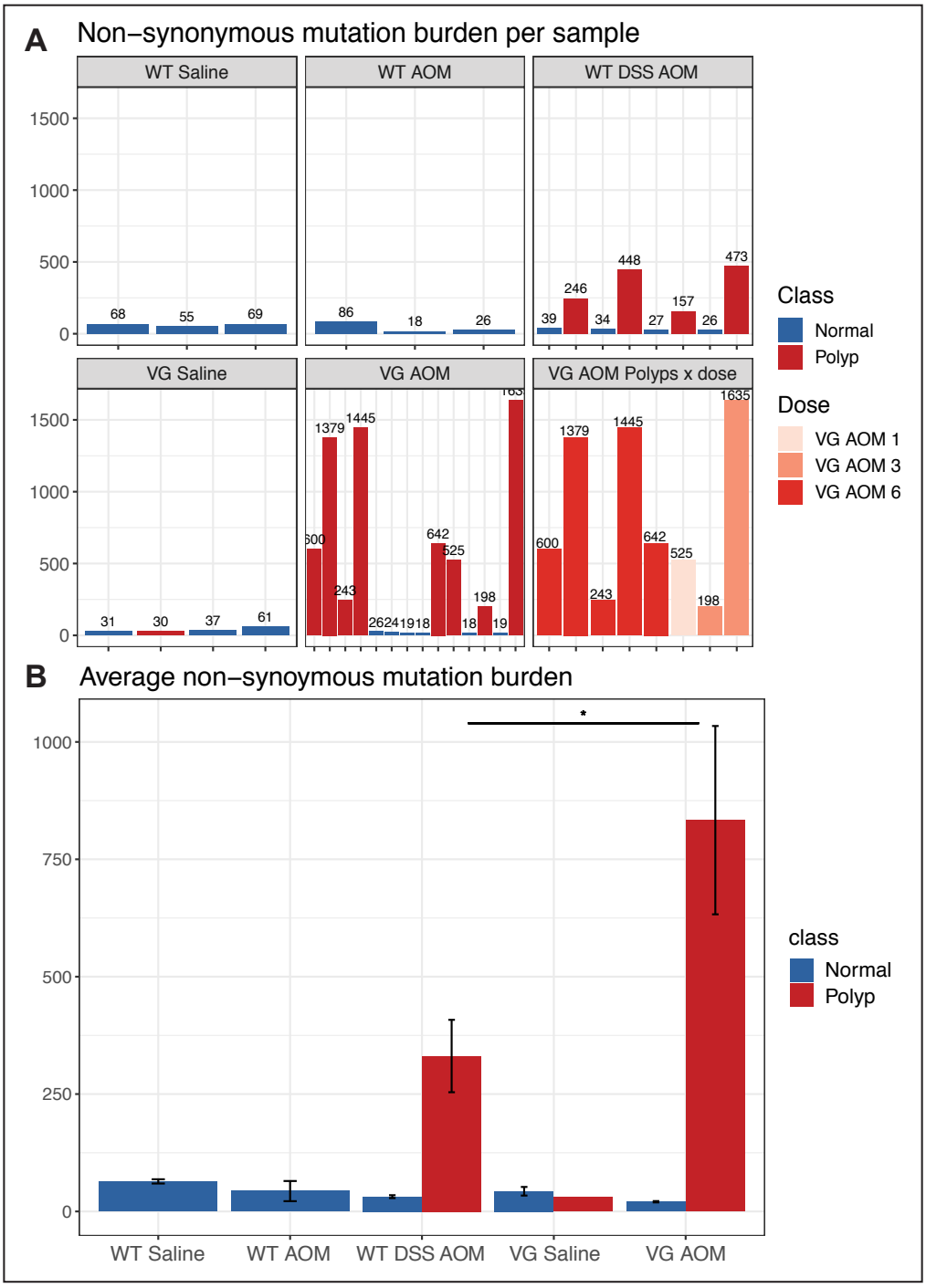


Figure 3.17: Non-synonymous mutation burden of polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as number of non-synonymous mutations per sample (A) and the average number of non-synonymous mutation burdens per experimental group (B). Error bars represent SEM, *p < 0.05 as calculated by one-way ANOVA with Tukey post-test.

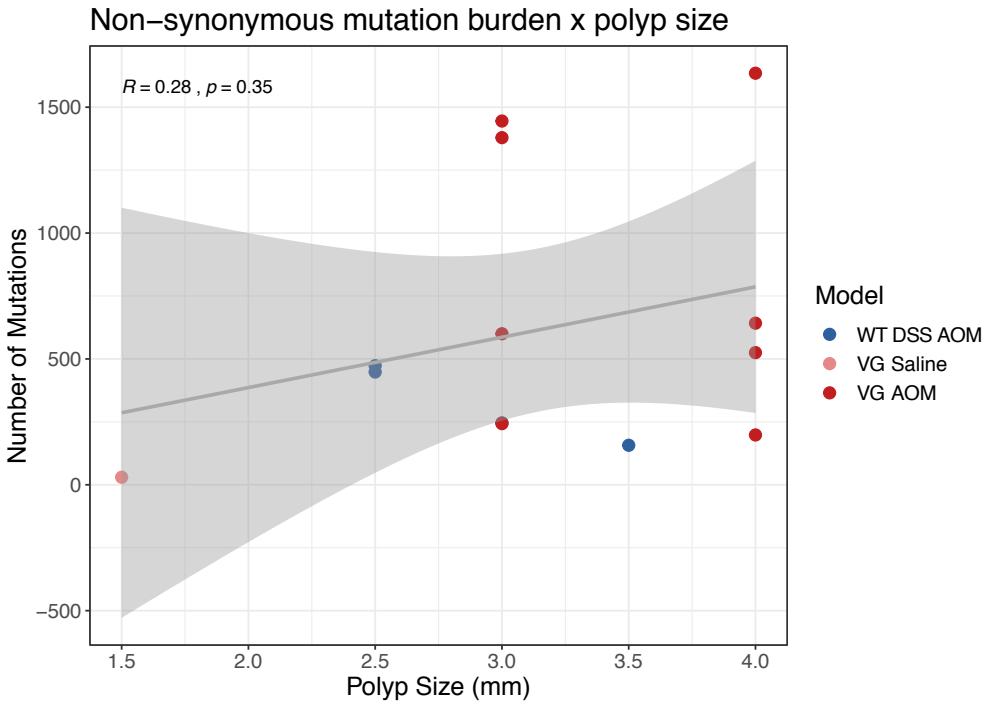


Figure 3.18: Correlation analysis of colon polyp size and non-synonymous mutation burden of wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

states, while in polyp samples this is not the case. The process of polyp formation thus involves an accumulation of mutations that dilutes the impact of largely age-associated progression observed in non-dysplastic tissue. This results in higher mutation burdens that are not correlated with age in polyps, compared with lower, strongly age-correlated burdens in normal samples. Taken together, these findings indicate that disease state impacts upon mutation accumulation following exposure to AOM and polyp formation in these animals involves an increase in mutation numbers that, in contrast to normal samples, does not correlate with age.

3.3.3.2 Characterisation of mutation types

Having confirmed a difference in polyp mutation burden between experimental groups, I next sought to investigate whether disease state influences the types of mutations present within normal and polyp samples. I therefore first characterised the mutation type contribution per normal sample (Figure 3.20; A) and found no stark differences in average mutation type contribution between experimental

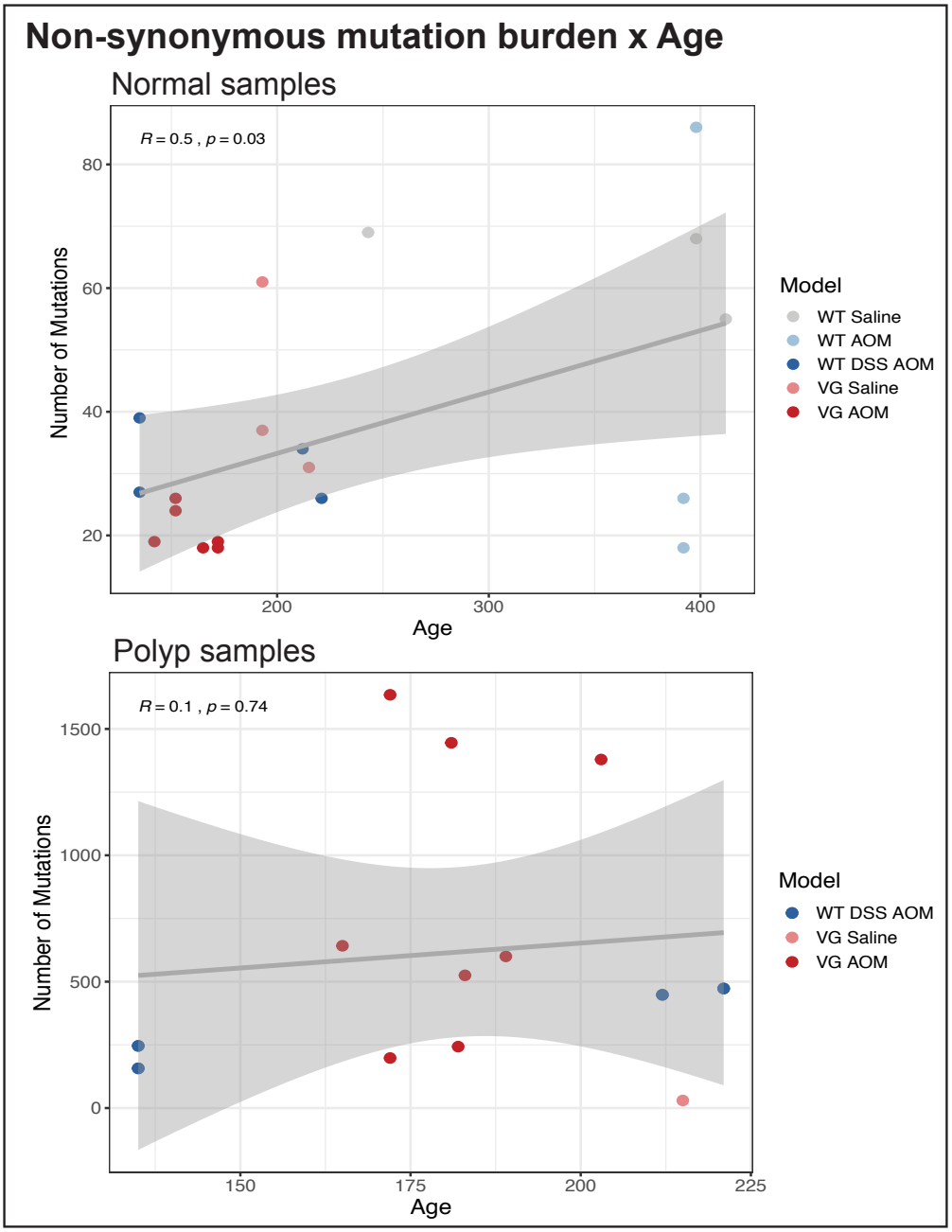


Figure 3.19: Correlation analysis of non-synonymous mutation burden and age of normal (top) and polyp (bottom) samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

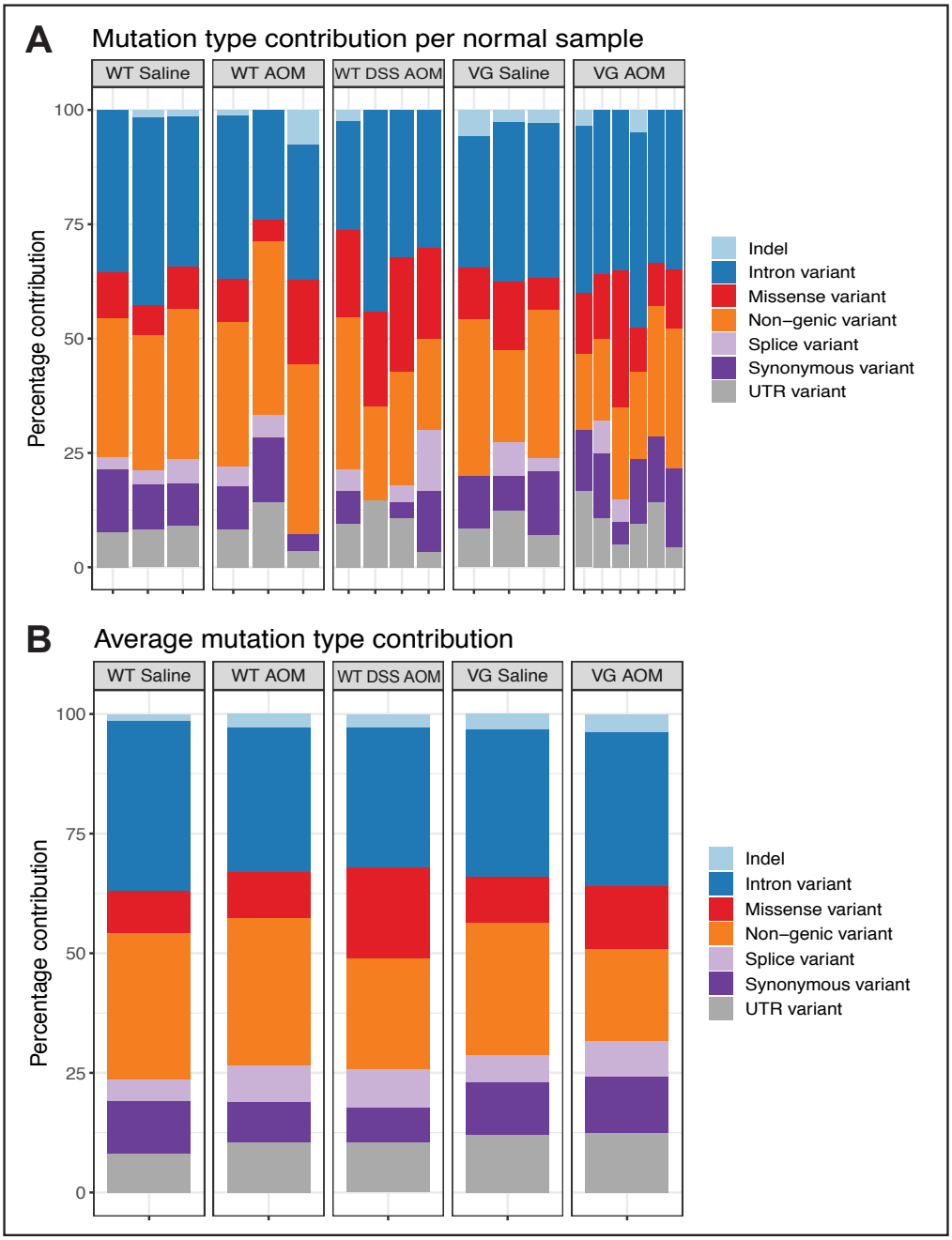


Figure 3.20: Characterisation of mutation types present in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as percentage contribution of each mutation type per sample (A) and average percentage contribution of each mutation type per experimental group (B).

groups. Normal samples from WT DSS AOM and VG AOM mice harboured slight increases in the proportion of missense variants and concurrent reductions in non-genic variants but the degree of overall variation between groups was low (Figure 3.20; B).

3. Underlying disease state impacts upon mutation acquisition

Polyp samples exhibited a slight increase in the number of mutation types due to the presence of nonsense mutations within WT DSS AOM and VG AOM polyps (Figure 3.21). However, there was no significant difference between the average contribution of each mutation type between polyps from WT DSS AOM mice compared to VG AOM mice, although both harboured increases in missense and splice variants and decreases in indels, splice and UTR variants compared to the single VG Saline polyp available for analysis. Taken together, these findings indicate that, despite a slight increase in the average contribution of missense variants in normal WT DSS AOM and VG AOM samples, there is no significant variation according to disease state in AOM-exposed, non-dysplastic colon tissue. Similarly, colon polyps from WT DSS AOM and VG AOM mice harboured a greater variety of mutation type due to the presence of nonsense variants, but there is no significant shift in mutation type contribution according to disease state within AOM-exposed colon polyps.

3.3.3.3 Mutational signature analysis

I next sought to investigate the impact of disease state on the mutational processes active within AOM-exposed normal and polyp samples. The 96-substitution classification (Section 2.4.3.8) of all mutations present within the normal and polyp samples from each experimental group are shown in Figure 3.22 and Figure 3.23, respectively. Applying this approach also enabled us to reveal distinct, validated mutational signatures, reported previously and documented in the COSMIC database, within normal and polyp samples from each experimental group. A clock-like mutational process (SBS1) contributed the majority of mutations in all normal samples, with the remainder attributed to defects in processes associated with nucleotide excision repair (NER), mismatch repair (MMR) and homologous recombination (HR) repair. Additionally, there was a weak alkylation-associated signature (SBS11) present within normal WT DSS AOM samples (Figure 3.24). The only signature unique to both VG disease groups was SBS5, a signature

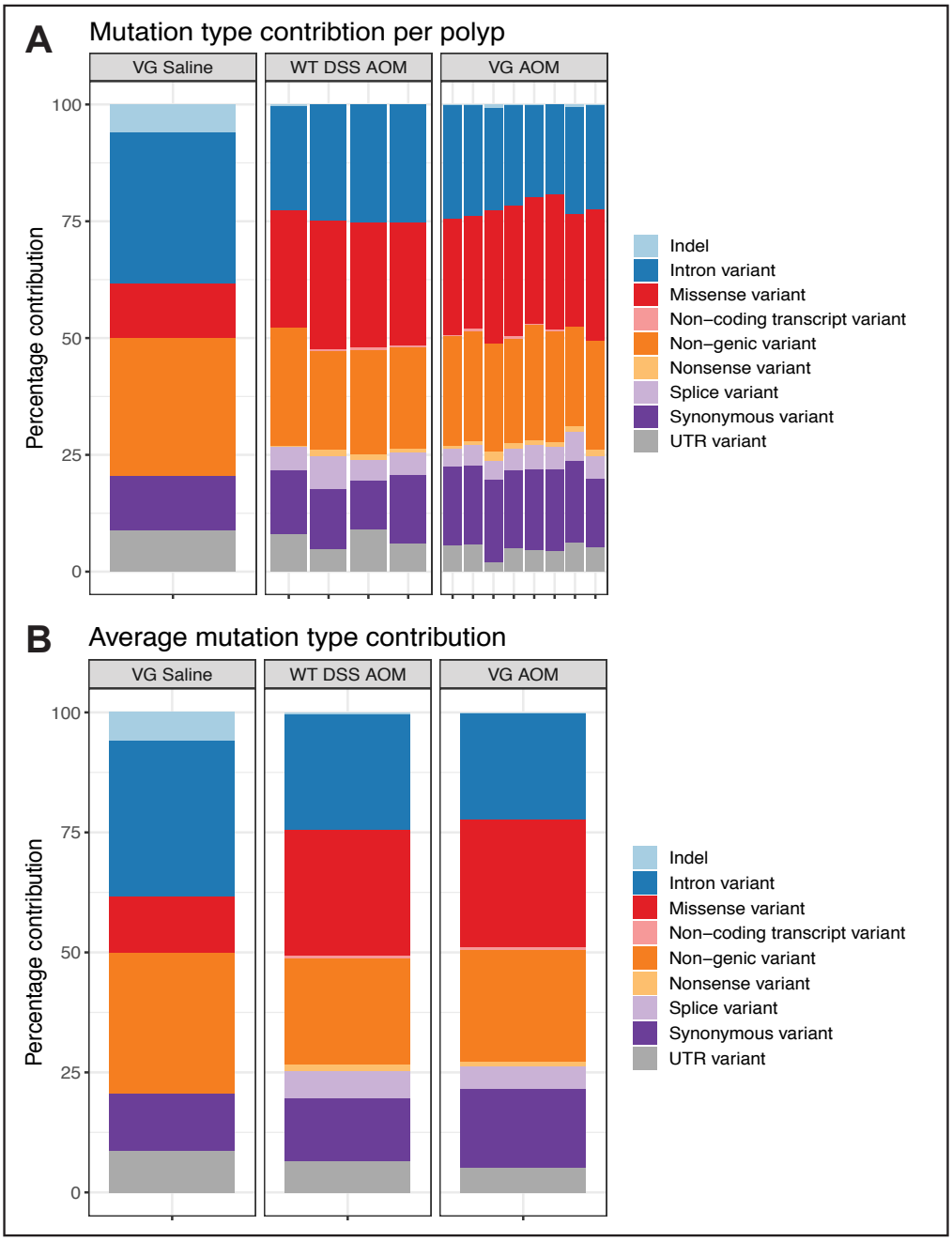


Figure 3.21: Characterisation of mutation types present in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as percentage contribution of each mutation type per sample (A) and average percentage contribution of each mutation type per experimental group (B).

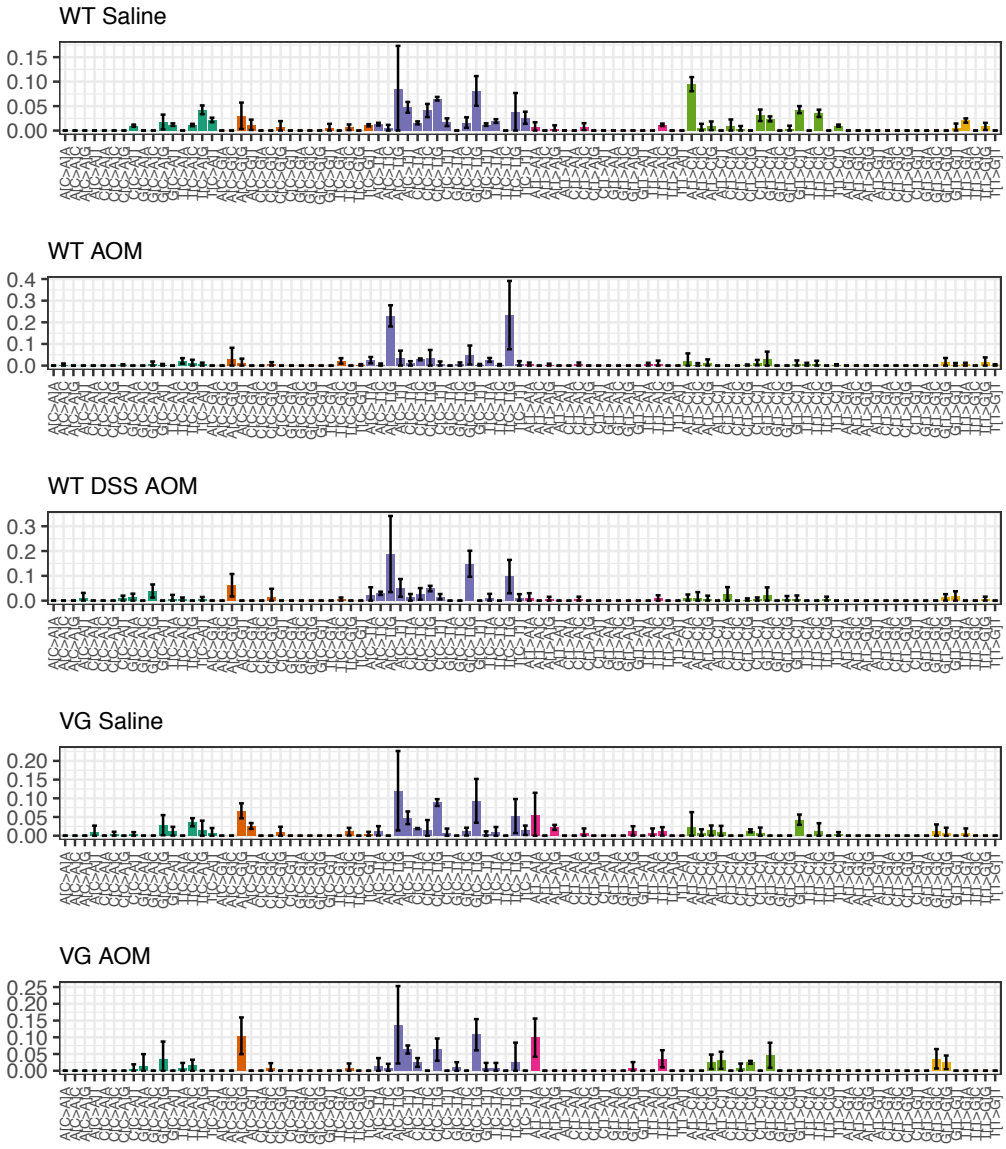


Figure 3.22: 96-substitution classification of mutations in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as percentage of all mutations per 96-substitution classification for each experimental group. Error bars represent SEM.

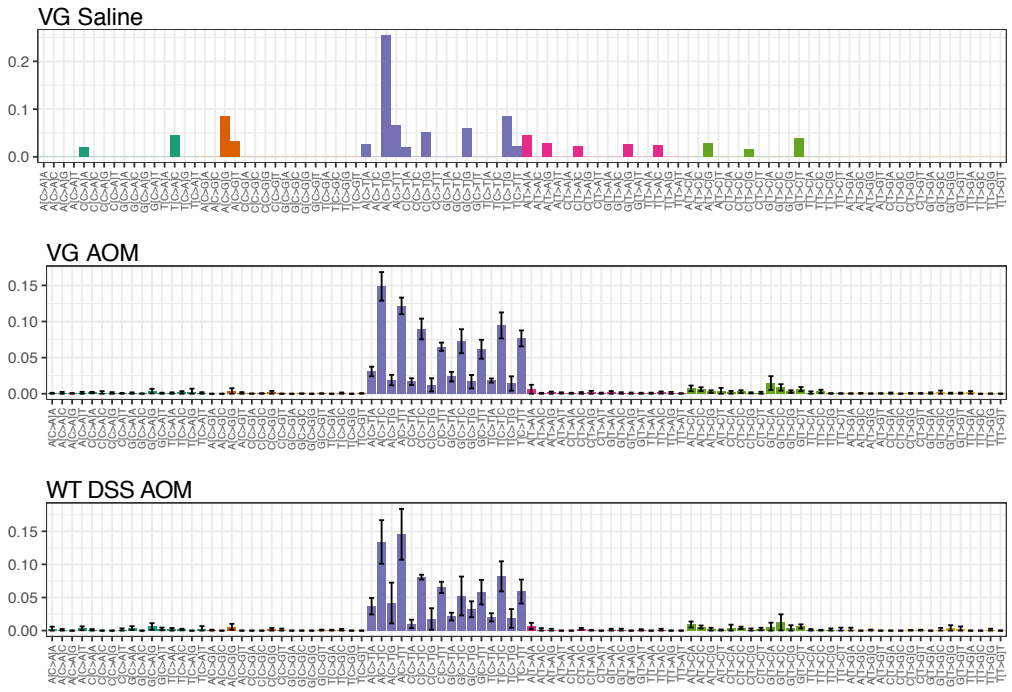


Figure 3.23: 96-substitution classification of mutations in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as percentage of all mutations per 96-substitution classification for each experimental group. Error bars represent SEM.

associated with defects in NER, while normal samples from VG AOM mice harboured the greatest variety of active mutational processes.

In polyps, 100% of the mutations within the single VG saline polyp available for analysis were contributed by a clock-like mutational process (SBS1), perhaps indicating inadvertent sampling of non-dysplastic epithelium as opposed to a true colon polyp. Contrastingly, the majority (~75%) of mutations in WT DSS AOM and VG AOM polyps were contributed by SBS11, an alkyation-associated signature (Figure 3.25).

Given that AOM is an alkylating mutagen, this suggests that in polyps from VG AOM and WT DSS AOM mice, age-associated, clock-like mutational processes are overtaken and diluted by alkylation repair-associated mutation acquisition as the predominant mutagenic driver. The remainder (~25%) of mutations in polyps from both groups were contributed by processes associated with NER and MMR, perhaps indicating a global increase in repair processes as a result of the AOM-driven increase in genome instability. Polyps from VG AOM mice

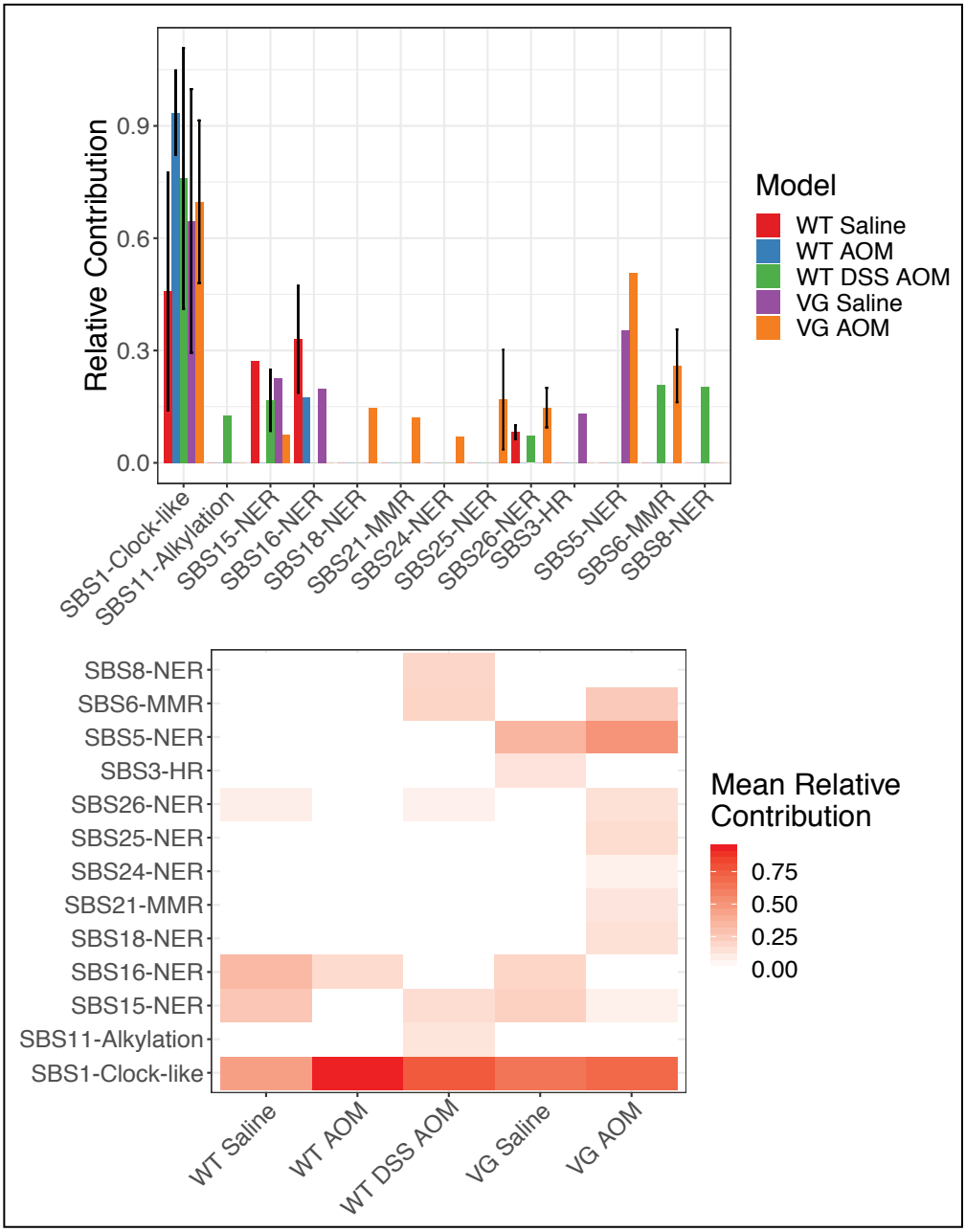


Figure 3.24: Analysis of mutational signature contribution in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as relative contribution of each detected mutational signature to all mutations per experimental group (top as bar plot, bottom as heat map). Error bars represent SEM.

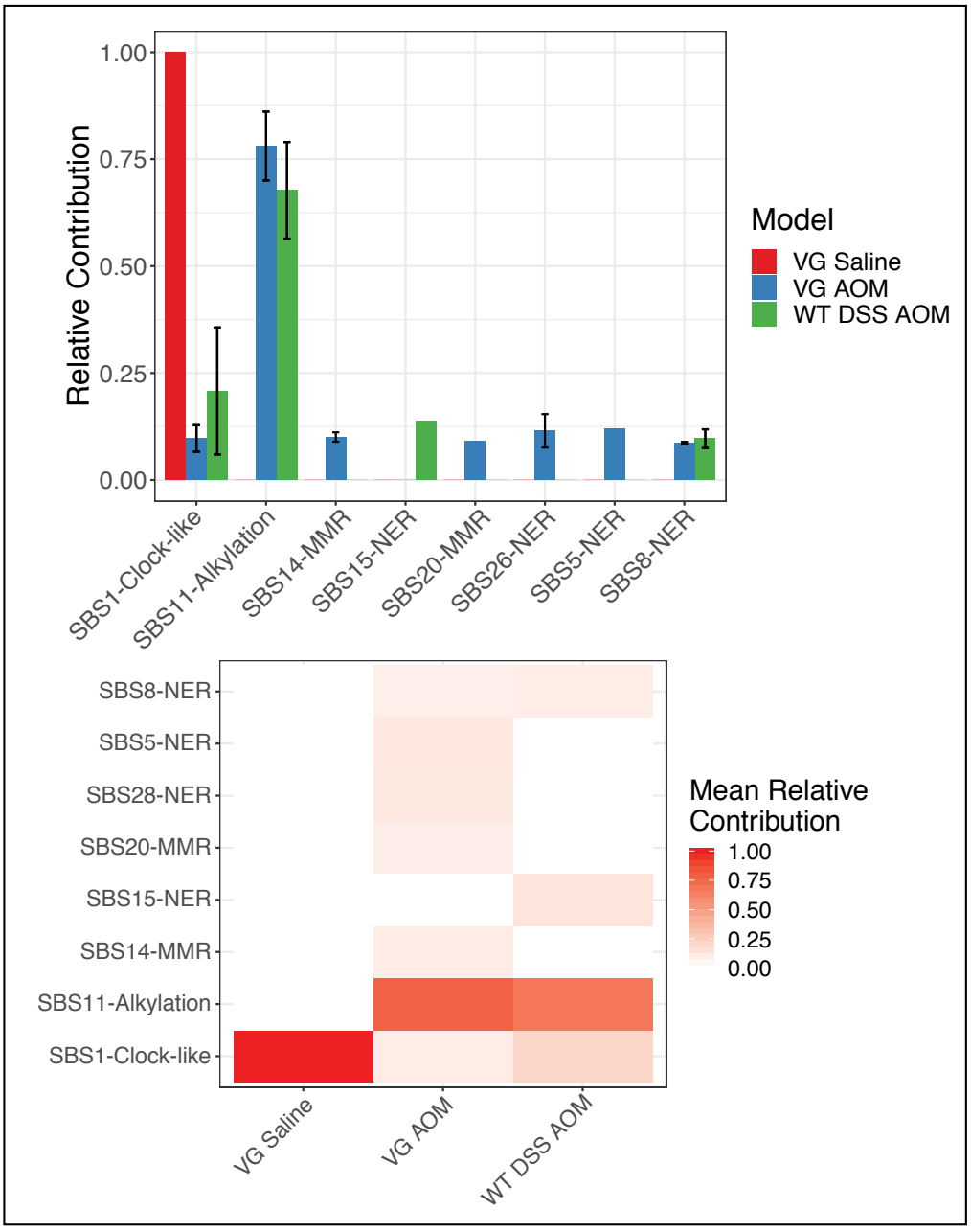


Figure 3.25: Analysis of mutational signature contribution in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Represented as relative contribution of each detected mutational signature to all mutations per experimental group (top as bar plot, bottom as heat map). Error bars represent SEM.

exhibited the greatest variety of active mutational processes compared to those from WT DSS AOM mice. Taken together, these results indicate that clock-like mutational processes contribute the majority of mutations in normal tissue, in line with our earlier observation of strong age-associated mutation accumulation in these samples. In polyps, I provide evidence of an AOM-associated mutational signature for the first time and show that this alkylation-associated process overtakes clock-like accumulation to contribute the majority of mutations in AOM-exposed polyps, with the remainder generated by NER and MMR processes that may vary according to disease state.

3.3.3.4 Characterisation of driver gene mutations

Given the differences in mutation acquisition between experimental groups described thus far, I next sought to determine whether disease state influences the acquisition of driver gene mutations in AOM-exposed normal and polyp samples. I therefore filtered non-synonymous mutations in normal and polyp samples using a set of 113 putative CRC driver genes comprised of 95 genes from the Intogen Colorectal Cancer (COREAD) list of driver genes identified from TCGA and ICGC analyses of COREAD data, along with 18 additional signalling and mismatch repair genes of interest to the Leedham lab (Section 2.4.3.7).

Interestingly, analysis of driver gene mutations in normal samples revealed that 8/13 harboured *Notch1* mutations (Figure 3.26), similar to the findings of Campbell *et al* [127], in which *Notch1* was frequently mutated in normal, non-dysplastic human eyelid tissue exposed to UV radiation. *Notch1* may therefore be similarly frequently mutated in non-dysplastic colon tissue, given that all VG AOM normal samples and at least one normal sample from every other disease group harboured a *Notch1* mutation. Conversely, *Notch1* mutations were not as frequent in polyp samples, with only 1/4 WT DSS AOM polyps and 3/8 VG AOM polyps harbouring *Notch1* mutations. *Notch1* mutation within normal murine colonic epithelium may therefore represent a functionally irrelevant driver mutation.

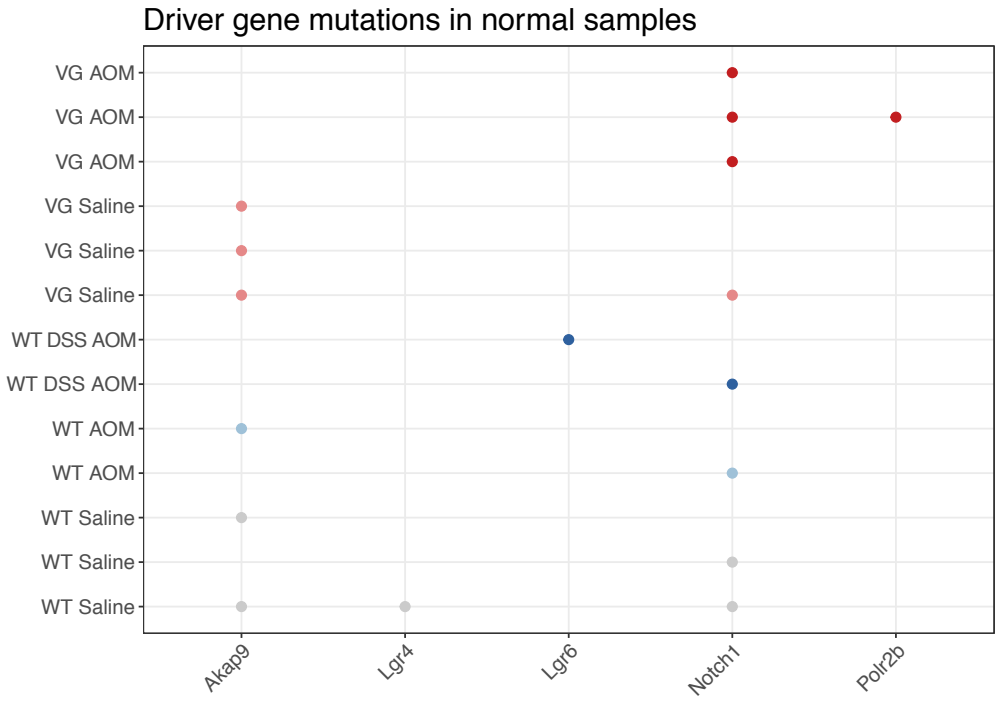


Figure 3.26: Distribution of driver gene mutations in normal samples from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Horizontal lines represent samples, vertical lines represent driver genes, dots indicate a mutation.

Analysis of driver gene mutations in polyp samples revealed that 5/8 VG AOM polyps and 4/4 WT DSS AOM polyps harboured mutations within the *Ctnnb1* gene (Figure 3.27), providing further evidence of the well-documented significance of *Ctnnb1* mutations in AOM-driven mutagenesis.

All detected *Ctnnb1* mutations were activating missense mutations within exon 3, suggesting that Wnt activation in AOM-exposed polyps is primarily through activating *Ctnnb1* mutations. Of the 3 VG AOM polyps that did not harbour an activating *Ctnnb1* mutation, 1 contained a 'stop-gained' *Apc* mutation as a presumptive alternative source of Wnt activation. Immunohistochemical analysis of WT DSS AOM and VG AOM colon polyps demonstrated the presence of areas of activated, nuclear β-catenin within polyps from both experimental groups (Figure 3.28).

Analysis of recurrently mutated driver genes within WT DSS AOM and VG AOM polyps confirmed *Ctnnb1* as the most frequently mutated gene in WT DSS AOM polyps and the second most frequently mutated in VG AOM polyps (Figure

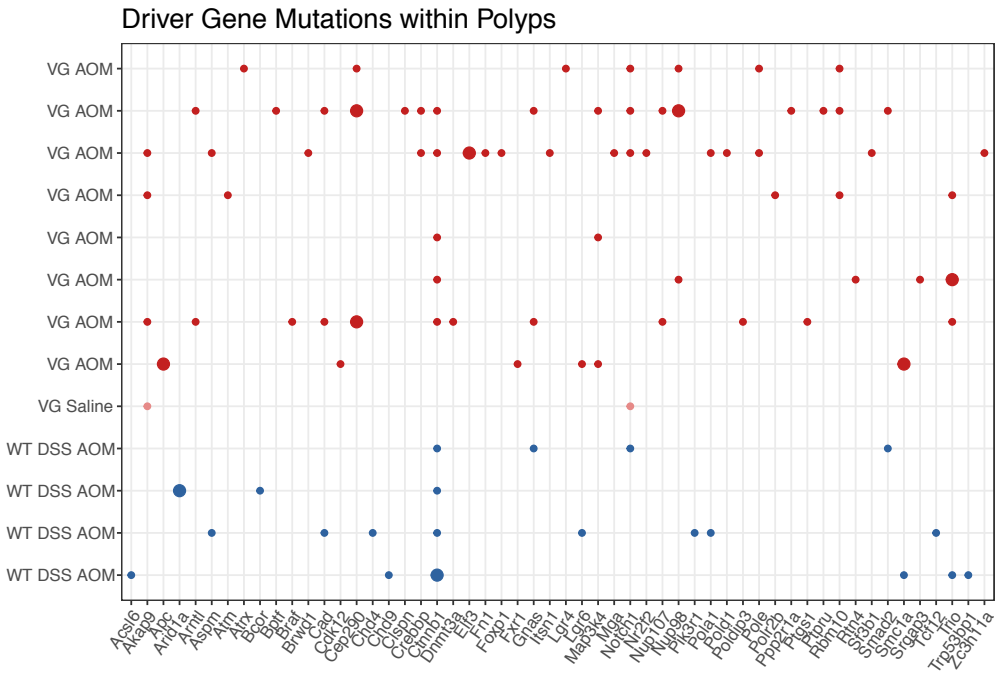


Figure 3.27: Distribution of driver gene mutations in polyps from wild-type and *Vil1-Grem1* mice exposed to AOM and DSS. Horizontal lines represent samples, vertical lines represent driver genes, dots indicate a mutation. Dot size represents the number of mutations (range 1-2).

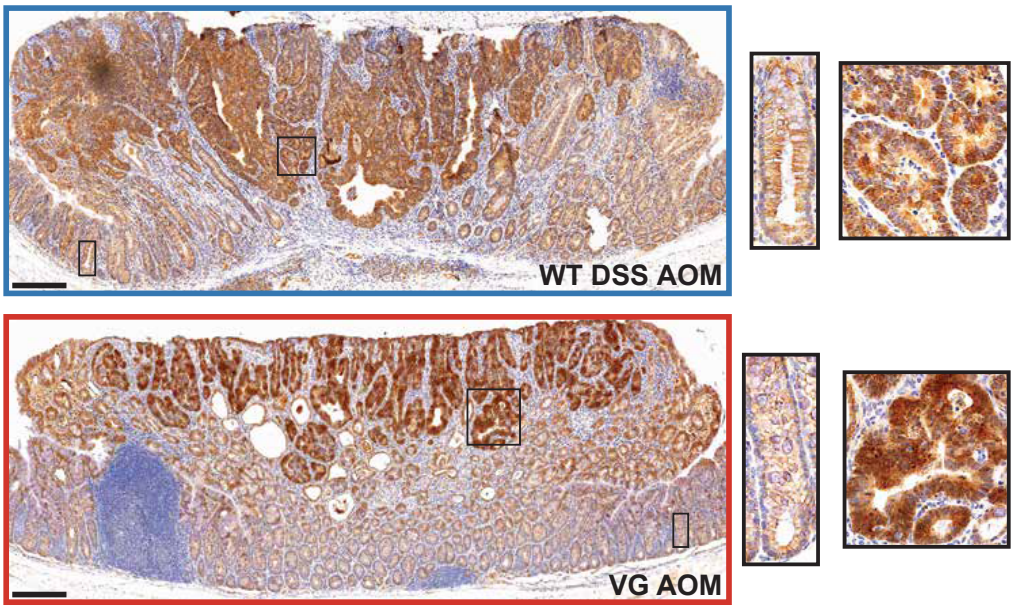


Figure 3.28: Representative images of IHC β-catenin staining in WT DSS AOM (top) and VG AOM (bottom) colon polyps. Inlays indicate membranous staining within colon crypts and nuclear staining within aberrant crypt foci. Scale bars represent 150 um.

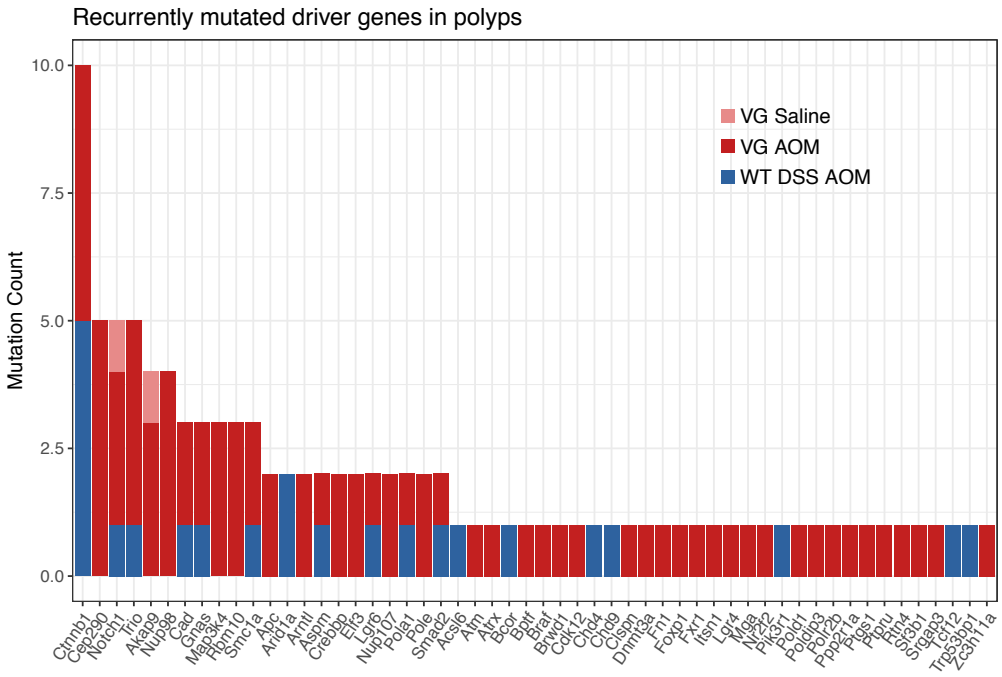


Figure 3.29: Recurrently mutated driver genes within polyps from VG AOM and WT DSS AOM mice.

3.29). However, this analysis also revealed the increased number of driver genes that were only mutated in VG AOM polyps, while a majority of those mutated in WT DSS AOM polyps were also mutated in VG AOM polyps.

In order to further investigate the effect of underlying disease state upon polyp mutations, I filtered driver gene mutations in order to identify only potentially damaging missense mutations. Comparison of the genes harbouring missense mutations between VG AOM and WT DSS AOM polyps indeed revealed a greater variety of mutated genes in VG AOM polyps compared to WT DSS AOM polyps (Figure 3.30).

Taken together, these findings indicate that differences in disease state correspond to variation in driver gene mutation and confirms the ubiquitous nature of activating *Ctnnb1* mutations to AOM-driven polyposis in both WT DSS AOM and VG AOM mice.

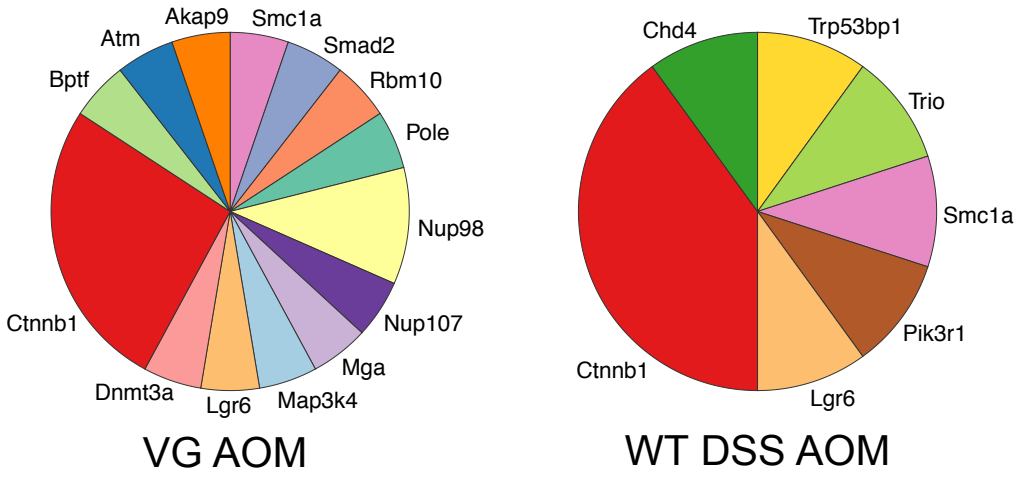


Figure 3.30: Comparison of missense mutations detected in polyps. Represented as the proportion of missense mutations per gene detected in all polyps from each experimental condition.

3.3.4 Summary of findings

- Differences in disease state proportionately impact survival and colon phenotype following exposure to AOM, with VG AOM mice exhibiting the greatest reduction in survival and most severe colon phenotype of all experimental groups

- Differences in disease state correspond to distinct histological changes within the colon following exposure to AOM, with WT DSS AOM mice exhibiting a reduction in proliferative and progenitor cells per colon crypt while, in VG AOM mice, increased colon crypt length is driven by an increase in the proportion of proliferative cells per crypt only

- I have demonstrated a link between survival and colon phenotype in VG mice exposed to AOM, such that the colonotropic effect of AOM is sufficient to overcome and accelerate the small bowel phenotype-driven reduction in survival usually observed in these mice

- Age associated mutation accumulation drives mutation acquisition in non-dysplastic tissue reglardless of AOM exposure. However, this clock-like

mutation accumulation is overtaken during polyp formation, in which exposure to AOM results in variable mutation burden regardless of age

- I have provided novel evidence associating the SBS11 mutational signature with AOM exposure in mice
- While activating *Ctnnb1* mutations provide a putative Wnt drive within AOM-exposed polyps, I have shown that differences in disease state are reflected in the variety of driver gene mutations present within polyps: VG AOM polyps harbour a greater number and variety of potentially damaging missense mutations in CRC driver genes than those from WT DSS AOM mice

Work presented in this section thus characterises for the first time the effect of AOM exposure in VG mice. I have found that underlying disease state skews the homeostatic capacity of the epithelium to cope with mutagenic insult, with different disease states impacting this ability to different degrees.

3.4 Discussion

3.4.1 Evidence of selective pressures in human colonic epithelium

The findings presented in the first section of this chapter shed light on the capacity of human colonic epithelium to 'tolerate' potentially damaging mutations within driver genes of colorectal cancer. In non-dysplastic tissue, the majority of mutations are contributed by clock-like processes, 'tolerance' of these is dependent upon the occurrence of potentially functionally damaging mutations at low VAFS. However, in samples of intestinal inflammation, clock-like mutation accumulation is overtaken by a greater variety of pathogenic mutational processes, with these samples harbouring a greater degree of potentially damaging missense mutations occurring at high VAFs. These findings are in line with those of larger scale studies that have utilised whole genome sequencing of single colonic crypts

to assess the landscape of somatic mutation within the intestine. The work of Lee-Six *et al*, mentioned previously, estimated that 1% of normal colonic crypts harbour a driver gene mutation [146], with a follow-up study by Olafsson *et al* within the same laboratory demonstrating that IBD-associated crypts exhibit an approximate doubling of this base substitution rate [148]. The work I have presented here is in line with these studies, given my finding that biopsies of inflamed epithelium harboured double the number of non-synonymous mutations compared to normal epithelium.

However, these studies of normal and IBD-associated crypts suggested that mutation accumulation occurs primarily through clock-like mutational processes, which are accelerated in IBD-associated crypts to drive increased mutation rates. In the work presented here, I have detected a far greater variety of active mutational processes within both normal and inflamed biopsies. This disparity may be a result of differences in sequencing methodology, given that the studies mentioned utilised whole genome sequencing at an average depth of 15x compared to my utilisation of targeted sequencing at an average depth of 255x.

Olafsson *et al* also demonstrated that, while clonal expansions outside of crypts are rare in normal colonic epithelium [146], inflamed tissue harbours widespread clonal expansions even in the absence of *KRAS*, *TP53* and *APC* mutations, thus indicating that clonal expansion in inflamed tissue occurs even prior to acquisition of mutations in these key driver genes [148]. Interestingly, exome and targeted sequencing of driver genes in colitis-associated cancers (CACs) by Fujita *et al* in 2018 revealed that somatic mutation of *RNF43* is associated with chronic IBD and is present in 10% of colitis-associated cancers (CACs) [151]. Furthermore, their analysis of *RNF43*-mutated CACs showed elevated expression of oncogenic genes, thus providing evidence supporting *RNF43* as a bona fide driver of CAC development. Somatic mutation of *RNF43* may thus provide the Wnt drive in inflammation-associated neoplastic lesions independently of the well characterised APC-initiated tumourigenic pathway. Detection of *RNF43* mutations occurring at high VAFs within a subset of our inflamed samples

may thus represent the detection of an early marker of inflammation-associated neoplasia. Given that these samples also harboured an increase in the number and variety of potentially damaging missense mutations in CRC driver genes, loss of heterozygosity of *RNF43* may represent an early marker of inflammation-associated increases in clonal expansion and, potentially, subsequent CAC.

The work presented here is, however, limited by several caveats. MIP sequencing is a relatively new method and the results presented here thus require confirmatory validation via an orthogonal sequencing method. The studies mentioned above were published during the time of experimentation, thus rendering validatory work a lower financial priority. The limited number of patients within each group also reduces the power of the analysis presented here. Furthermore, the technical requirement for pooled crypts isolated from intestinal biopsies introduces uncertainty given variability in the number of crypts per biopsy. Further work utilising high depth sequencing in combination with single crypt resolution will therefore shed more light on mutation dynamics within normal and inflamed colon, given that the VAF and mutation signature analysis presented is inherently limited by our ability to detect small clone sizes, as determined by sequencing depth. The lack of single crypt sampling and subclonal sequencing depth resolution therefore does not allow us to address questions regarding the origin of the mutations detected in this chapter: are they single mutations across many crypts that arose together and are not fixed? Or, are these mutations that swept through a single crypt but no further? Improving clonal resolution through spatially defined sampling and increased sequencing depth could thus address these questions and provide subclonal resolution to this analysis. Given that the findings of this chapter do not rely upon subclonal resolution, the effect of this caveat is limited for the analyses I have presented here. However, further work I have outlined here that could shed light on the earliest stages of transformation within the gut would benefit from the reliable detection of small clone sizes. Additional analysis of copy number variation

and methylation will further reveal the impact of differences in intestinal disease state on genomic integrity.

3.4.2 Evidence of selective pressures in mouse colonic epithelium

Work presented in this chapter characterised the differential effect of underlying disease state on mutation acquisition. Interestingly, the impact of AOM upon colon phenotype varies according to disease state, with AOM exposure in WT DSS AOM mice associated with a reduction in proliferative and progenitor cells, but AOM exposure in VG AOM mice associated with an increase in proliferative cells per colon crypt. This difference is in line with the nature of each intestinal condition, whereby WT DSS AOM mice represent a model of regenerative colitis while VG AOM mice harbour a constitutively expanded proliferative cell compartment. These differences in the mutationally susceptible cells expanded within each disease state may thus explain the increased number of colon polyps and reduction in survival of VG AOM mice compared to WT DSS AOM mice.

Interestingly, the observation that the severity of disease phenotype correlates with increased non-synonymous mutation burden was reversed within normal tissue, whereby non-dysplastic samples from VG AOM and WT DSS AOM mice harboured fewer non-synonymous mutations than untreated mice. Whether these differences in mutation number are simply age-related, as I have suggested given the earlier sampling of normal tissue from animals with more severe disease phenotypes, or whether they are evidence of a greater degree of stabilising selection in VG AOM and WT DSS AOM mice represents an interesting avenue for future work. It may be that the increased number of proliferative cells within the colonic crypts of VG AOM mice result in mutated cells with neutral or disadvantageous mutations being rapidly out competed through niche succession. Despite this resulting in a lower number of mutations overall, it would also increase the likelihood of selectively advantageous mutations occurring that

result in polyp formation, thus underpinning the observable increase in polyp numbers in VG AOM mice. These possibilities could be explored through modelling of crypt dynamics according to different disease state conditions, which would enable manipulation of the number of mutation susceptible cells and the likelihood of a mutated cell achieving fixation within crypts.

VG AOM polyps also harboured significantly more mutations than WT DSS AOM polyps. Referencing our proposed model of the genotype-phenotype axis in intestinal disease, the difference in mutation acquisition between VG AOM and WT DSS AOM mice could thus be mechanistically driven by either disease-specific shifts in the cell of origin and/or distinct microenvironmental disruption of homeostatic mutation protection mechanisms. Given that previous work has shown that aberrant *Grem1* expresssion increases the number of putative *Lgr5*-stem/progenitor cells [51], and I have demonstrated expansion of the colonic proliferating cell population here, this raises the intriguing notion of different cells of origin harbouring distinct optimal mutational spectra. There is clinical evidence of this in humans, with the formation of morphologically distinct colonic polyp subtypes thought to occur via conventional or non-conventional mutational pathways that are dependent upon the cell of origin. Conventional pathways involve the acquisition of *APC* or *CTNNB1* mutations within *LGR5*-positive stem cells at the crypt base, while non-conventional pathways involve the acquisition of different activating mutations within non-*LGR5* or ectopic crypt progenitor cells (Section 1.5.3).

It could also be the case that the specific perturbation of intestinal signalling pathways within each disease state impacts differentially upon mutation protection mechanisms. This would result in distinct acquisition of mutations that is reflected in differential mutational spectra within tissue from each disease state. Given that I have shown that mutations in VG AOM polyps were contributed by distinct mutational processes that resulted in the acquisition of a greater variety of driver gene mutations, it may thus be that constitutively deranged BMP signalling

impacts upon mutational protection mechanisms to drive distinct patterns of mutation acquisition in these mice.

The work presented here may, however, have been limited by mouse numbers within each experimental group limiting the number of polyps available for genomic analysis, in addition to the potential for error in the identification and sampling of colonic polyps. Future work could incorporate increased sampling of non-dysplastic colon from each group in order to shed light on the dynamics of mutation selection prior to neoplastic transformation. Additionally, multi-regional sequencing of polyps from each experimental group could further reveal the impact of disease state on mutation timing, given the limited ability of whole polyp exome sequencing to reveal whether mutations are polyclonal or monoclonal in origin.

3.4.3 Conclusion

Taken together, the findings presented in this chapter suggest that underlying disease state influences the acquisition and persistence of mutations within colonic epithelium. In humans, I have demonstrated that driver gene mutations are present even in non-dysplastic crypts and that the accumulation of these mutations appear to vary according to underlying disease state, with inflamed epithelium exhibiting frequent *RNF43* mutations occurring at high VAFs, indicative of potential loss of heterozygosity, and an increased variety of accompanying driver gene mutations. Investigation of somatic mutation within murine colonic epithelium from different disease states following exposure to AOM further indicated that differences in intestinal disease state are associated with differential responses to mutagen exposure. This raises interesting questions regarding the role of cell-extrinsic, microenvironmental factors in mutation protection mechanisms, given that in both human and mouse models of intestinal disease, mutations occur to a greater degree in tissue from more severe intestinal phenotypes. In VG mice, BMP signalling may thus not only impact upon cell fate

decisions, as evidenced by the generation of ectopic crypts, but also upon cellular responses to mutation, previously considered a cell intrinsic property determined by position along the crypt-villus axis (Section 1.4.4.1). This intriguing notion is explored in the next chapter.

"Forget about your house of cards" - Radiohead

4

Cell intrinsic mutation susceptibility in *Vil1-Grem1* mice

Contents

4.1 Introduction 114
4.2 Results 116
4.2.1 IHC analysis of DDR pathway markers within the ectopic crypts of *Vil1-Grem1* mice 116
4.2.1.1 DNA damage sensing within ectopic crypts . . 117
4.2.1.2 DNA damage translation within ectopic crypts 118
4.2.1.3 Activation of apoptotis within ectopic crypts . . 120
4.2.2 Increased miR-96 expression in the SB1 villus region of VG mice 124
4.2.3 Summary of findings 125
4.3 Discussion and Future work 128
4.3.1 Discussion 128
4.3.1.1 Out with the old axis, in with the new 130
4.3.1.2 Limitations 131
4.3.1.3 Future Work 131

4.1 Introduction

Having shown in Chapter 3 that differences in intestinal disease state influence mutation acquisition, I next sought to specifically investigate somatic mutation susceptibility in VG mice. As discussed earlier (Section 1.4.5), given the variety of possible genotoxic insults, ranging from UV radiation to harmful radical oxygen species, there is a broad repertoire of responses to the specific kinds of DNA damage induced by different types of agents. DDR research is therefore a broad and quickly evolving field, comprehensive overview of which is beyond the remit of this thesis. I have chosen to stratify the numerous DDR pathways that have been described to date through a lens of genomic instability, separating repair processes into stringent pathways that result in accurate, effective repair of DSBs and those that are less effective and therefore contribute to genomic instability:

- Homologous Recombination (HR) represents a high-fidelity repair pathway, where a template strand primarily available during the S and G2 stages of the cell cycle is utilised to ensure sequence accuracy of the repaired strand (Section 1.4.2.1)

- Non-Homologous End Joining (NHEJ) represents faster but somewhat less accurate repair involving template-free end-resection and ligation that can introduce sequence inaccuracies and insertions and deletions (Section 1.4.2.2)

- Alternative End Joining (AEJ) pathways are relatively recently characterised and thought to represent backup repair processes, available to all cells at all times. AEJ repair is associated with frequent, potentially large-scale errors and is thus thought to contribute significantly to genomic instability through inaccurate repair of DNA damage (Section 1.4.2.3)

Given this broad variation in outcomes, cells employ different DDR pathways in a context-dependent manner. Within the intestine, DDR has been shown to vary along the crypt-villus axis. Stem cells within crypts preferentially empoy

stringent HR repair to preserve the genomic integrity of actively dividing cells within this compartment. Proliferative cells within the transit amplifying zone utilise fast, efficient NHEJ repair given that they are also protected by high rates of cell turnover. Finally, post-mitotic differentiated cells exhibit minimal activation of DDR pathways given their 'soon to be shed' position further up the villus (Section 1.4.4.1).
Given that the crypt-villus axis is maintained by tightly regulated morphogen gradients, it is thus unsurprising that numerous studies have also linked signalling pathways to activation of the DNA damage response. Head and neck squamous cell carcinoma (HNSCC) cells negative for TGF-β signalling activity exhibit compromised HR-directed DNA repair [104], while reduced IL-22 signalling within intestinal epithelial cells leads to suppression of DNA damage-induced apoptosis [49].

I thus surmised that the constitutively deranged intestinal morphogen gradients caused by aberrant epithelial *Grem1* expression in *Vil1-Grem1* mice would also compromise DDR along the crypt-villus axis. In health, proliferative cells are normally restricted to the transit amplifying compartment of the intestinal crypt. Our lab has previously shown that aberrant epithelial *Grem1* epression in VG mice results in expansion of a putative stem/progenitor compartment, with cell proliferation occurring within ectopic crypt foci situated within the villus of the small intestinal epithlieum [51]. Cells within ectopic crypts are thus actively proliferating within the genotoxic environment of the intestinal lumen. Given that results from Chapter 3 indicate that VG mice are more susceptible to mutagenesis, I hypothesised that this is underpinned by Grem1-driven impairment of DDR within ectopic crypts, resulting in increased genomic instability within this compartment. The overall aim of this chapter was therefore to characterise intestinal DDR pathways in *Vil-Grem1* mice following irradiation, and propose a mechanism by which aberrant epithelial Grem1 results in increased mutation susceptibility. In order to explore this, I took an *in vivo* approach, exposing

wild-type and *Vil1-Grem1* mice to ionising radiation (IR) prior to characterising DDR pathways within each intestinal compartment. Specific aims included:

- To determine whether aberrant epithelial *Grem1* expression results in an altered DDR in VG mice
- To characterise altered DDR pathways within the ectopic crypt compartment of VG mice
- To propose a mechanism by which altered *Grem1* expression impairs the DDR

4.2 Results

4.2.1 IHC analysis of DDR pathway markers within the ectopic crypts of *Vil1-Grem1* mice

I first sought to specifically characterise the expression of DDR-associated proteins within ectopic crypts via immunohistochemical analysis. This 'on-slide' approach allows for direct comparison of the response to irradiation of cells within ectopic crypts and those within the basal crypt niche. I thus characterised the expression of protein markers that correspond specifically to the sensing, translating and effecting stages of DDR, as outlined in Table 4.1, within the intestinal epithelial compartments of WT and VG mice.

Mice were irradiated and sacrificed at various time points post-irradiation, as described earlier. Initial experiments that analysed γH2AX, p-p53 and caspase-3 expression utilised one mouse sacrificed at each timepoint of 0 h, 2 h, 5 h, 8 h and 24 h post-irradiation; for later analysis of pATM, RAD51 and BAG1 expression, three mice per timepoint were sacrificed at 0 h, 2 h and 5 h post-irradation. The intestinal tract was collected (Section 2.2.6) and sections generated for IHC analysis (Section 2.6.1) of DDR marker expression. Quantification of the epithelial cell expression of γH2AX, p-p53, caspase-3 per intestinal compartment was carried out by Dr Viktor Koelzer using the HALO imago analysis platform (Section

Sensing	*Phosphorylation of these proteins corresponds to the earliest, 'sensing' step of the DNA damage response*
γH2AX	Phosphorylation of the histone H2AX at residue Ser139 is one of the earliest steps in the initiation of DNA repair [70]
p-ATM	ATM is recruited to break sites by the MRE11-RAD50-NBS1 (MRN) complex [152] and undergoes auto-phosphorylation at Ser1981 in the vicinity of DSBs
Translating	*Subsequent recruitment of DDR machinery and activation of response pathway proteins corresponds to 'translation' of the DNA damage signal to a cellular response*
RAD51	RAD51 is a DNA binding protein that is central to the HR DNA repair pathway and is recruited to DSBs in order to coordinate homology-directed DNA repair according to a template strand [153]. Following ionizing radiation, cytoplasmic RAD51 is redistributed to discrete nuclear DNA damage-induced foci to coordinate stringent DDR pathways [154]
p-p53	The 'translation' step also involves phosphorylation of the tumor suppressor protein p53 at Ser15 [98]. This leads to reduced interaction between p53 and its negative regulators, thereby promoting the accumulation and activation of p53 in response to DNA damage.
Effecting	*The final stages of DDR involve an activation cascade of caspase proteins, which are responsible for the execution of apoptosis within cells harbouring DNA damage*
caspase-3	As part of this cascade, cleavage of caspase-3 converts it into its active form and thus enables execution of its cell death function [98]
BAG1	BAG1 is an anti-apoptotic protein that markedly increases the anti-cell death function of BCL2 and represents a link between signalling pathways and anti-apoptotic mechanisms [155]

Table 4.1: Analysed proteins corresponding to the three stages of DDR

2.6.2.3); I carried out the equivalent quantification of pATM, RAD51 and BAG1 expression using QuPath image analysis software (Section 2.6.2.2). I utilised R for all subsequent analysis of expression data.

4.2.1.1 DNA damage sensing within ectopic crypts

There was a rapid increase in the percentage of γH2AX$^+$ nuclei within the crypt and villus compartments of both WT and VG mice at 2 h post-irradiation (Figure 4.1, A). The ectopic crypt compartment of VG mice, however, exhibited a striking absence of accumulation of γH2AX$^+$ cells, harbouring only 6% γH2AX$^+$ cells at 2 h post-irradiation, versus 30% observed at the crypt base (Figure 4.1; B).

Characterisation of baseline pATM levels in non-irradiated mice (Figure 4.2; A) showed that both WT and VG crypts harboured high expression of pATM,

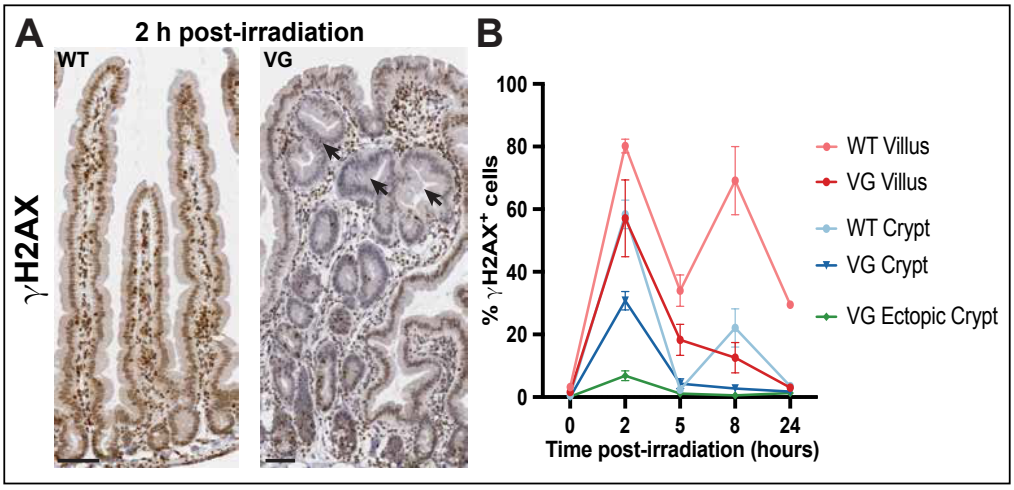


Figure 4.1: Analysis of γH2AX expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of γH2AX IHC staining of SB1 of the indicated experimental mouse 2 h post-irradiation. Arrows indicate reduced expression within ectopic crypts, scale bars represent 50 μm. B) Quantification of the percentage of γH2AX$^+$ cells per indicated compartment at each timepoint post-irradiation. Error bars represent SEM of 5 technical repeats per compartment, N = 1 mouse per timepoint.

as expected given the greater proportion of proliferating cells within these compartments, while pATM levels within WT and VG villi remained low. The ectopic crypt compartment exhibited a significant reduction in baseline pATM$^+$ nuclei, with an average of 13% of ectopic crypt pATM$^+$ cells versus 37% at the crypt base. In terms of response to IR, ectopic crypt cells exhibited an increase in the percentage of pATM$^+$ nuclei 2 h post-irradiation that was also observed in all other compartments. Ectopic crypt cells therefore exhibit reduced levels of pATM$^+$ nuclei at baseline and reduced H2AX phosphorylation at 2 h post-irradiation that, taken together, indicate an impairment of the 'sensing' stage of DNA repair.

4.2.1.2 DNA damage translation within ectopic crypts

Activation of p53 was also reduced within ectopic crypts, which harboured 22% p-p53$^+$ cells 2 h post-irradiation versus 72% at the crypt base (Figure 4.3). Figure 4.3 also illustrates the previously described (Section 1.4.4.1) difference in DDR outcomes along the crypt-villus axis, with both VG and WT crypt compartments activating p53 in response to radiation compared to reduced p-p53 levels within their respective villus compartments. This finding indicates that

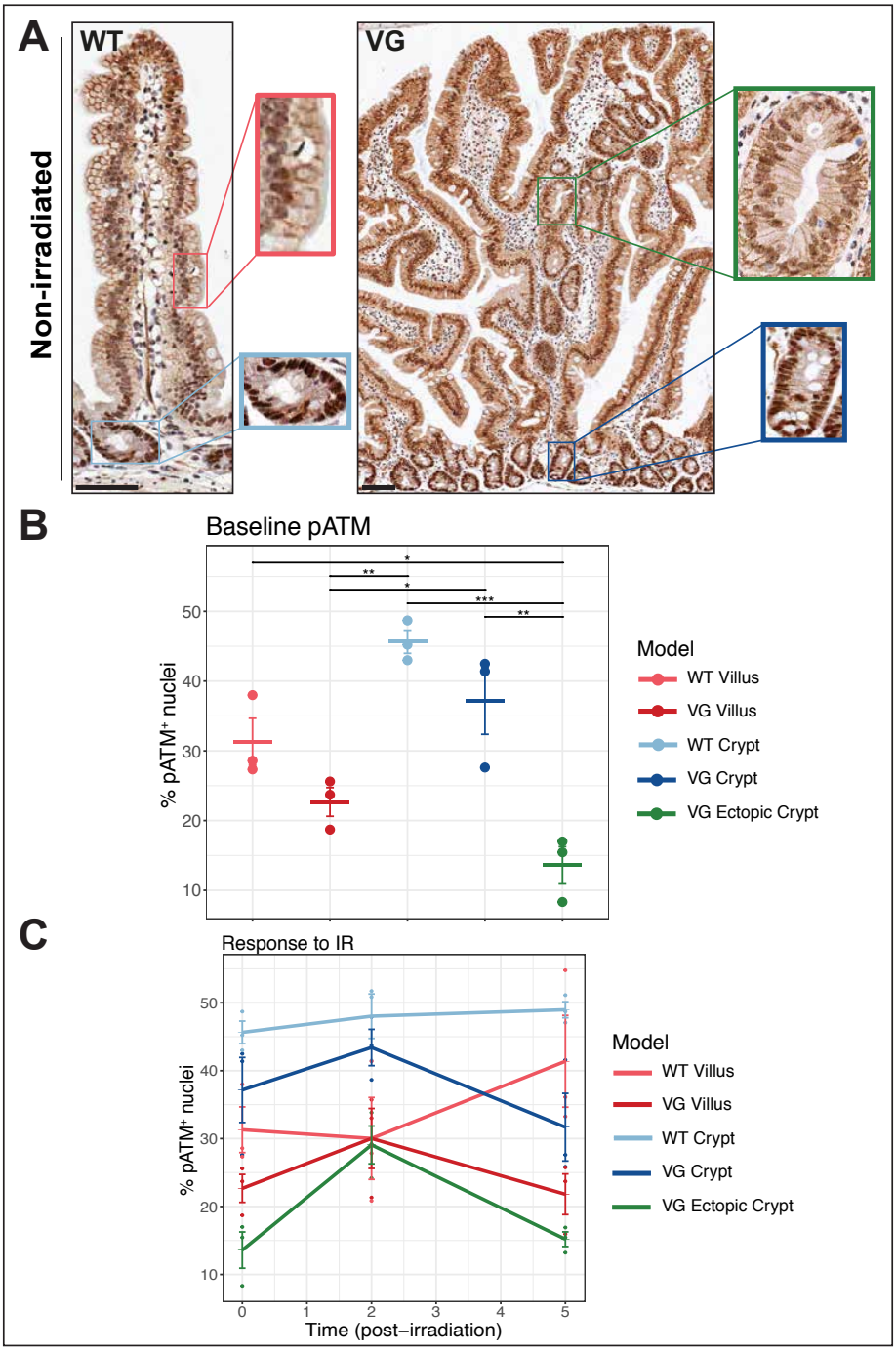


Figure 4.2: Analysis of pATM expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of baseline pATM IHC staining in SB1 of the indicated experimental mouse. Scale bars represent 50 μm, inlays of WT Villus (pink), WT Crypt (light blue), VG Crypt (dark blue) and VG Ectopic Crypt (green) compartments, respectively. B) Quantification of the percentage of pATM$^+$ nuclei per indicated compartment at baseline. Error bars represent SEM, N = 3 mice, *p < 0.05, **p < 0.01, ***p < 0.001 as calculated by one-way ANOVA with Tukey post-test. C) Quantification of the percentage of pATM$^+$ nuclei per indicated compartment following irradiation. Error bars represent SEM between N = 3 biological replicates, each comprising 30 technical replicates.

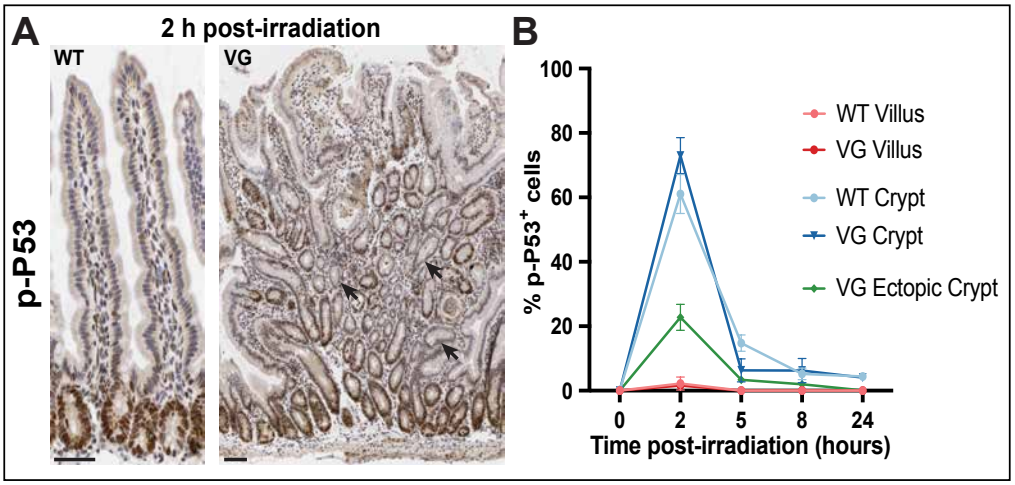


Figure 4.3: Analysis of phospho-p53 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of p-p53 IHC staining of SB1 of the indicated experimental mouse 2 h post-irradiation. Arrows indicate reduced expression within ectopic crypts, scale bars represent 50 μm. B) Quantification of the percentage of p-p53$^+$ cells per indicated compartment at each timepoint post-irradiation. Error bars represent SEM of 5 technical repeats per compartment, N = 1 mouse per timepoint.

cells within ectopic crypts harbour an impaired ability to activate p53 following IR compared to those at the crypt base.

The percentage of RAD51$^+$ cells in non-irradiated ectopic crypts was markedly lower than both WT and VG crypts (Figure 4.4). However, there was no difference in RAD51 protein expression following irradiation in VG villi compared to all other compartments, with the percentage of RAD51$^+$ cells remaining steady or slightly reduced in all compartments at 2 and 5 h post-irradiation (Figure 4.4; C). These findings, taken together, suggest a reduction in the availability of RAD51 within ectopic crypts at baseline compared to all other compartments.

4.2.1.3 Activation of apoptotis within ectopic crypts

The translation of DNA repair signals to produce an apoptotic response, indicated by IHC staining of cleaved caspase-3, was reduced and delayed in the VG ectopic crypt compartment, which peaked at 5 h post-irradiation while WT and VG crypt cells peaked at 2 h (Figure 4.5). The difference in DDR outcomes along the crypt-villus axis is again evident here, with mutation-susceptible proliferating

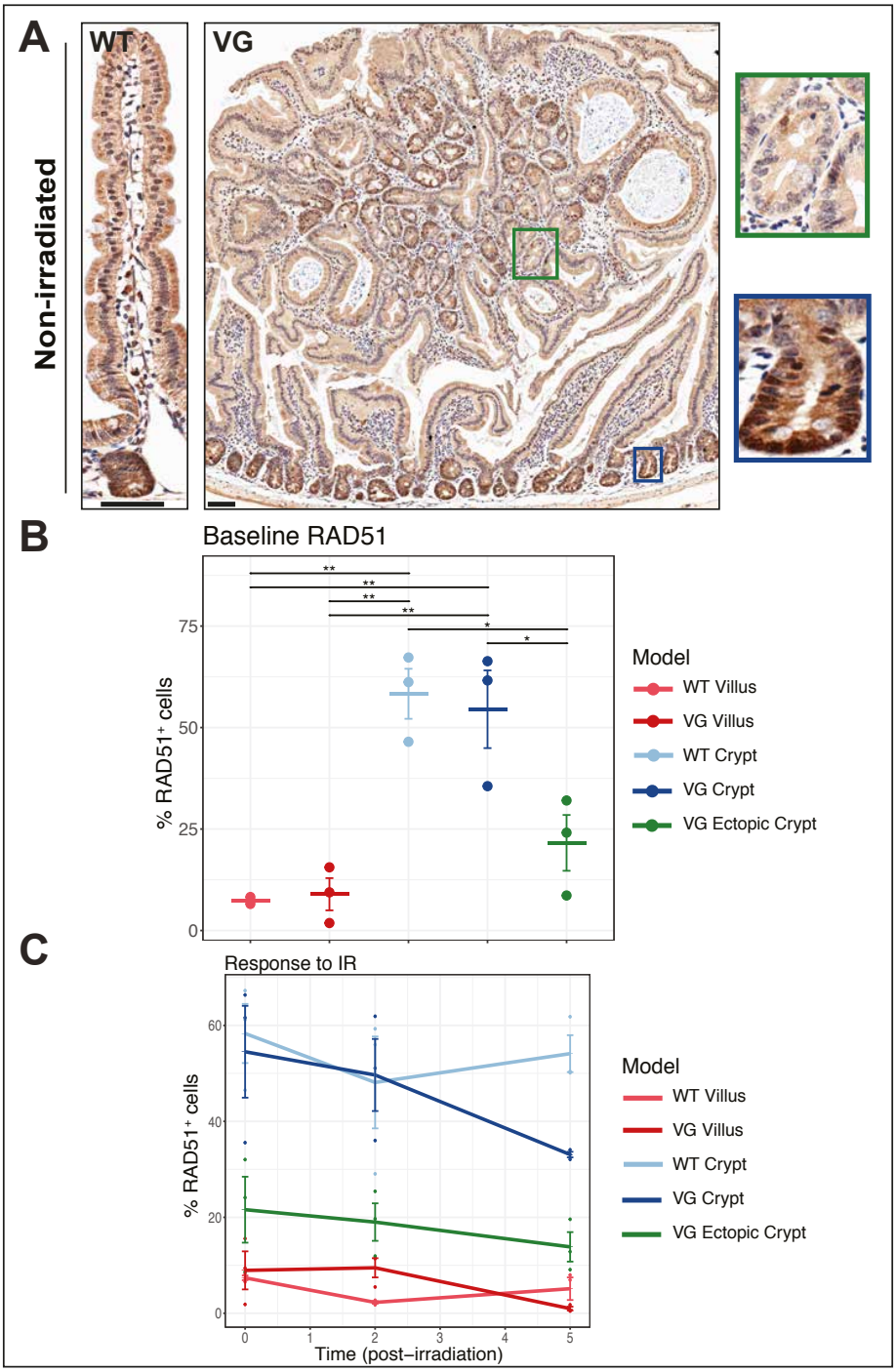


Figure 4.4: Analysis of RAD51 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of baseline RAD51 IHC staining in SB1 of the indicated experimental mouse. Scale bars represent 50 um, inlays of WT Villus (pink), WT Crypt (light blue), VG Crypt (dark blue) and VG Ectopic Crypt (green) compartments, respectively. B) Quantification of the percentage of $RAD51^+$ cells per indicated compartment at baseline. Error bars represent SEM, N = 3 mice, *$p < 0.05$, **$p < 0.01$ as calculated by one-way ANOVA with Tukey post-test. C) Quantification of the percentage of $Rad51^+$ cells per indicated compartment following irradiation. Error bars represent SEM between N = 3 biological replicates, each comprising 30 technical replicates.

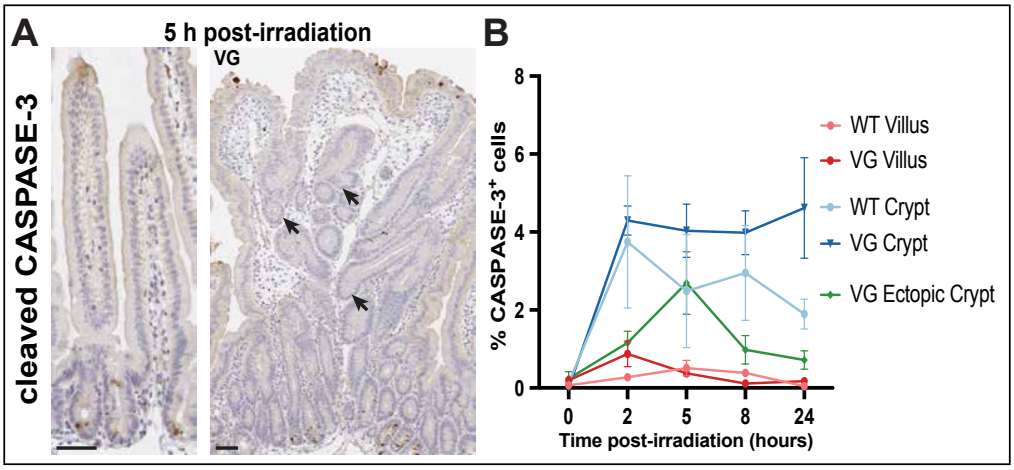


Figure 4.5: Analysis of cleaved caspase-3 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of cleaved caspase-3 IHC staining of SB1 of the indicated experimental mouse 5 h post-irradiation. Arrows indicate reduced expression within ectopic crypts, scale bars represent 50 μm. B) Quantification of the percentage of cleaved caspase-3^+ cells per indicated compartment at each timepoint post-irradiation. Error bars represent SEM of 5 technical replicates per compartment, N = 1 mouse per timepoint.

cells within WT and VG crypts activating apoptosis in response to 10 Gy IR while soon to be shed differentiated cells within villus compartments do not.

Interestingly, expression of the anti-apoptotic protein BAG1 was high in both normal WT and VG crypts and VG ectopic crypts at baseline (Figure 4.6;A), with no significant difference between ectopic crypt levels and normal crypt levels (Figure 4.6;B). However, there was no difference in BAG1 protein levels in response to irradiation between compartments, with all compartments exhibiting a slight increase in $BAG1^+$ cells at 2 h post irradiation that was ameliorated by 5 h.

Taken together, these results indicate that non-irradiated ectopic crypts harbour reduced levels of nuclear pATM and cellular RAD51 proteins compared to their proliferative counterparts at the crypt base; while both ectopic and basal crypts exhibit high levels of anti-apoptotic protection, mediated by BAG1. In response to irradiation, however, ectopic crypts exhibit reduced phosphorylation of the sensing protein H2AX, accompanied by reduced p-53 activation and a delayed, reduced activation of caspase-3 compared to both WT and VG crypts. Cells within basal crypts therefore harbour intact DDR in tandem with high levels of anti-

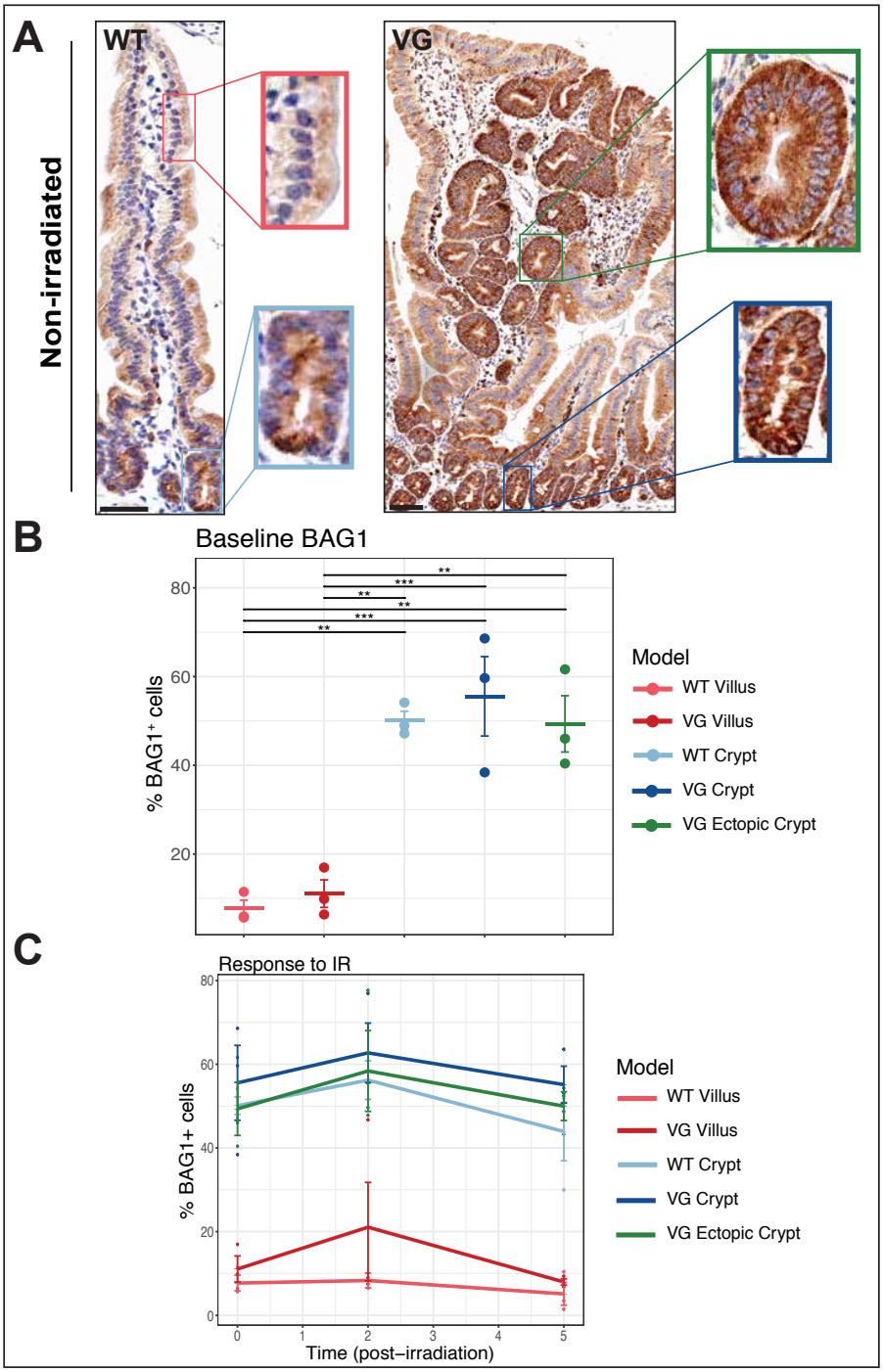


Figure 4.6: Analysis of BAG1 expression in intestinal tissue from WT and VG mice exposed to 10 Gy IR. A) Representative images of baseline BAG1 IHC staining in SB1 of the indicated experimental mouse. Scale bars represent 50 um, inlays of WT Villus (pink), WT Crypt (light blue), VG Crypt (dark blue) and VG Ectopic Crypt (green) compartments, respectively. B) Quantification of the percentage of $BAG1^+$ cells per indicated compartment at baseline. Error bars represent SEM, N = 3 mice, **p < 0.01, **)p < 0.001 as calculated by one-way ANOVA with Tukey post-test. C) Quantification of the percentage of $BAG1^+$ cells per indicated compartment following irradiation. Error bars represent SEM between N = 3 biological replicates, each comprising 30 technical replicates.

apoptotic proteins in order to facilitate and protect the high level of proliferation in this compartment. However, cells within ectopic crypts, despite containing the same level of anti-apoptotic protection, exhibit reduced levels of key DNA repair proteins and a detectable impairment in three key stages of DDR.

Intact DDR at the crypt base therefore results in the successful detection of DNA damage within proliferative cells that also harbour intact apoptotic pathways, whereas in ectopic crypts cells accrue DNA damage that fails to be accurately 'sensed' and translated into an appropriate apoptotic response, thereby allowing cells with potentially harmful mutations to replicate and mutate further. These results illustrate for the first time that proliferating cells within ectopic crypts of VG mice do not respond to DNA damage in the same way as their crypt-based counterparts, thus indicating that an intact DNA damage response is not a cell-intrinsic property of proliferative cells.

4.2.2 Increased miR-96 expression in the SB1 villus region of VG mice

MiRNAs are small (21-25mer), endogenous RNA molecules that negatively regulate gene expression at a post-transcriptional level. Furthermore, the expression of some miRNAs has been shown to change in response to DNA damage [156]. miR-96 is a relatively well characterised miRNA that has been shown to suppress translation of *Rad51* [110] and is part of the polycistronic miR-183-96-182 miRNA cluster that is rapidly downregulated in response to IR [156]. Further, miR-96 has been shown to be downregulated by BMP4 signaling in vascular endothelial cells [111]. Given that Grem1 shows high affinity for BMP4, over that of BMP7 [30], I hypothesised that aberrant epithelial Grem1 binds BMP4 and thus prevents BMP-mediated repression of miR-96 levels, facilitating miR-96 driven repression of RAD51 translation within the ectopic crypt compartment of VG mice. I therefore carried out qPCR analysis of miR-96 expression within the SB1 villus compartment of VG mice. miR-96 expression was elevated within the VG SB1 villus compartment compared to both WT and VG crypts and

WT villi (Figure 4.7; A). Furthermore, comparison of compartmental miR-96 expression between intestinal sections confirmed that elevated miR-96 expression was confined to SB1 VG vill (Figure 4.7; B), the region with the most profound overexpression of Grem1 [157].

Conversely, there was no sigificant difference in miR-96 expression in response to IR between regions, with the SB1 region of all compartments exhibiting similar miR-96 expression levels at 2 h and 5 h post-irradiation as that observed within the combined SB2-4 regions (Figure 4.8). The relatively low level of miR-96 expression within both WT and VG crypts remained unchanged following IR, while, in line with the literature, there was a downregulation of miR-96 expression detected within WT and VG villus compartments 2 h post-irradiation. Taken together, these results indicate that ectopic crypts harbour sufficiently increased expression of miR-96 to drive a detectable increase in miR-96 expression within VG SB1 villus compared to WT SB1 villus. This baseline increase in miR-96 expression does not, however, alter miR-96 expression in response to IR, with both WT and VG villus compartments exhibiting decreased miR-96 expression 2h post-irradiation.

Earlier results showed that RAD51 protein expression is reduced within non-irradiated ectopic crypts relative to WT and VG basal crypts. However, RAD51 expression in all compartments is largely unchanged in response to radiation. Given that *Rad51* is a well-documented target of the miR-183-96-182 cluster, it could thus be that aberrant epithelial GREM1 prevents BMP-mediated repression of miR-96 in non-irradiated crypts, with the subsequent increase in miR-96 levels facilitating increased repression of RAD51 translation. This is therefore detectable as decreased RAD51 protein expression within ectopic crypts.

4.2.3 Summary of findings

In view of my original hypothesis that increased mutation susceptibility in *Vil1-Grem1* mice is caused by Grem1-driven impairment of DDR within ectopic crypts, a summary of the key findings presented within this chapter are below.

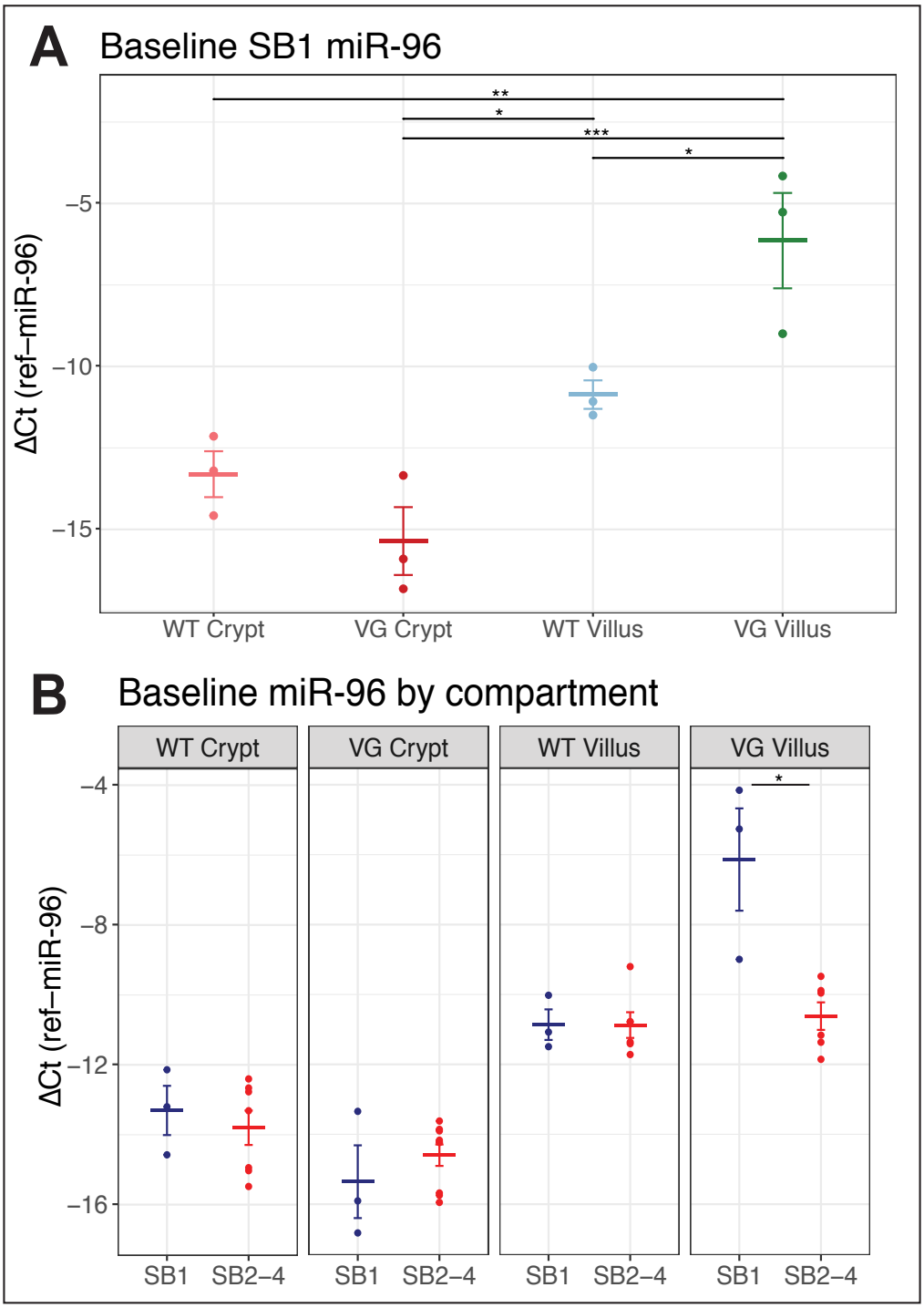


Figure 4.7: Expression analysis of homeostatic miR-96 within WT and VG mice. A) Baseline miR-96 expression within the SB1 intestinal region shown as normalised to the average expression of three reference genes U6, Sno202 and Sno234. B) Baseline miR-96 expression within the SB1 region compared to regions SB2-4 per intestinal compartment. Error bars represent SEM between N = 3 biological replicates, comprising 5 technical repeats. *p<0.05, **p<0.01, ***p<0.001 as calculated by one-way ANOVA with Tukey post-test.

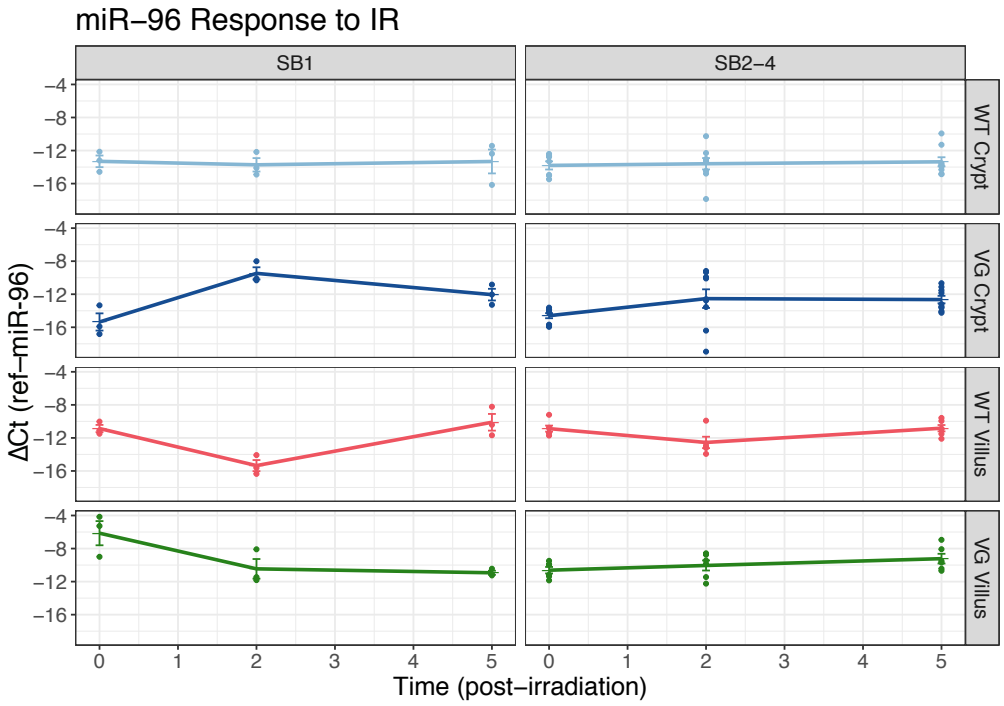


Figure 4.8: Expression analysis of miR-96 in response to 10 Gy IR within WT and VG mice. miR-96 expression shown as normalised to the average expression of three reference genes U6, Sno202 and Sno234 within the SB1 intestinal region (left panels) and SB2-4 regions (right panels) per intestinal compartment. Error bars represent SEM between N= 3 biological replicates, comprising 5 technical replicates.

- I have characterised the expression of key DDR pathway components following irradiation in *Vil1-Grem1* mice

- Cells at the crypt base protect their genomic integrity by mounting a rapid, robust response to the detection of DNA damage. Results presented here indicate that physiological cell proliferation is facilitated by high levels of anti-apoptotic protein. However, in the event of irrepareble DNA damage, these cells are capable of bypassing or overcoming this mechansim, and undergo appropriate apoptosis to prevent transmission of mutations to daghter cells.

- Proliferating cells within ectopic crypts also exhibit high levels of anti-apoptotic protein. However, these cells also harbour reduced expression of key DNA repair proteins at baseline and exhibit impaired activation of DDR pathways in response to irradiation. Proliferation within this compartment

thus occurs in the absence of protective mechanisms against DNA damage and with reduced capacity to mount an apoptotic response, rendering ectopic crypt cells more susceptible to genomic instability and providing a mechanistic basis for the results presented in Chapter 3.

– Within ectopic crypts, miRNAs may represent a post-transcriptional link between dysregulated BMP signalling and impaired DDR. I have therefore suggested a novel mechanism by which a Grem1-miR-96-Rad51 axis impacts DDR within the ectopic crypt compartment of VG mice

4.3 Discussion and Future work

4.3.1 Discussion

There is a fine and evolved balance to be played between controlled cell proliferation in a stem cell-supported system, with stereotypical architecture and stringent cell fate determination, and the need to prevent mutations passing into daughter cells. Previous work investigating DNA repair within the intestine has demonstrated that the populations of cells along the crypt-villus axis employ distinct DDR pathways. Furthermore, the efficiency and speed of these pathways is thought to correlate with the 'dispensability' of cells: stem cells at the crypt base preferentially employ stringent HR repair and mount a rapid, robust apoptotic response to irreparable damage; TA cells employ slightly less efficient NHEJ repair and a delayed apoptotic response, and soon to be shed villus cells fail to exhibit activation of DDR. However, it remained unclear whether pathway choice is a therefore cell-intrinsic property linked to cell position and fate determination or whether cell-extrinsic, microenvironmental factors influence DDR regulation along the homeostatic crypt-villus axis. Work presented here suggests that the evolved fine line between proliferation and mutation protection is not a cell-intrinsic property of a proliferating cell *per se*, but that microenvironmental signalling can skew this balance to favour the persistence of mutated daughter cells.

4. Cell intrinsic mutation susceptibility in Vil1-Grem1 mice

Results presented here have shown that cells proliferating within ectopic crypts exhibit an altered response to DNA damage compared to their counterparts at the crypt base, thus indicating that cell-extrinsic microenvironmental dysregulation can influence DDR regulation. These findings provide evidence, through the use of a representative population of cells proliferating outside of the stereotypically controlled zones of the crypt-villus axis, that mutation protection mechanisms are not simply a cell intrinsic property of proliferating cells, which would have been evidenced by proliferative cells within basal and ectopic crypts exhibiting equivalent DDR activation in response to IR. Instead, deranged morphogen gradients within the *Vil1-Grem1* microenvironment appear to not only result in the uncoupling of cell fate determination from position along the crypt villus axis (resulting in the formation of ectopic crypts) but also a shift in the mechanisms that protect against somatic mutation acquisition within these crypts. It appears then that homeostatic regulation of DDR along the crypt villus axis is also influenced by signalling pathway gradients.

This raises the intriguing prospect of dual roles for intestinal signalling pathways in not only influencing cell fate along the crypt-villus axis, as has been widely documented, but also impacting upon the protective mechanisms employed by cells against mutation accumulation. There is precedent for this, with both TGF-β and IL-22 signalling implicated in the maintenance of genomic stability and the regulation of DDR pathway activation, respectively. Head and neck squamous cell carcinoma (HNSCC) cells negative for TGF-β activity exhibit compromised HR-directed DNA repair and an increase in genomic instability as a result of increased reliance upon AEJ repair. Furthermore, selective removal of an IL-22 receptor from intestinal epithelial cells resulted in suppression of DNA damage-induced apoptosis and increased tumour formation in an inflammation-driven tumour model. The findings presented within this chapter thus specifically implicate BMP signalling in dysregulated DDR activation and raise numerous questions regarding the mechanisms of this regulation and the clinical impact of these findings.

4.3.1.1 Out with the old axis, in with the new

Within this context, I have proposed a novel Grem1-miR-96-RAD51 axis, whereby aberrant epithelial Grem1 within the SB1 region of Vil1-Grem1 mice results in increased miR-96-driven repression of *Rad51* in non-irradiated ectopic crypts. Given that a similar, TGF-β-driven increase in miR-182-mediated repression of HR-associated *BRCA1* results in an increase in genomic instability and AEJ repair [104], it may be that a similar skewing of the BMP-miR-96-RAD51 axis in *Vil1-Grem1* mice results in increased genomic instability within ectopic crypts. This could represent a mechanistic basis for the increased mutation susceptibility of VG mice demonstrated in Chapter 3. Furthermore, the presence of a BMP-miR-96-RAD51 axis would have translational relevance for rectal cancer, which comprises ~29% of all CRC. Rectal cancer is unique in that the tumours are fixed in the pelvis and are thus amenable to radiation therapy. However ~15% of rectal cancers are completely resistant to radiotherapy, meaning that pateints are exposed to this damaging therapy for no discernible clinical benefit. The biological basis for this resistance to radiation response has been very difficult to discern and predict in advance of treatment decisions. Sporadic rectal cancers arise from either conventional adenomas or traditional serrated adenomas (TSAs). The latter of these are rare but aggressive lesions that have a particular predilection for the distal colon. The findings presented in this chapter could thus provide a basis for aberrant epithelial Grem1 in a subset of serrated rectal lesions contributing to impaired DDR activation in response to radiotherapy. Future work to confirm the link between BMP signalling and miR-96 expression, and explore this in serrated lesions, is proposed at the end of this section.

Work presented within this chapter thus raises the intriguing possibility of BMP signalling activity not only influencing intestinal cell fate determination but also how cells respond to DNA damage. This raises fundamental questions regarding the factors that determine whether or not a cell is dispensable (and thus the DDR pathways it employs). How cells 'know' this is a fascinating puzzle: is it a cell intrinsic property determined by cell type or position within

the crypt-villus axis or is it the culmination of what a cell is 'told' by the signals it is exposed to? Moreover, how set are these processes, given that we know cells 'told' to differentiate into mature lineages are able to dedifferentiate and fulfil a different function within the crypt given the appropriate intestinal signals (Section 1.4.4.2). Given that these findings implicate aberrant epithelial Grem1 expression in disrupted mutation protection, these notions have implications for *Vil1-Grem1* and HMPS lesion initiation. In these models, polyposis is driven by mutation acquisition within an expanded population of proliferative cells. Work presented here thus indicates that cells within these populations also harbour microenvironmentally disrupted mutation protection pathways. Importantly, this represents a therapeutic opportunity, such that neutralisation of aberrant epithelial Grem1 and subsequent restoration of cell fate determination may also restore homeostatic mutation protection. This premise is explored in the next chapter.

4.3.1.2 Limitations

The results discussed are, however, limited by some caveats of the analytical approach utilised. The proposed link between Grem1 and miR-96 requires functional validation *in vitro* or *in vivo*. However, there are numerous difficulties recapitulating the complex and cross-talking intestinal cell populations in vitro, given a high degree of functional redundancy due to the importance of protecting the genomic integrity of proliferating cells. Mir-ISH is a potential approach, although initial experiments failed and qPCR indicates these miRNAs are present at very low levels and so require highly sensitive detection. We could also isolate RNA from ectopic crypts, but this has limited success and the utility of gene expression to interrogate complex DDR networks is limited.

4.3.1.3 Future Work

Work in this field is thus moving towards the investigation of microenvironmental influence on mutation protection as a mechanism for increased genomic stability,

especially in cancers, which have been shown to reside within specifically perturbed tumour microenvironments. The gold standard is therefore recapitulating these complex systems in vivo in order to accurately characterise the relationship between signalling pathways and DNA repair mechanisms. However, given the evolutionary requirement for mutation protection mechanisms, these are complex, cross-talking pathway with multiple redundancies built in. Recapitulating these pathways in vitro has thus proved challenging. A promising emerging method involves the use of injected xenografts of specific cell populations into the colons of mice that are then irradiated, with the subsequent response of injected cells monitored. This provides a means of assessing the response of specific cell populations to an intact intestinal microenvironment.

There is mounting evidence that miR-96 is associated with oncogenesis and increased radiosensitivity [158], with several studies specifically showing that miR-96 is upregulated in colorectal cancer [159–161]. Given the work presented here focuses on the earliest stages of tumourigenesis, where ectopic crypts and subsequent polyp formation are driven by microenvironmental disruption, this may explain the relatively small but significant increase in miR-96 expression, localised to the compartment with most profound microenvironmental disruption, presented in this chapter. Could the degree of miR-96 de-repression thus be a marker of tumour progression and likelihood of radioresistance? There is already evidence for this in HNSCC [162] and glioblastoma [163]. Further work exploring this exciting possibility could therefore examine miR-96 expression in radioresistant versus radiosensitive rectal cancers alongside Grem1 expression, since a subset of rectal cancers originate from serrated lesions associated with high epithelial Grem1 [164]. Combining Grem1 expression with miR-96 derepression could be a useful clinical marker for serrated lesions with impaired DDR, thus enabling identification of radioresistant lesions. Silencing miR-96 could also represent a potential therapeutic avenue to sensitise these lesions to radiotherapy.

"Everything not saved will be lost." - Nintendo

5

The impact of therapy on mutation acquisition in *Vil1-Grem1* mice

Contents

5.1 Introduction 134
5.2 Results: Anti-Grem1 in *Vil1-Grem1* mice 136
5.2.1 Survival analysis 136
5.2.2 Phenotypic characterisation 137
5.2.3 Genomic analysis 144
5.2.3.1 Mutation burden analysis 144
5.2.3.2 Characterisation of mutation types 145
5.2.3.3 Mutational signature analysis 147
5.2.3.4 Characterisation of driver gene mutations . . . 152
5.2.4 Summary: Anti-Grem1 therapy in *Vil1-Grem1* mice . . . 159
5.3 Discussion 162

5.1 Introduction

Results in Chapter 3 have shown that differences in disease state correspond to altered mutation susceptibility and acquisition. Chapter 4 went on to investigate increased mutation susceptibility within *Vil1-Grem1* mice and proposed a novel mechanism by which aberrant Grem1 within the epithelial microenvironment drives increased genomic instability within ectopic crypts. Here, I sought to further explore somatic mutation in disease states by investigating the impact of therapy, and restoration of cell fate determination, on mutation persistence. As described previously, the widespread intestinal polyposis observed within *Vil1-Grem1* mice is driven by aberrant epithelial expression of *Grem1*. Our lab, in collaboration with UCB Pharmaceuticals, therefore sought to use an antibody against GREM1 in order to explore the effect of neutralising excess GREM1 to see whether this would return cell fate determination to normal. This antibody, Ab7326 (anti-Grem1 from here), is an allosteric inhibitor, with a binding site distal from the known BMP binding regions on GREM1. Anti-Grem1 is therefore able to access the allosteric binding site even when BMP is complexed to GREM1, resulting in significant inhibition of GREM1 activity (UCB Pharmaceuticals, unpublished data). Our lab has shown that twice weekly, subcutaneous administration of anti-Grem1 antibody (30 mg/kg) for a period of 6 weeks results in a significant, dose-dependent abrogation of the pan-intestinal polyposis phenotype in these mice (Figure 5.1, leftmost panel). Furthermore, immunohistochemical staining of treated animals showed reversal of the deranged cell fate determination and resultant ectopic crypt formation that characterised untreated tissue (Figure 5.1). In the context of this thesis, I was therefore eager to explore the impact of anti-Grem1 therapy, and subsequent abrogation of the proliferative phenotype, on mutation susceptibility and acquisition within *Vil1-Grem1* mice. I hypothesised that the damaging, pro-carcinogenic mutations brought on by AOM treatment in *Vil1-Grem1* mice are either 'arrested' by therapy and so remain, in a 'dormant'

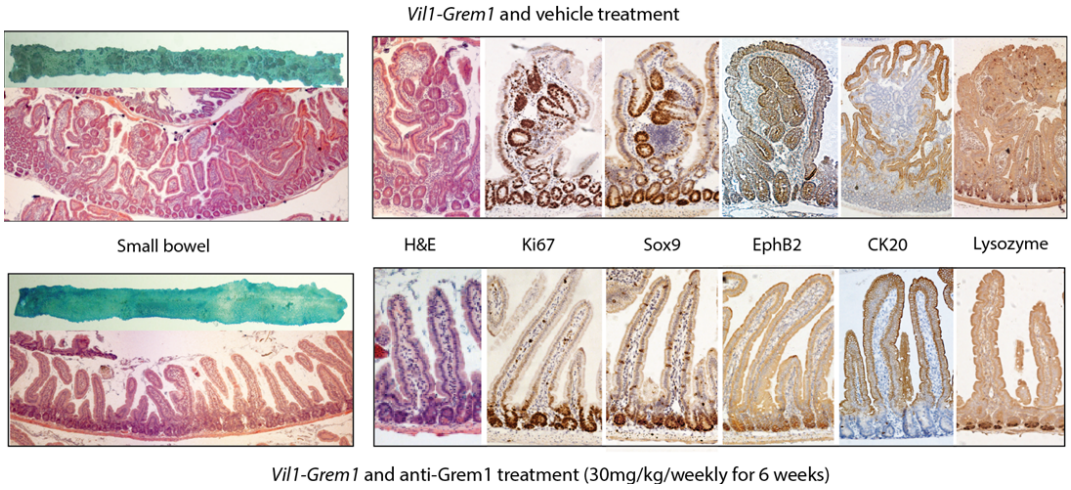


Figure 5.1: Abrogation of *Vil1-Grem1* intestinal phenotype following anti-Grem1 antibody treatment

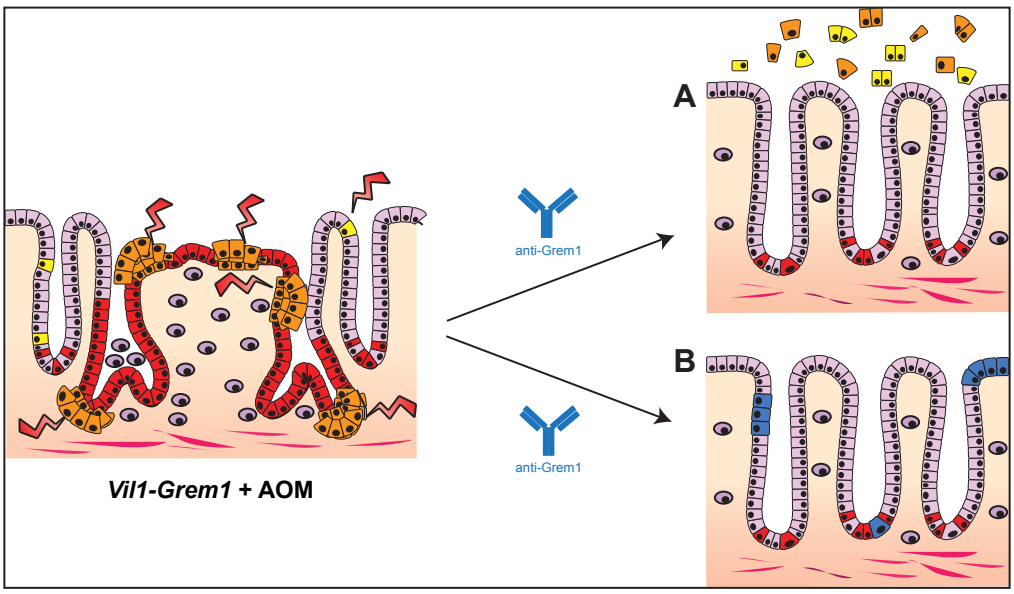


Figure 5.2: A model of therapy-driven protection against mutagenesis. Red represents proliferative cells, yellow and orange represent those harbouring damaging mutations, blue represent those harbouring 'dormant' mutations following therapy.

or inactive state, within treated tissue or instead 'shed' through therapy-driven remodelling of dysplastic epithelium (Figure 5.2).

In order to explore this hypothesis, I took an *in vivo* approach that utilized *Vil1-Grem1* mice treated with an anti-Grem1 antibody prior to, during, and following treatment with the DNA-damaging agent, AOM, described previously in this thesis. Specific aims were:

- To initiate cell fate restoration prior to mutagen exposure in order to assess whether anti-Grem1 therapy conferred protection against AOM-driven mutagenesis and any subsequent associated neoplasia
- To characterise the intestinal phenotype arising from *Vil1-Grem1* mice treated with anti-Grem1
- To explore the impact of restoration of cell fate determination on the genomic profile of treated and untreated colonic tumours

5.2 Results: Anti-Grem1 in *Vil1-Grem1* mice

In order to explore the effect of anti-Grem1 therapy, *Vil1-Grem1* mice were first aged 3 months to ensure the development of ectopic crypts. Mice were then treated with nine weeks of 30 mg/kg anti-Grem1 (UCB Pharmaceuticals) or isotype control (mIgG1) antibody (Figure 5.3). Following the first three weeks of antibody treatment, mice were treated with 10 mg/kg AOM once weekly for three weeks, alongside ongoing antibody therapy. Following three weeks of combined AOM and antibody therapy, mice underwent a further three weeks of antibody-only treatment, for a total treatment period of nine weeks. After the 9-week treatment period, mice were observed daily and sacrificed upon exhibiting symptoms of intestinal illness, such as anaemia and hunching. The intestinal tract was collected prior to sampling of colonic polyps and adjacent normal tissue for DNA extraction (Section 2.2.6.2). The small intestine, divided into three regions (SB1, 2 and 3), and colon were then swiss rolled and processed for immunohistochemical analysis of intestinal epithelial markers (Section 2.2.6).

5.2.1 Survival analysis of *Vil1-Grem1* mice treated with AOM and anti-Grem1

I first sought to characterise any difference in survival between experimental groups. Survival analysis, shown in Figure 5.4, confirmed a significant reduction in survival of *Vil1-Grem1* mice exposed to AOM, shown previously in Chapter

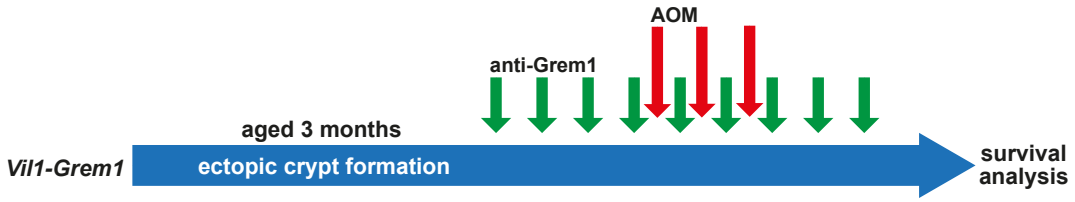
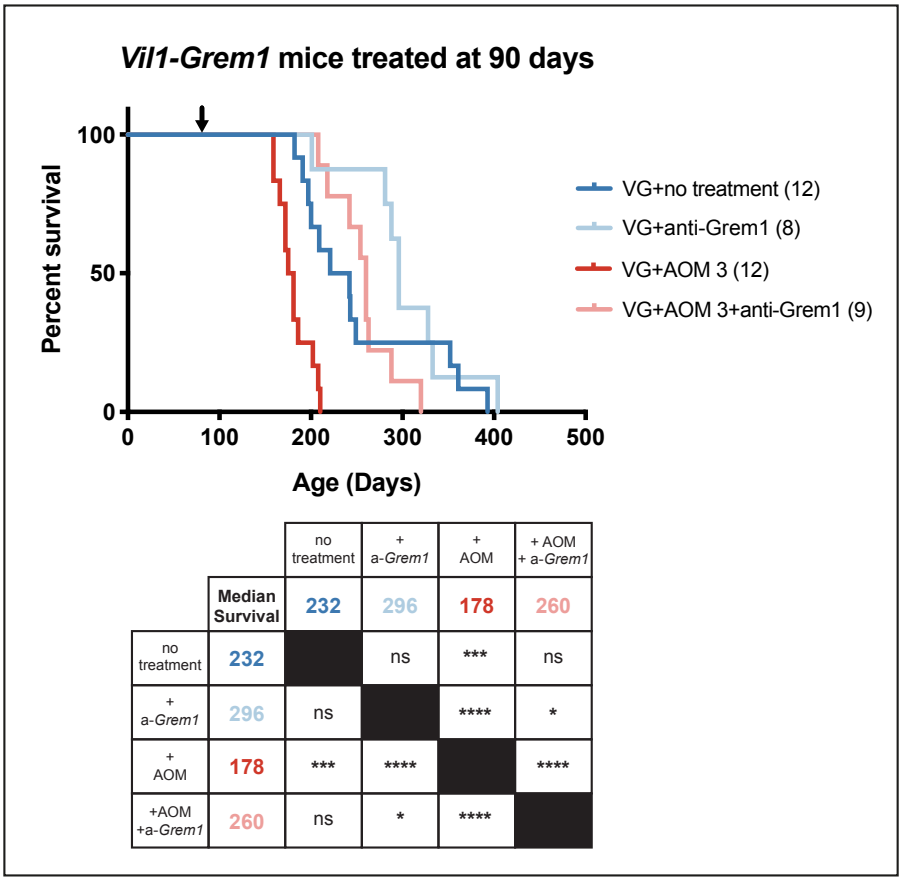


Figure 5.3: Treatment scheme for *Vil1-Grem1* mice treated with an anti-Grem1 antibody and AOM.

3, as compared to all other experimental groups. The data also demonstrated a significant increase in median survival exhibited by animals exposed to AOM and treated with anti-Grem1 antibody (260 days) compared to those exposed to AOM and treated with an isotype-control (178 days). Furthermore, both experimental groups treated with anti-Grem1, VG+anti-Grem1 (296 days) and VG+AOM+anti-Grem1 (260 days), exhibited higher median survival than that of untreated *Vil1-Grem1* animals (232 days), although these differences in survival curves did not reach statistical significance. Taken together, these results demonstrate that anti-Grem1 therapy can rescue the significant reduction in median survival of *Vil1-Grem1*+AOM mice to that of untreated *Vil1-Grem1* animals.

5.2.2 Phenotypic characterisation of anti-Grem1-treated *Vil1-Grem1* small intestine and colon

Upon establishing that treatment with anti-Grem1 rescues median survival of *Vil1-Grem1* mice exposed to AOM, I next sought to characterise any differences in intestinal phenotype between experimental groups. As shown in Figure 5.5, there was no significant difference in small intestine or colon length between all experimental conditions. There was, however, a significant increase in colon crypt length between untreated *Vil1-Grem1* mice and those exposed to AOM, as initially observed in Chapter 3. Although the average crypt length of animals treated with anti-Grem1 was less than those exposed to AOM, this difference did not reach statistical difference. This result indicates that exposure to AOM results in a thickening of *Vil1-Grem1* colonic epithelium that is partially resolved by anti-Grem1 therapy.



		no treatment	+ a-Grem1	+ AOM	+ AOM + a-Grem1
	Median Survival	232	296	178	260
no treatment	232		ns	***	ns
+ a-Grem1	296	ns		****	*
+ AOM	178	***	****		****
+AOM +a-Grem1	260	ns	*	****	

Figure 5.4: Kaplan-Meier survival plot (top) and analysis (bottom) for *Vil1-Grem1* mice treated with the mutagen AOM and an anti-Grem1 antibody. Arrow indicates start of treatment, n as indicated. *p < 0.05, ***p < 0.001, ****p < 0.0001 as calculated by log-rank Mantel-Cox test.

Despite observing no significant difference in macroscopic intestinal characteristics, H&E staining of small intestine and colon swiss rolls indicated variability in the number of polyps per mouse. I thus utilized Ki67 staining of proliferative cells to quantify the total number of small intestine polyps per mouse within each experimental group (Figure 5.6; A). Both experimental groups that received anti-Grem1 therapy exhibited significant reductions in small intestine polyp burden compared to both untreated *Vil1-Grem1* mice and those exposed to AOM (Figure 5.6; B). Further characterisation of these polyps within each experimental group revealed no significant difference in the number of proliferative, apoptotic or progenitor cells per polyp, as calculated by the percentage of Ki67, cleaved caspase-3 and Sox9 positive cells per polyp per mouse (Figure 5.6; C). Taken

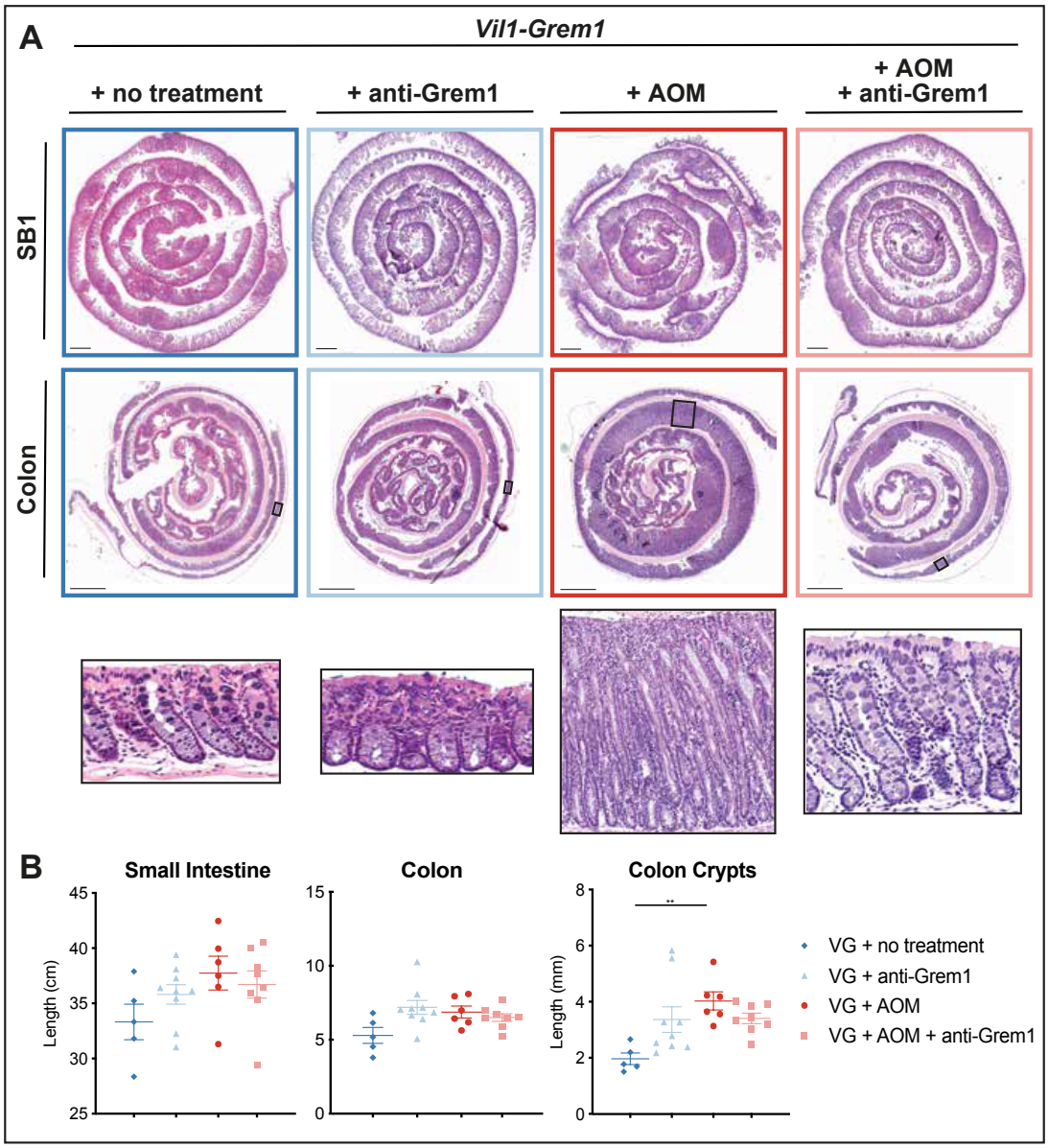


Figure 5.5: Macroscopic intestinal phenotype of *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. A) Representative images of hematoxylin and eosin staining of small intestine section 1 (SB1) and colon of experimental mice. Scale bars represent 2 mm. B) Quantification of small intestine, colon and colonic crypt length of experimental mice. Points represent individual mice or (for colon crypts) the average of 30 technical replicates per mouse, error bars represent SEM. N as indicated. *p < 0.05 as calculated by one-way ANOVA with Tukey post-test.

together, these findings indicate that anti-Grem1 therapy rescues polyp burden within the small intestine of *Vil1-Grem1* and *Vil1-Grem1*+AOM mice, although polyps that escape rescue do not exhibit any significant difference in intestinal phenotypic markers. Small bowel phenotype is used here to indicate the efficacy of antibody treatment in *Vil1-Grem1* mice. The remainder of this analysis will focus on the colon, given that AOM is a colonotropic mutagen.

VG+AOM mice exhibited an increase in colon crypt length and an expansion of the proliferative Ki67+ cell compartment per crypt, as previously shown in Chapter 3. Treatment with anti-Grem1, however, produced a statistically significant reduction in average colon crypt length(Figure 5.7). This is driven by a statistically significant reduction in the percentage of proliferative $Ki67^{+}$ cells per colon crypt, accompanied by an increase in the proportion of $Sox9^{+}$ progenitor cells, presumably as a consequence of the reduction in overall crypt size. This indicates that anti-Grem1 therapy partially rescues the AOM-driven increase in mutationally susceptible cells per colon crypt in VG mice. I next quantified the colon polyp histology within each experimental condition. As discussed previously in Chapter 3, untreated VG mice rarely produce colonic tumours at the age these mice were sacrificed, but those mice exposed to AOM exhibit significant colonic polyp burden. In keeping with this, the majority of untreated VG mice, and those treated only with anti-Grem1, exhibited no colon polyps (Figure 5.7). VG+AOM mice, however, exhibit a significant increase in colonic polyp burden that is partially abrogated by treatment in VG+AOM+anti-Grem1 mice. Further characterisation of the size of colon polyps within each experimental group revealed that VG+AOM mice harboured polyps 2 mm (75%) or larger (25%), while those treated with anti-Grem1 exhibited a greater variety of polyp sizes, with 25% 1mm or less, 25% 2 mm and the remainder 3 mm or larger. Anti-Grem1 therapy in VG+AOM mice thus results in a reduced number of colonic polyps that exhibit greater variety in size. Interestingly, the effect of anti-Grem1 on colon crypts is also evident within colon polyps, with analysis of the percentage of proliferative and progenitor cells in VG+AOM+anti-Grem1 polyps showing a reduction in

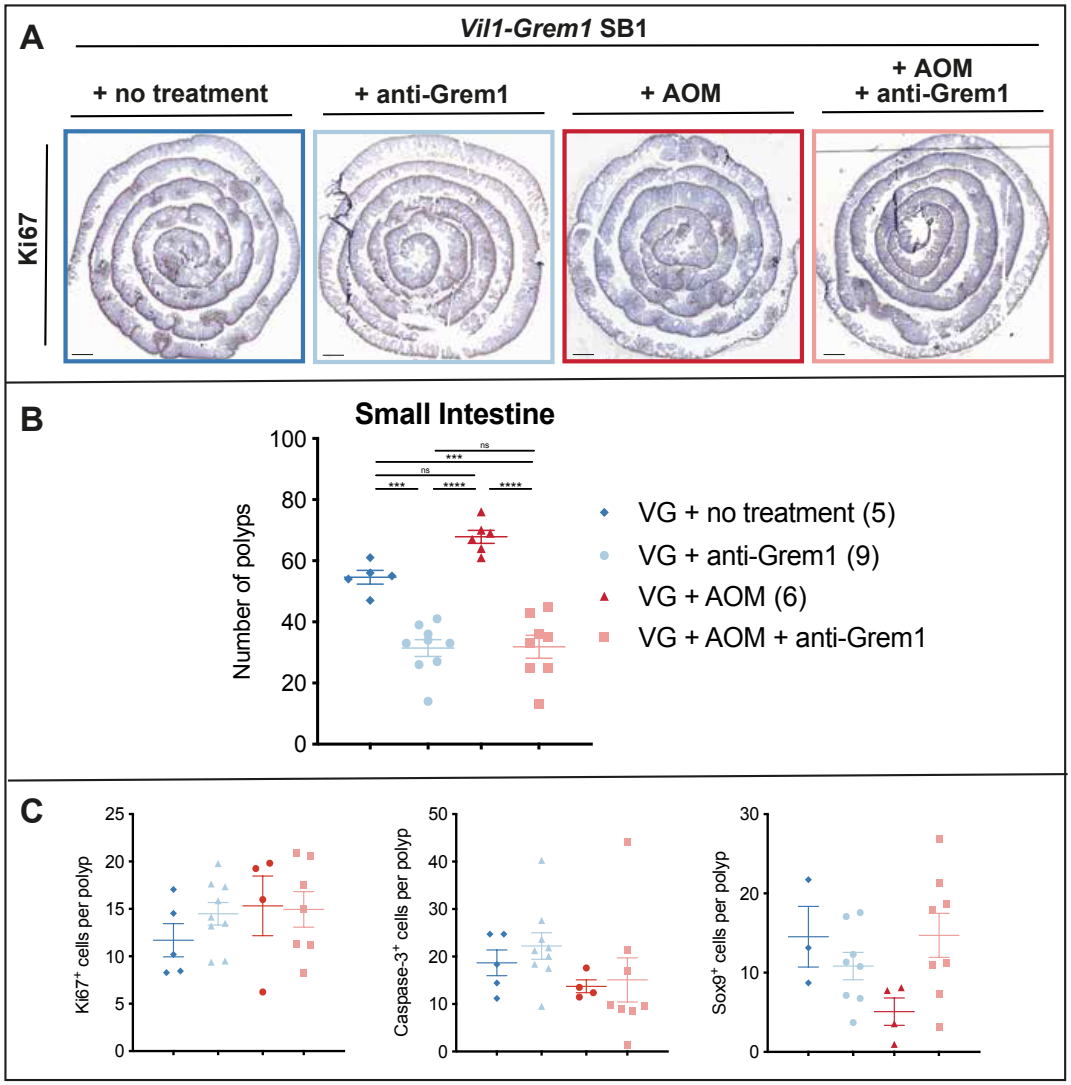


Figure 5.6: Characterisation of small intestine phenotype of *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. A) Representative images of Ki67 IHC staining of small intestine section 1 (SB1) of experimental mice. Scale bars represent 2 mm. B) Quantification of the total number of small intestine polyps per experimental condition (points represent individual mice) C) Quantification of the number of Ki67, cleaved caspase-3 and Sox9 positive cells per small intestinal polyp per experimental condition (points represent individual polyps). All error bars represent SEM. *p < 0.05, ***p < 0.001, ****p < 0.0001 as calculated by one-way ANOVA with Tukey post-test.

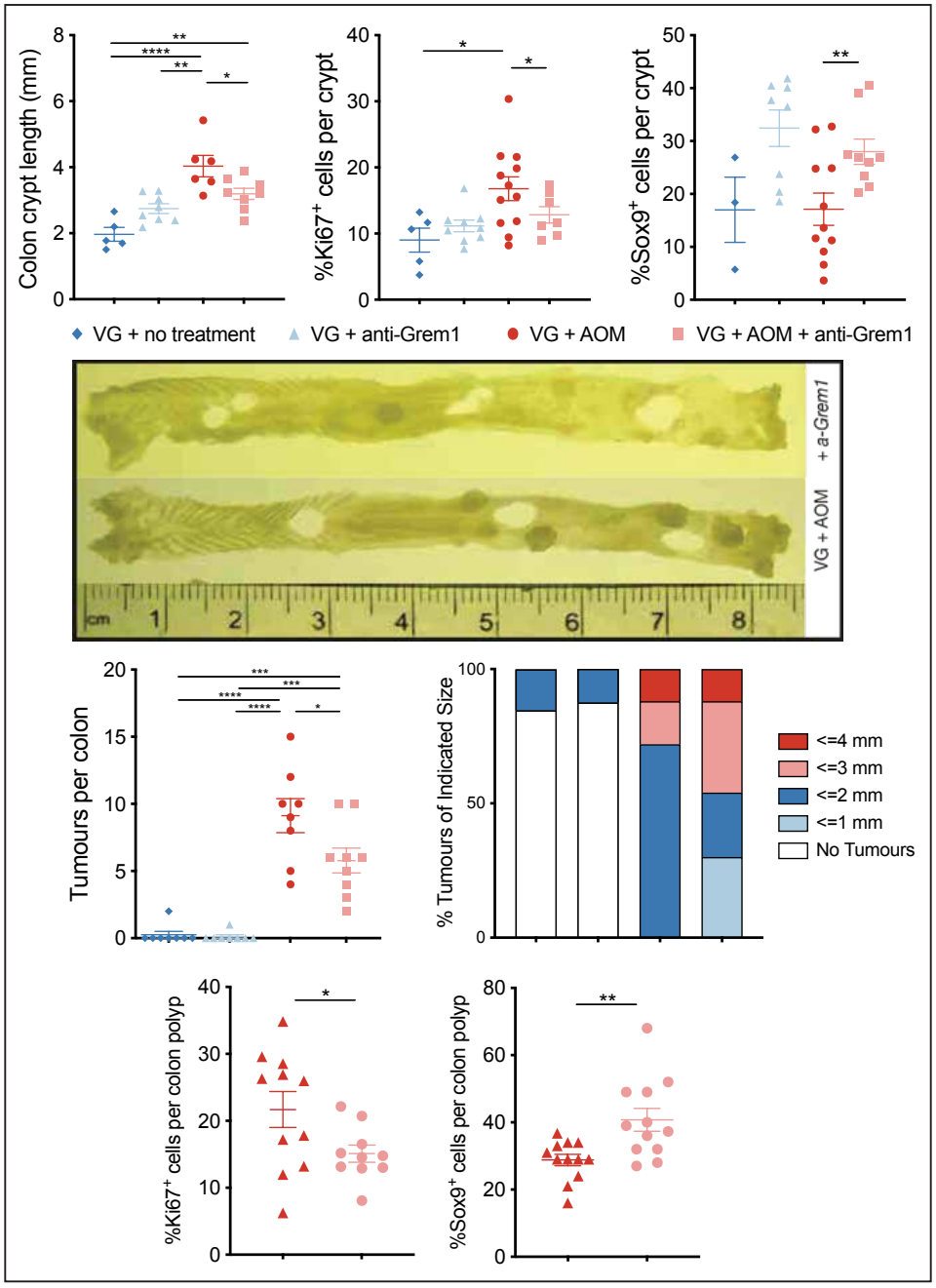


Figure 5.7: Characterisation of colon phenotype in *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Top: Quantification of colon crypt length and percentage of proliferative Ki67$^+$ and progenitor Sox9$^+$ cells (points represent average per mouse). Middle: Representative images of colons from VG AOM and VG+AOM+anti-Grem1 mice and quantification of the number and size (calculated as percentage of all colon polyps recorded) of colonic polyps per mouse per experimental condition. Bottom: Quantification of the percentage of proliferative Ki67$^+$ and progenitor Sox9$^+$ cells per colon polyp (points represent poylps). All error bars represent SEM. *p < 0.05, **p < 0.001, ***p < 0.001, ****p < 0.0001 as calculated by one-way ANOVA with Tukey post-test.

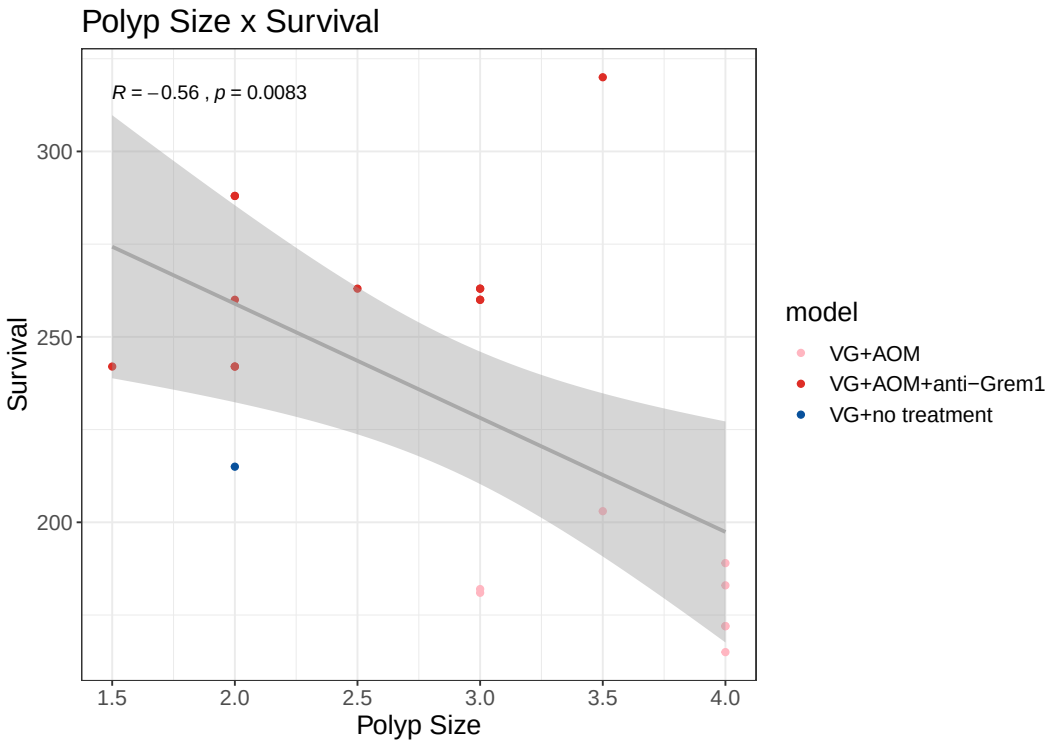


Figure 5.8: Correlation analysis of colon polyp size and mouse survival. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

Ki67$^+$ cells per polyp, accompanied by an increase in the proportion of Sox9$^+$ cells, compared to untreated polyps (Figure 5.7). Taken together, these results thus indicate that anti-Grem1 therapy partially rescues the AOM-driven increase in crypt length and polyp burden in VG mice through a reduction in the number of mutationally susceptible proliferative cells within the colon.

I also found a significant negative correlation (R = -0.56, p = 0.0083) between the size of colon polyps sampled for further analysis and mouse survival (Figure 5.8). The clustering of samples according to experimental condition further demonstrates that VG+AOM+anti-Grem1 mice exhibit increased survival and smaller polyps compared to untreated VG+AOM mice. This is especially pertinent given that the increased survival of VG+AOM+anti-Grem1 mice meant treated polyps were sampled from older mice on average, yet these remained smaller than VG+AOM polyps sampled at earlier timepoints, due to their reduced survival. These findings therefore further confirm that anti-Grem1 therapy increases survival and reduces polyp size and burden in treated *Vil1-Grem1* mice.

5.2.3 Genomic analysis

Having shown that anti-Grem1 therapy drives an increase in median survival and a reduction in polyp burden in VG+AOM mice, I next sought to investigate whether this difference in survival and intestinal phenotype was caused by an anti-Grem1-driven shift in mutation acquisition. I therefore extracted DNA (Section 2.4.1.1) from 12 colon polyps sampled from *Vil1-Grem1*+AOM+anti-Grem1 mice for genomic analysis. Sample quality control and subsequent Whole Exome Sequencing (WES) was carried out by the High-Throughput Genomics Group at the Wellcome Centre for Human Genetics, University of Oxford (Section 2.4.3.2). Pre-processing, variant calling and filtering of provided sequencing files was carried out by Ryan Schenck, a bioinformatician within the Leedham lab (Section 2.4.3.2-6). I carried out all subsequent variant analysis using R and various publicly available resources.

5.2.3.1 Analysis of anti-Grem1-treated *Vil1-Grem1* polyp mutation burden

In order to determine whether anti-Grem1 protected against AOM-induced mutagenesis through putative normalisation of epithelial cell fate, I assessed non-synonymous mutation burden in polyps from mice within each experimental group. To contextualise sequencing information from VG+AOM+anti-Grem1 polyps, I also included data from 7 VG+AOM polyps exposed to 1, 3 or 6 doses of AOM and 1 rare polyp from an untreated VG mouse, discussed previously in Chapter 3. There was significant variation in both the total (not shown) and non-synonymous mutation burden within all polyps profiled, ranging from 138 to 1635 non-synonymous mutations per polyp (Figure 5.9, A). The single polyp collected from the untreated VG mouse contained 30 non-synonymous mutations. Calculation of the average non-synonymous mutation burden per polyp within each experimental condition further displayed the wide variation in mutation burden observed within experimental conditions (Figure 5.9, B) and confirmed that although the mean number of non-synonymous mutations per anti-Grem1-treated polyp (666) was less than that of VG+AOM polyps exposed to

3 and 6 doses of AOM (916 and 862, respectively), this did not reach statistical significance (p = 0.41). This indicates that the increase in survival and decrease in polyp burden driven by our experimental anti-Grem1 treatment program either does not occur as a result of a reduction in the number of mutations per polyp. Alternatively, it may also be that polyps that persist following therapy have escaped a putative anti-Grem1 mutation abrogation effect.

To further explore the relationship between the observed phenotype and mutation burden in these animals, I investigated whether there was any correlation between mutation burden and phenotypic markers such as polyp size and mouse survival. Correlation analysis in Figure 5.10 revealed that although treated and untreated *Vil1-Grem1* polyps cluster according to survival, the y-axis variability in mutation burden per polyp results in only a weak negative correlation between mutation burden and survival (R = -0.21) that did not reach statistical significance (p = 0.37). Similarly, Figure 5.11 shows that although, as expected, polyps treated with anti-Grem1 populate the lower range of polyp size, there is only a weak positive correlation between polyp size and mutation burden (R = 0.31) that did not reach statistical significance. This further suggests that the phenotypic outcomes observed in *Vil1-Grem1* mice treated with anti-Grem1 are not correlated with or driven by changes in mutation burden.

5.2.3.2 Characterisation of mutation types within anti-Grem1-treated *Vil1-Grem1* polyps

I next sought to investigate whether the improvement in phenotypic outcomes in VG+AOM+anti-Grem1 mice was driven by a shift in the types of mutation present within treated versus untreated polyps. I therefore characterised the mutations within each polyp according to impact (Figure 5.12, A). VG+AOM polyps exhibited decreased average contributions of UTR, non-genic, intron and indel variants and increases in the average contribution of synonymous and missense mutations compared to the untreated *Vil1-Grem1* polyp (Figure 5.12, B). However, given that only a single untreated polyp was available for analysis, these findings require further validation to achieve significance. Analysis

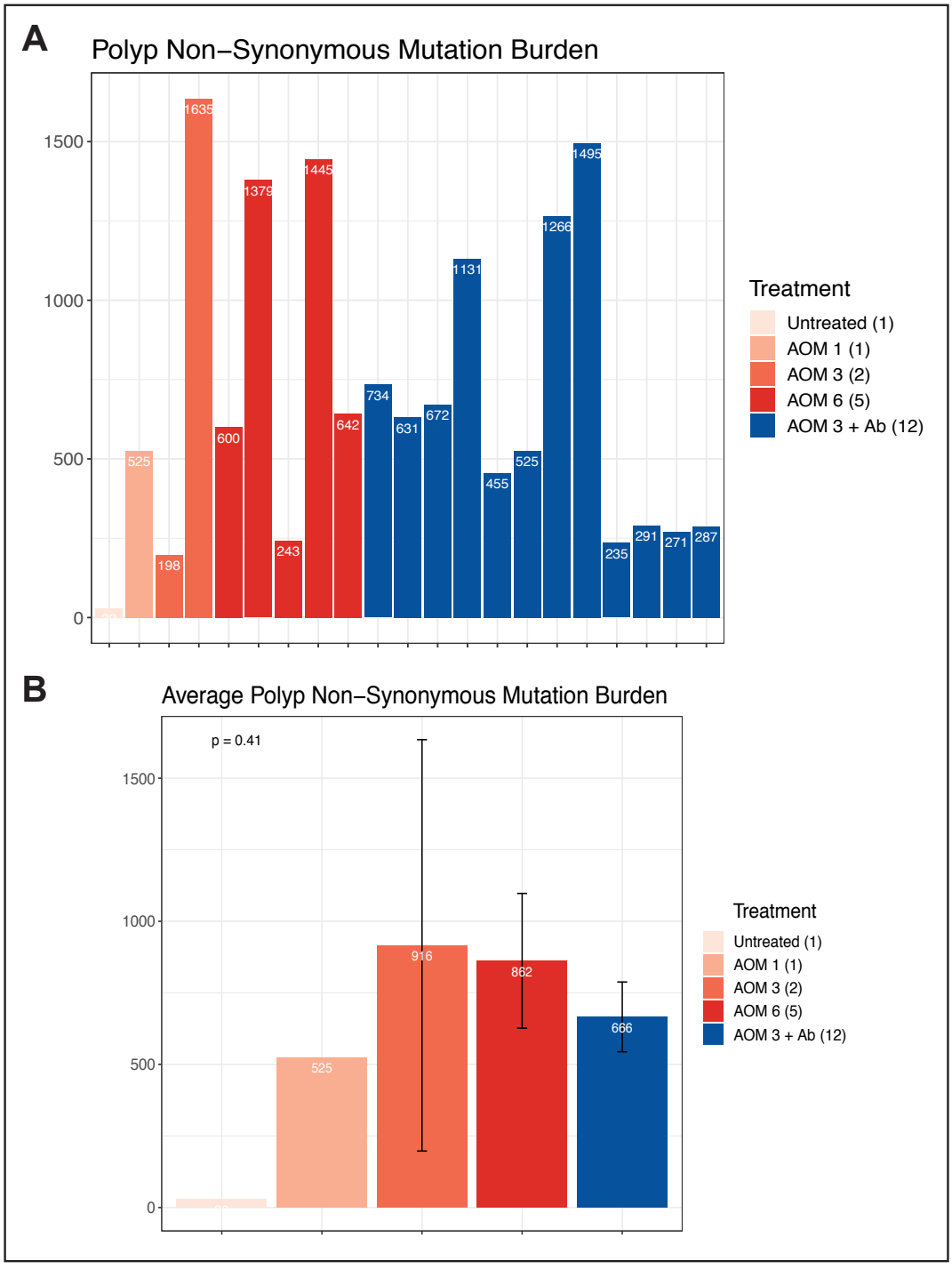


Figure 5.9: Non-synonymous mutation burden of polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Represented as number of non-synonymous mutations per sample (A) and the average number of non-synonymous mutation burdens per experimental group (B). N as indicated, error bars represent SEM, p value as calculated by one-way ANOVA.

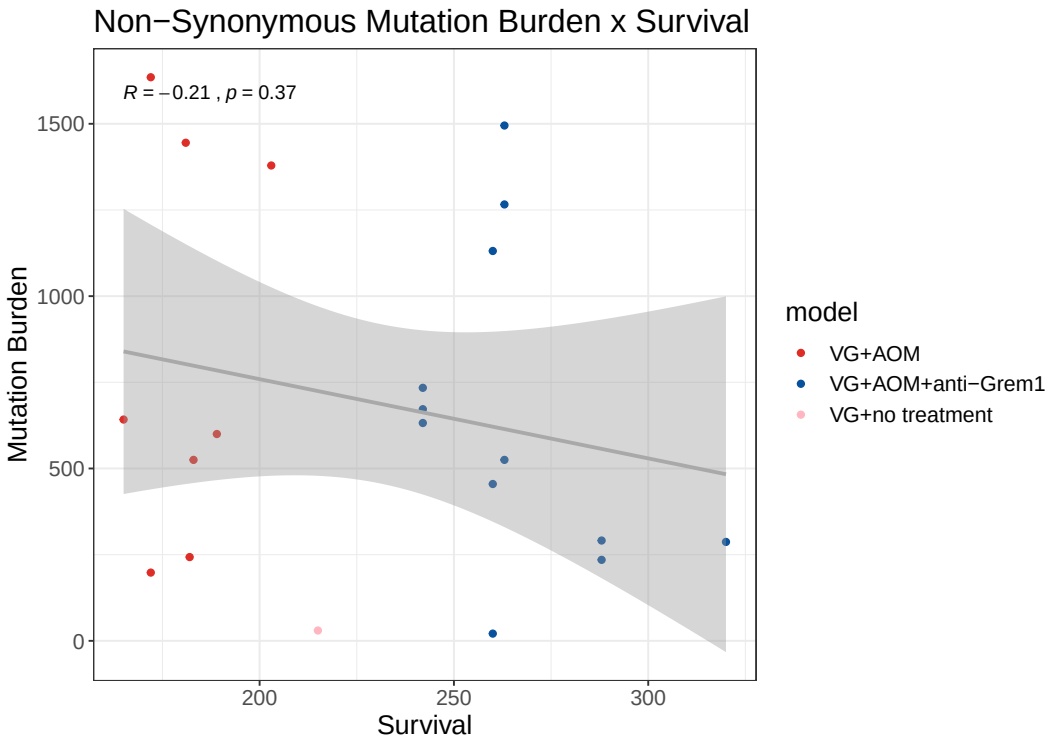


Figure 5.10: Correlation analysis of experimental polyps according to non-synonymous mutation burden and mouse survival. Points represent individual colon polyps sampled for genomic analysis, with the age of the mouse at collection represented on the y-axis. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

of the average contribution of each mutation type in VG+AOM+anti-Grem1 polyps, however, revealed a 13% increase in the average contributon of intron variants accompanied by 5% decreases in the average contribution of missense and non-genic variants compared to untreated VG+AOM polyps. The average contribution of UTR, synonymous and splice variants in anti-Grem1-treated polyps remained largely unchanged from that observed in VG+AOM polyps, indicating that anti-Grem1 therapy results in an increased proportion of intron variants within VG+AOM polyps.

5.2.3.3 Mutational signature analysis of anti-Grem1-treated *Vil1-Grem1* polyps

In order to further characterise the genomic impact of anti-Grem1 therapy, I next sought to investigate any difference in mutational processes within treated and untreated polyps from VG+AOM mice. Mutation signatures within polyps were extracted using base substitutions and information on the sequence context

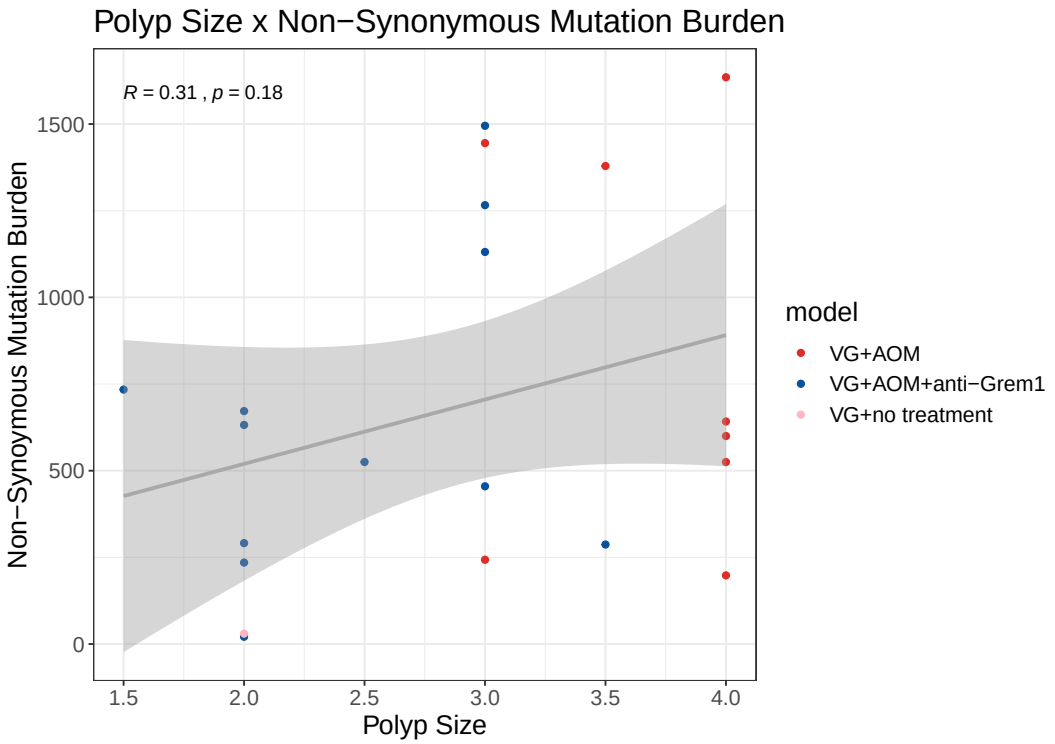


Figure 5.11: Correlation analysis of polyp size and non-synonymous mutation burden. Points represent individual colonic polyps sampled for genomic analysis. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

of each mutation (Section 2.4.3.8; Figure 5.13, A). Applying this approach to our polyp analysis revealed the presence of 8 distinct, validated mutational signatures that have been reported previously and documented in the COSMIC database (Figure 5.14. Of these, SBS1, a clock-like, aging-associated mutational process found in almost all human cancer samples, contributed 100% of mutations detected in the single untreated *Vil1-Grem1* polyp and ~10% of those detected in *Vil1-Grem1*+AOM+anti-Grem1 and *Vil1-Grem1*+AOM polyps. Interestingly, and in line with findings presented in chapter 3, the alkylation-associated SBS11 contributed the majority (~75%) of mutations in all *Vil1-Grem1*+AOM and *Vil1-Grem1*+AOM+anti-Grem1 polyps and was absent in the untreated *Vil1-Grem1* polyp. This further supports our proposal of SBS11 as an AOM-associated mutational signature. These data therefore indicate that exposure to AOM results in an alkylation-associated mutational process overtaking clock-like mutation accumulation as the primary driver of mutagenesis within AOM-exposed polyps.

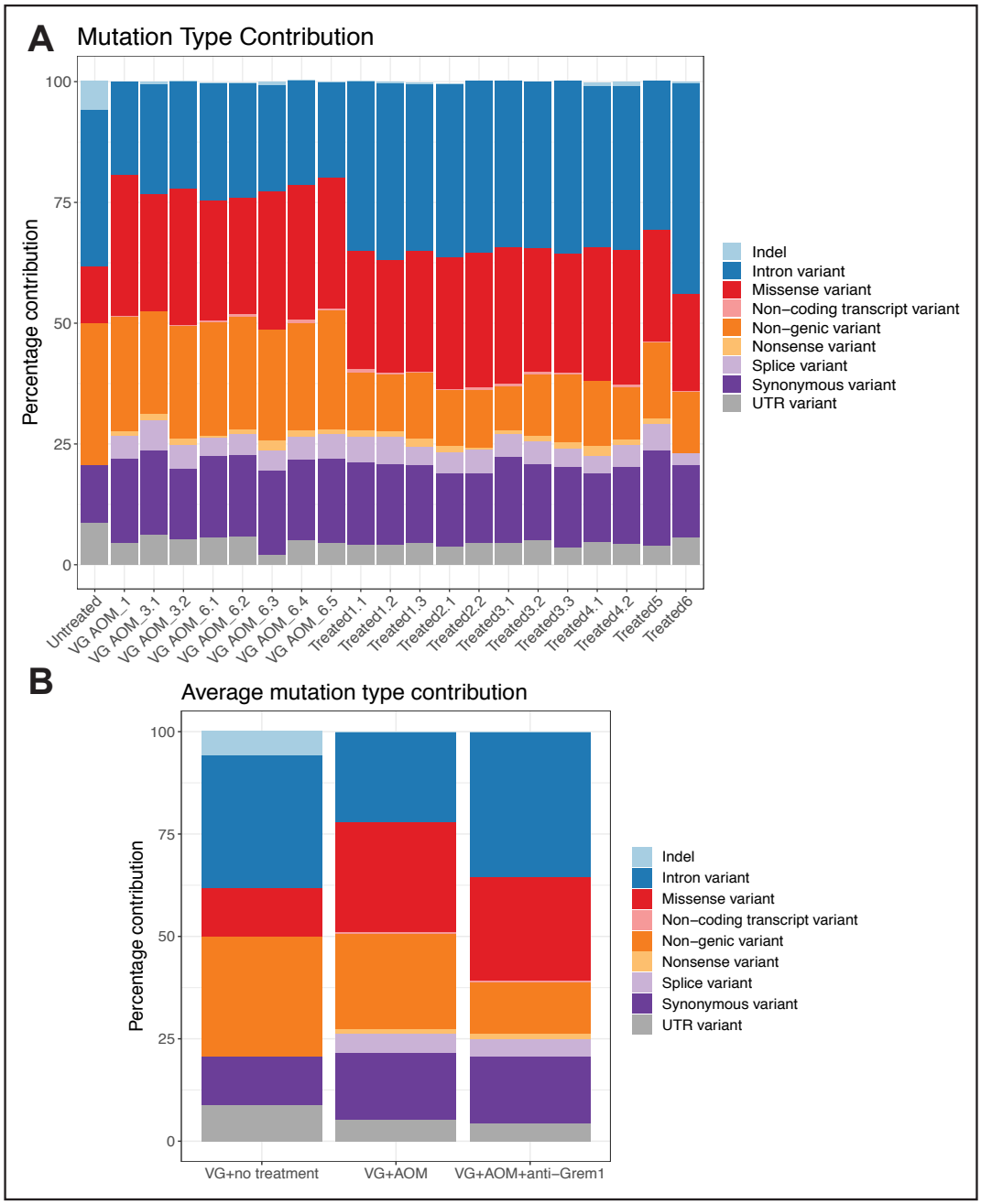


Figure 5.12: Characterisation of mutation types present in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Represented as percentage contribution of each mutation type per sample (A) and average percentage contribution of each mutation type per experimental group (B).

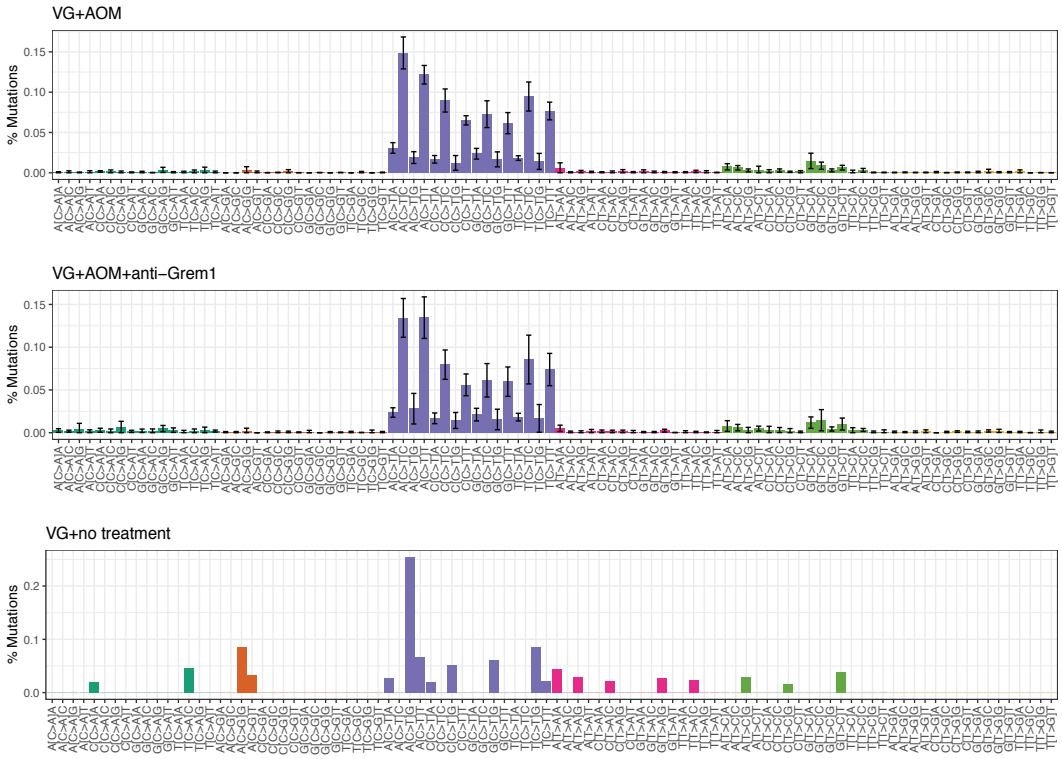


Figure 5.13: Characterisation of mutational spectrum in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Represented as percentage of all mutations per 96-substitution classification for each experimental group. Error bars represent SEM.

The remaining mutations in both *Vil1-Grem1*+AOM and *Vil1-Grem1*+AOM+anti-Grem1 polyps were largely generated by signatures associated with errors in transcription-associated nucleotide excision repair and DNA mismatch repair (SBS5, 8, 14, 20, 21 and 26), both of which may be elevated due to the global increase in genomic instability caused by AOM mutagenesis. Although SBS21 was detected in polyps from *Vil1-Grem1*+AOM+anti-Grem1 mice and not *Vil1-Grem1*+AOM mice, it accounted for less than ~10% of all mutations detected. Given that the majority (~95%) of mutations detected in both *Vil1-Grem1*+AOM+anti-Grem1 and *Vil1-Grem1*+AOM polyps were generated by SBS11, we can conclude that anti-Grem1 therapy does not result in a significant shift in the mutational processes active within polyps.

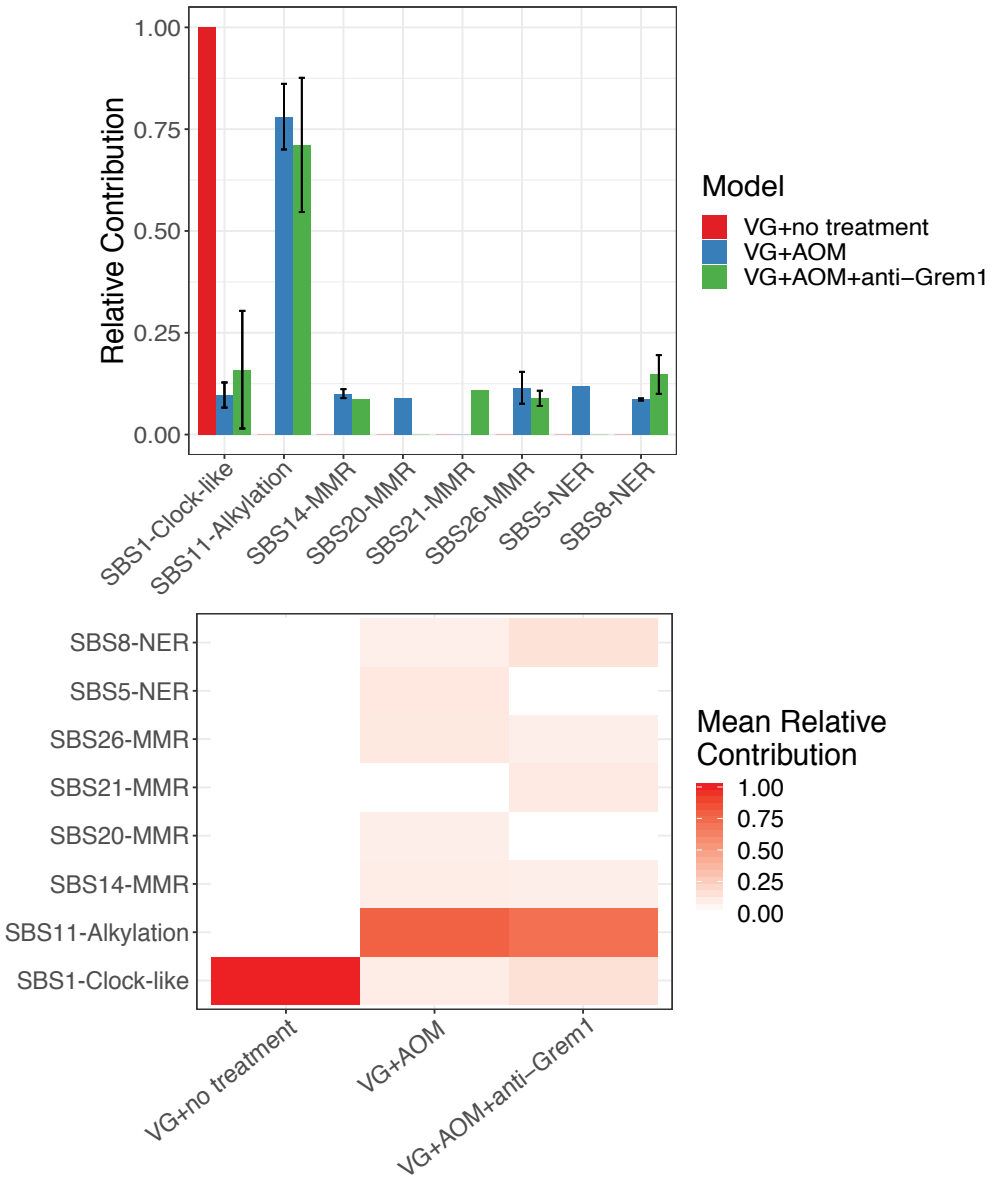


Figure 5.14: Mutational process contribution in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Represented as the relative contribution of each detected mutational signature to each experimental group (top as bar plot, bottom as heat map). Error bars represent SEM.

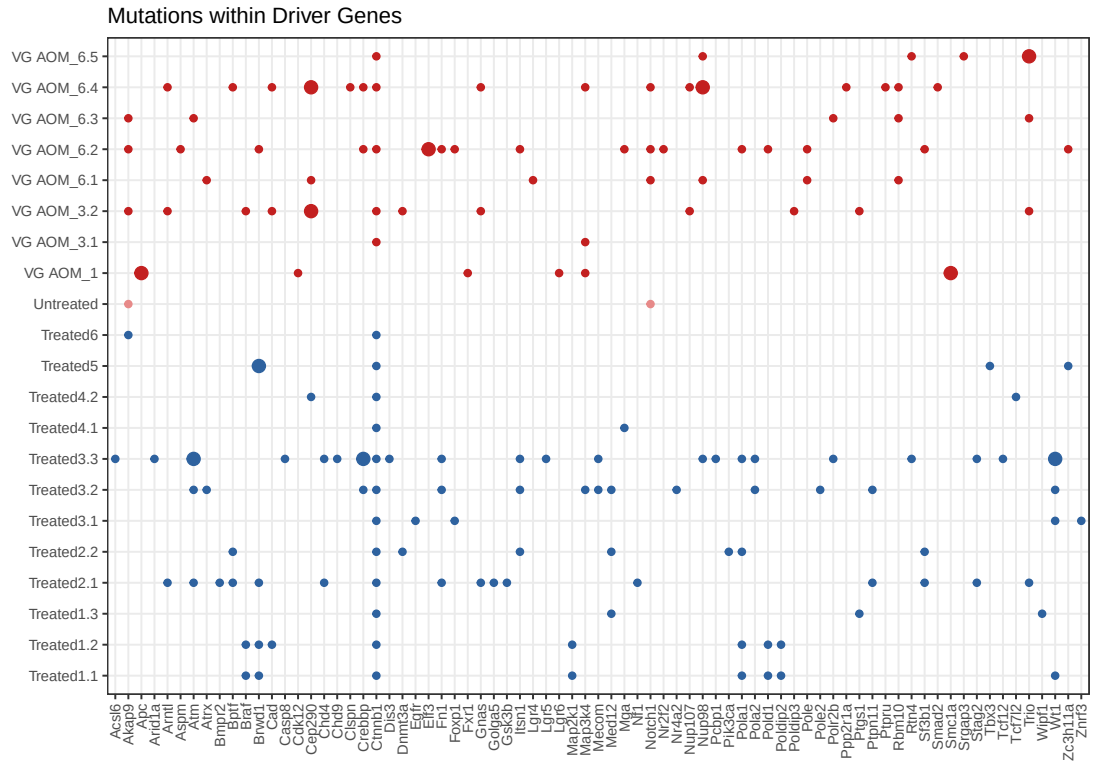


Figure 5.15: Distribution of driver gene mutations in polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Polyps treated with AOM only (blue), AOM+anti-Grem1 (red) or no treatment (pink) are indicated on the Y axis, driver genes on the x axis. Smaller dots indicate a single mutation, larger dots indicate two mutations.

5.2.3.4 Characterisation of driver gene mutations in anti-Grem1-treated *Vil1-Grem1* polyps

I next sought to identify whether a difference in mutations within putative colorectal cancer driver genes was driving the observable difference in phenotype between *Vil1-Grem1*+AOM and *Vil1-Grem1*+AOM+anti-Grem1 mice. I thus filtered total non-synonymous mutations using a set of 113 putative CRC driver genes described previously (Section 2.4.3.7). 100% of polyps analysed contained driver gene mutations (Figure 5.15), although there was a high degree of variation in the number of driver gene mutations per polyp within and between experimental groups. Overall, despite the average number of driver mutations per polyp being lower in VG+AOM+anti-Grem1 polyps than VG+AOM polyps, this reduction did not reach statistical significance (Figure 5.16), indicating that anti-Grem1 therapy does not impact upon the acquisition of driver gene mutations.

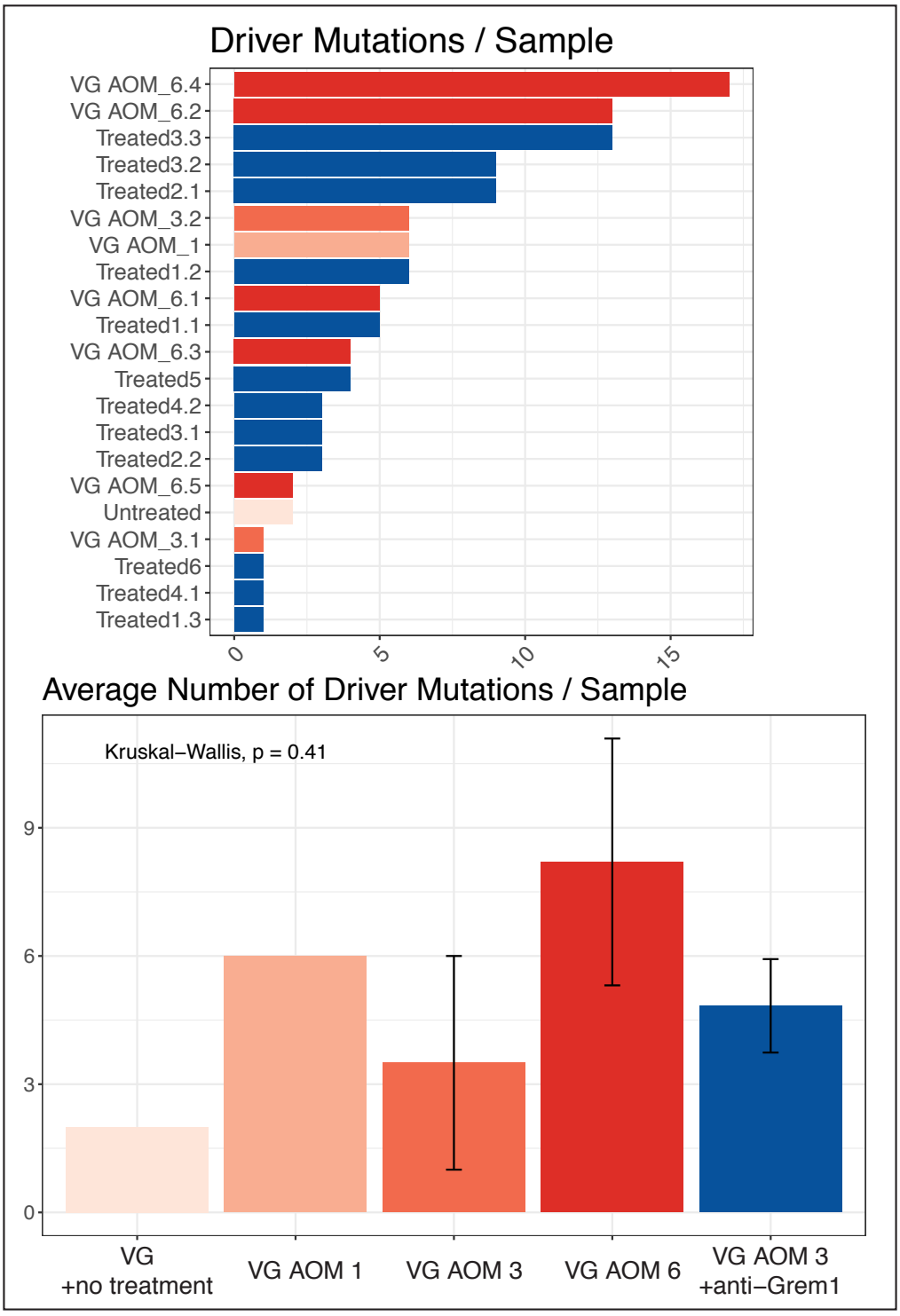


Figure 5.16: Number of driver gene mutations within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. Represented as the number of driver gene mutations per polyp (A) and the average number of driver gene mutations per polyp within each experimental condition. Error bars represent SEM, p calculated by Kruskal-Wallis test.

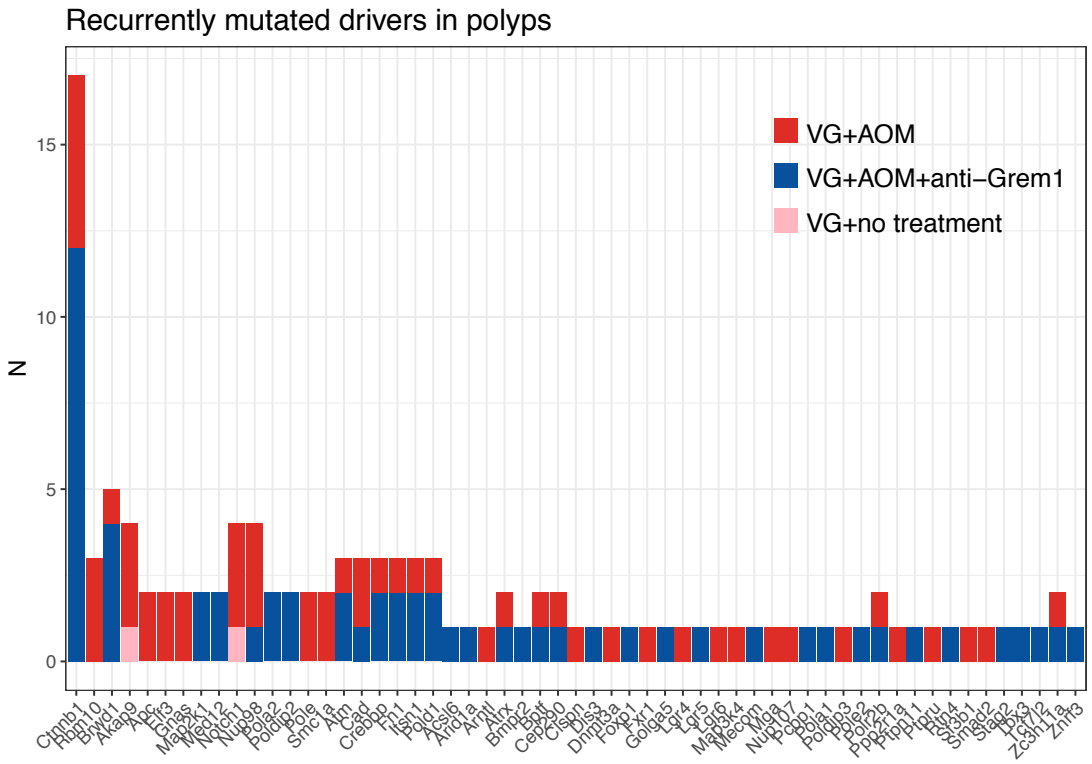


Figure 5.17: Recurrently mutated driver genes within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1. The number of contributions per driver gene for each experimental group is indicated by colour.

Analysis of recurrently mutated driver genes revealed a low degree of overlap between treated and untreated VG+AOM polyps, with the exception of *Ctnnb1*, which was mutated in all VG+AOM+anti-Grem1 polyps and 5/8 VG+AOM polyps. The untreated *Vil1-Grem1* polyp harboured only mutations within *Akap9* and *Notch1*, with *Notch1* only mutated in 3 other polyps, all from VG+AOM mice. Mutations within the *Lgr* genes (*Lgr4*, *Lgr5*, *Lgr6*) were present in 1/12 polyps from VG+AOM+anti-Grem1 mice and 3/8 of VG+AOM mice (Figure 5.17). 100% of *Ctnnb1* mutations were missense variants that result in amino acid changes at residues 32, 33, 34, 37 and 41 within the GSK3β binding site region of the resultant CTNNB1 protein (Table 5.1).

Modification of these residues impairs GSK3β binding to the N-terminus of β-catenin and signalling its degradation, leading to an accumulation of cytoplasmic β-catenin and elevated Wnt signalling levels. The amino acid substitutions caused by all *Ctnnb1* missense mutations detected here are present and validated

Mutation	Protein Change	COSMIC Entries	FATHMM prediction
94G>A	Asp32Asn	136	Pathogenic (0.98)
98C>T	Ser33Phe	158	Pathogenic (0.96)
101G>A	Gly34Glu	119	Pathogenic (0.98)
110C>T	Ser37Phe	302	Pathogenic (0.96)
122C>T	Thr41Ile	141	Pathogenic (0.98)

Table 5.1: Missense *Ctnnb1* mutations detected within *Vil1-Grem1* polyps

within the COSMIC database of human cancer mutations, with a range of 119 to 302 entries per mutation. Furthermore, all 5 mutations have been predicted to be 'pathogenic' by the FATHMM-MKL algorithm, with functional p-value scores of 0.96 or 0.98. This algorithm predicts the functional, molecular and phenotypic consequences of protein missense variants using hidden Markov models [165]. Functional missense mutations within the *Ctnnb1* gene therefore appear to provide a proliferative Wnt drive in VG mice exposed to AOM, as widely documented and demonstrated previously in Chapter 3.

Interestingly, however, missense mutations within genes involved in DNA repair, such as those encoding proofreading polymerase proteins (*Pola1*, *Pola2*, *Pold1*, *Poldip2*, *Pole*, *Pole2*) and the DNA damage response protein ATM, were more prevalent in polyps from mice treated with anti-Grem1, present in 6/12 anti-Grem1 treated polyps and 2/8 AOM-only polyps. As discussed previously, *Atm* is a central gene involved in the regulation of DNA repair, such that *Atm*-deficient cells exhibit high rates of intrachromosomal recombination and error-prone recombination [166]. Similarly, *Pola1*, *Pola2*, *Pold1*, *Poldip2*, *Pole*, *Pole2* are all polymerase-associated genes with important roles in DNA repair and replication. Indeed, somatic loss of *POLE* in human cancers is associated with a hypermutator phenotype [167]. Assessment of the missense mutations detected within these genes showed that the resultant amino acid changes fell within conserved functional or regulatory domains of the resultant proteins (Table 5.2) and were predicted to be 'damaging' by the PROVEAN functional impact analysis tool [150].

Gene	Protein Change	PROVEAN Prediction	Notes
Atm	Thr1615Ala	Damaging	Coiled coil region implicated in protein interactions [168]
Atm	Leu2920Phe	Damaging	Highly conserved regulatory PI3K/PI4K catalytic domain [157]
Atm	Ser2505Pro	Damaging	Highly conserved regulatory FAT catalytic domain [169]
Pola1	Glu1200Gln	Damaging	Alpha zinc finger helix motif within DNA binding domain
Pola2	Thr130Ile	Damaging	N-terminal domain involved in complex formation
Pola2	Gly89Glu	Damaging	N-terminal domain involved in complex formation
Poldip2	Pro363Ser	Damaging	apaG domain involved in protein-protein interactions and substrate-specificity
Poldip2	Pro363Ser	Damaging	apaG domain involved in protein-protein interactions and substrate-specificity
Pole	Gly395Arg	Damaging	Highly conserved DNA, dNTP and pyrophosphate binding domain
Pole2	Ser516Asn	Damaging	C terminus catalytic sub-unit domain involved in regulating polymerase activity [170]

Table 5.2: Missense mutations in genes involved in DNA repair

The missense mutations detected within these genes thus correspond specifically to processes involved in the maintenance of genomic integrity. Furthermore, these mutations were present at lower VAFs (0.12 - 0.2) in untreated VG+AOM polyps compared to much higher VAFs in VG+AOM+anti-Grem1 polyps (0.32 - 0.5), indicating that these are clonally expanded, and thus selected and functional, 'truncal' mutations in treated polyps (Figure 5.18).

Stratification of the mutational processes within polyps according to DNA repair mutations revealed that SBS14, associated with defective DNA mismatch repair, was unique to polyps harbouring a missense DNA repair mutation (Figure 5.19). These findings suggest that the missense mutations in DNA repair-associated genes documented above may have a functional impact upon mutation acquisition within the polyps they are detected in.

Taken together, these findings indicate that VG+AOM+anti-Grem1 polyps disproportionately harbour clonally expanded, potentially functionally damaging

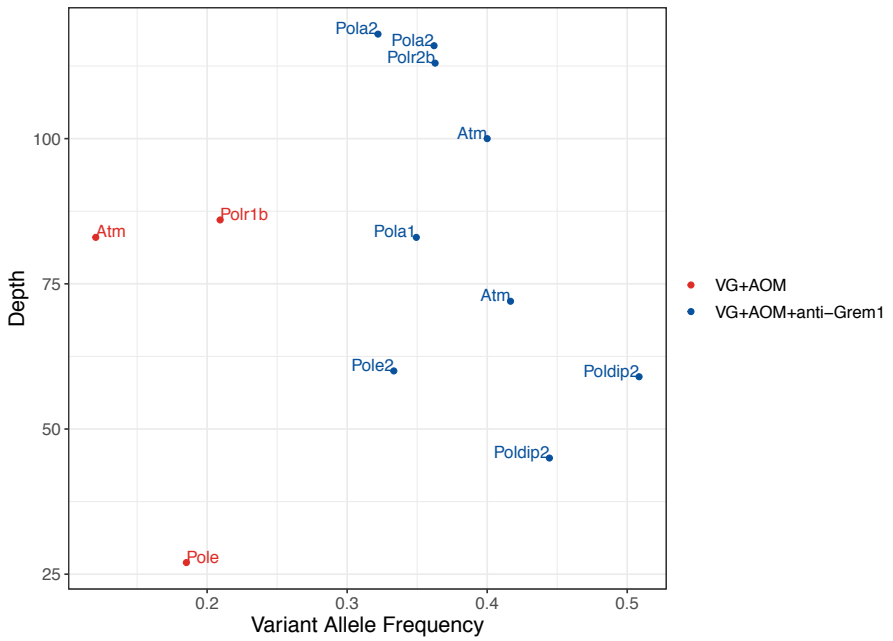


Figure 5.18: Depth and variant allele frequency of missense mutations detected in DNA repair-associated genes.

missense mutations in genes involved in DNA repair compared to untreated VG+AOM polyps. This somewhat perplexingly indicates that a higher proportion of VG+AOM+anti-Grem1 polyps are 'repair impaired' compared to untreated VG+AOM polyps. It may be that acquisition of these mutations in a subset of VG+AOM polyps is an early event that facilitates evasion from therapy. In order to explore this further, I sought to re-assess polyp mutation burden while controlling for DNA repair status. I therefore analysed non-synonymous and driver gene mutation burdens according to the number of predicted 'damaging' missense mutations in DNA repair-associated genes (Figure 5.20).

The majority of *Vil1-Grem1*+AOM polyps (6/8) did not harbour any missense DNA repair mutations yet exhibited a wide range of non-synonymous mutation burdens. The VG+AOM polyps that each harboured one missense DNA repair mutation exhibited no relationship between non-synonymous mutation burden and missense DNA repair mutation ($R = -0.02$, $p = 0.95$). This indicates that VG+AOM mice generate polyps with varied mutation burdens largely in the absence of missense DNA repair mutations. In polyps from *Vil1-Grem1*+AOM+anti-Grem1

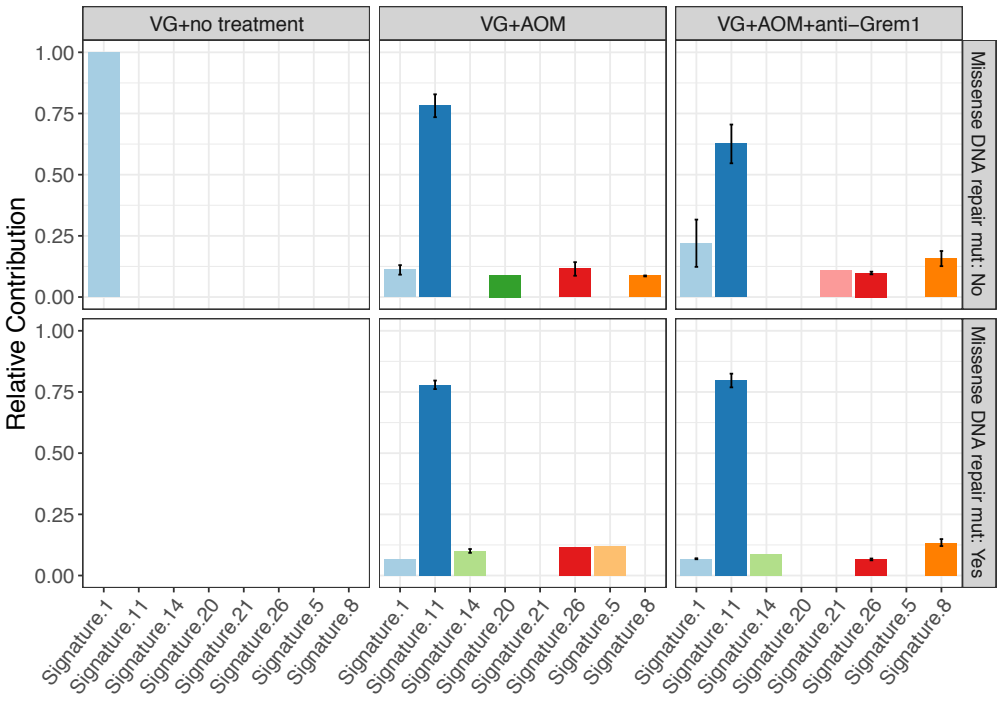


Figure 5.19: Average relative mutation signature contribution within polyps from each experimental model according to DNA repair mutations (No = No DNA repair mutation, Yes = Missense DNA repair mutation)

mice, however, there was a strong, statistically significant positive correlation between non-synonymous mutation burden and DNA repair mutations (R = 0.89, p = 0.00012), driven by a greater number of polyps harbouring at least one missense DNA repair mutation (6/8), with two polyps harbouring 2 and 3 mutations, respectively. The missense mutations within DNA repair genes detected within anti-Grem1-treated polyps are therefore associated with a putative 'hypermutated' phenotype. This observation held when analysing mutations within driver genes (Figure 5.21), with polyps from *Vil1-Grem1*+AOM mice exhibiting no relationship between the number of driver gene mutations and missense DNA repair mutations (R = 0.19, p = 0.65) despite harbouring a wide range of driver gene mutations per polyp (2-18), while those from *Vil1-Grem1*+AOM+anti-Grem1 mice exhibited a strong and statistically significant positive relationship (R = 0.93, p = 9.8e-06).

I went on to reveal that anti-Grem1 therapy results in a reduction in non-synonymous mutation burden that is dependent upon the absence of missense DNA repair genes. DNA repair-proficient polyps (defined as those not harbouring

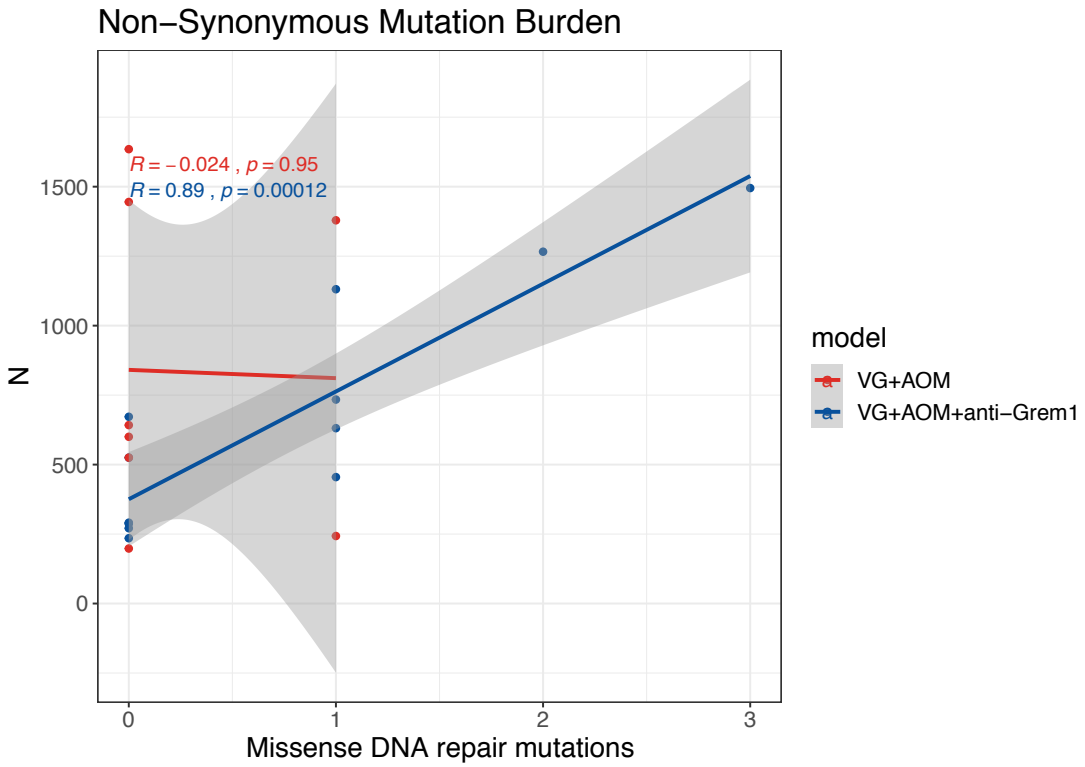


Figure 5.20: Correlation analysis of polyp non-synonymous mutation burden and number of missense DNA repair-associated mutations. Each point represents a colon polyp sampled for genomic analysis. Shading represents 95% CI, R and p-value as calculated by Pearson correlation analysis.

the missense DNA repair mutations described above) from VG+AOM mice exhibited an average of 840 mutations, while those from VG+AOM+anti-Grem1 mice harboured an average of 380 mutations (Figure 5.22). Given that *Vil1-Grem1*+AOM mice treated with anti-Grem1 exhibit fewer polyps, these results indicate that anti-Grem1 therapy can abrogate macroscopic polyp burden and reduce mutation burden in some residual lesions, but may promote accrual of functional DNA repair-associated mutations in others, leading to hypermutation and polyp escape.

5.2.4 Summary: Anti-Grem1 therapy in *Vil1-Grem1* mice

Addressing the aims set out at the beginning of this chapter, the key findings presented here are:

– Anti-Grem1 therapy rescues survival and polyp burden of VG+AOM mice,

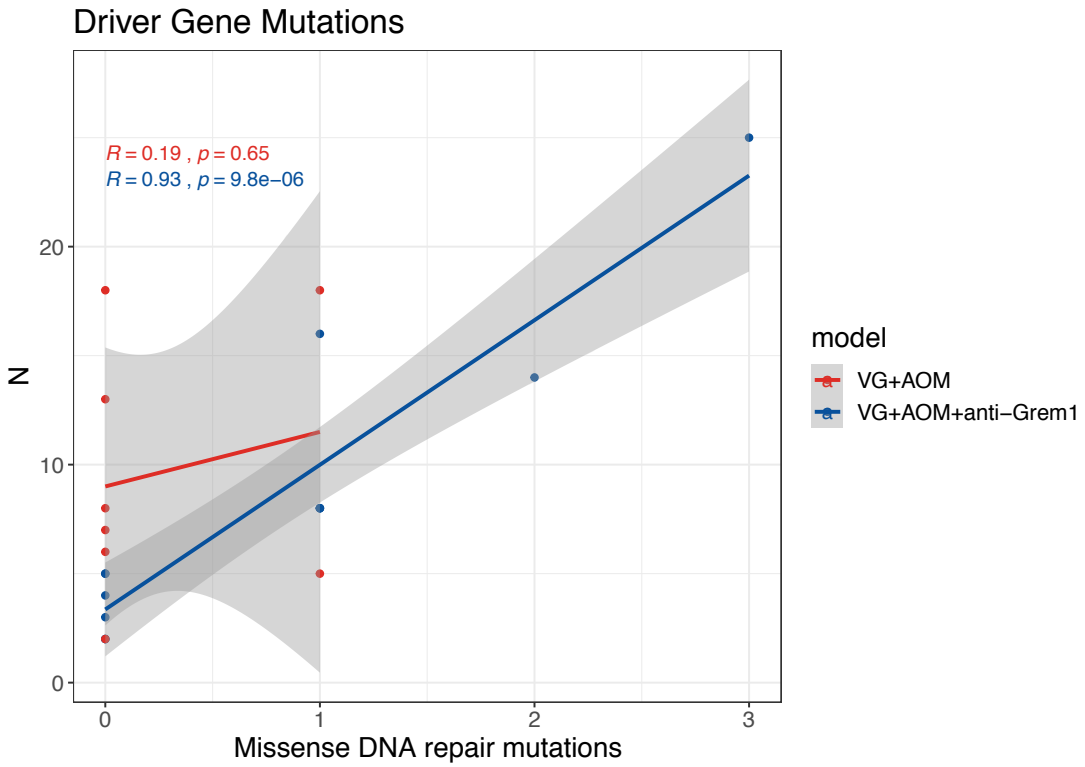


Figure 5.21: Correlation analysis of polyp driver gene mutation burden and number of missense DNA repair mutations. Each point represents a colon polyp sampled for genomic analysis. Shading represents 95% CI, R and p as calculated by Pearson correlation analysis.

predominantly through abrogation of SB phenotype but also through a reduction in the number of proliferative cells within the colon that are susceptible to AOM-driven mutagenesis

– *Ctnnb1* mutation is universal in VG+AOM polyps, in line with earlier findings, but VG+AOM+antiGrem1 polyps harbour increased missense mutations in DNA repair and polymerase genes

– In VG+AOM polyps that do not harbour missense mutations in DNA repair genes, treatment with anti-Grem1 results in a reduction in non-synonymous mutation burden

– A majority of VG+AOM+anti-Grem1 colonic polyps that evade therapeutic rescue, despite remaining smaller on average than untreated polyps, harbour apparently functionally impactful mutations within DNA repair genes.

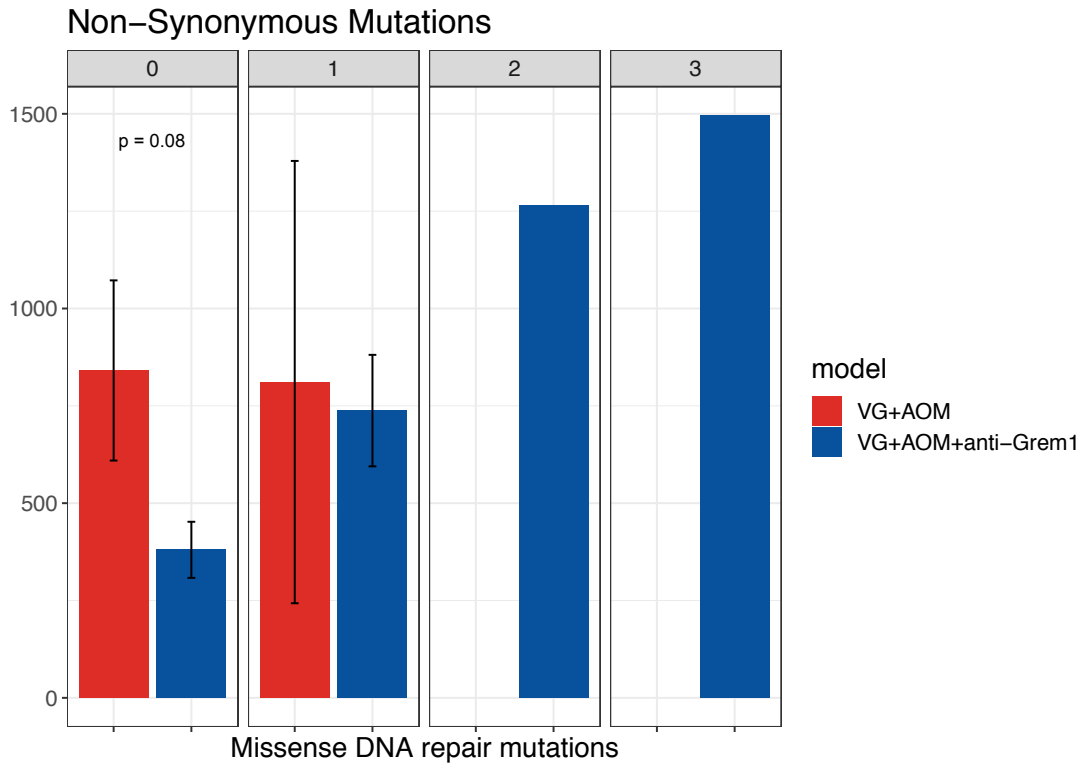


Figure 5.22: Average non-synonymous mutation burden within polyps from *Vil1-Grem1* mice exposed to AOM and treated with anti-Grem1 according to number of missense DNA repair mutations. Represented as average non-synonymous mutation burden per polyp for each experimental group, panels represent indicated number of missense DNA repair mutations. Error bars represent SEM, p value as calculated by unpaired t-test.

These polyps also exhibit an increased proportion of intronic variants, presumably as a consequence of a global increase in genomic instability as a result of their 'repair impaired' status

– Anti-Grem1-treated polyps harbouring missense DNA repair-mutations exhibit a putative 'hypermutated' phenotype in which non-synonymous mutation burden correlates strongly with the number of DNA repair-mutations. These 'repair impaired' polyps exhibit a unique, DNA repair-associated mutational signature not detected in polyps lacking missense DNA repair mutations

5.3 Discussion

Pathogenesis in VG AOM polyps appears to be driven primarily by disruption in cell fate determination, with the resultant increase in proliferative cells driving increased susceptibility to AOM-induced mutagenesis and subsequent colon phenotype severity. This is evidenced by therapeutic restoration of cell fate determination resulting in the partial rescue of colon phenotype, polyp burden and survival in VG+AOM+anti-Grem1 mice. In view of my original hypothesis, therapy thus drives a reduction in the mutationally susceptible proliferative cell compartment, resulting in reduced numbers of colon polyps, which also harbour fewer mutations.

Interestingly, whilst anti-Grem1 treatment may thus reduce the number of mutations and mutation susceptible cells initially, it may also skew cells towards DNA repair mutation acquisition and a secondary, 'hypermutator' phenotype thereafter that enables escape from a therapeutic selective pressure. This is evidenced by a majority of the VG+AOM+anti-Grem1 polyps that evade therapeutic rescue harbouring mutations that may functionally impact genomic stability. This additional mutagenic drive may thus be sufficient to overcome therapeutic restoration of cell fate determination, so that these polyps continue to accumulate mutations and remain detectable with increased mutation burdens, as compared to both untreated VG+AOM polyps and treated VG+AOM polyps with no DNA repair mutations.

The findings presented in Chapter 4, whereby crosstalk between BMP signalling and DDR pathways drives increased genomic instability within ectopic crypts, may also be relevant here. Anti-Grem1 therapy in VG mice restores both cell fate determination and, on the basis that DDR is not simply a cell-intrinsic property but can in fact be microenvironmentally influenced, Grem1-driven impairment of homeostatic mutation protection mechanisms. These therapeutic outcomes increase survival, stall polyp progression and produce a detectable reduction in proliferative cells and mutation burden in anti-Grem1-treated mice.

Under the mutagenic pressure of AOM exposure, however, somatic mutations that impact genomic stability are optimally selected in treated polyps. These polyps thus remain detectable, harbouring a reduction in proliferative cells but significantly increased mutation burdens.

The missense, DNA repair-associated mutations disproportionately detected in anti-Grem1-treated VG AOM polyps are either a result of an ectopic crypt cell-intrinsic skew to DNA repair mutations or an AOM-driven skew to DNA repair mutations. The latter of these options would be supported in the literature if AOM exposure were associated with impaired DNA repair and a hypermutator phenotype. However, the genomic instability caused by AOM mutagenesis has been proposed to be independent of DNA repair mutations [171]. We are therefore left with the alternative hypothesis, whereby ectopic crypts in VG mice harbour a disease-intrinsic skew to DNA repair mutations. Alternatively, therapy may apply a selective pressure which enhances the emergence of an alternative pathway to genomic instability that involves the acquisition of mutations in DNA repair-associated genes.

There is putative evidence of a similar skew to genomic instability in humans. Traditional serrated adenomas within the colon, the sporadic equivalent of VG polyps, are thought to occur via 'non-conventional' carcinogenic pathways involving a non-*LGR5* cell of origin. TSAs are also more prone to DNA repair-deficiency [172], although in human disease this is thought to be the sequelae of hypermethylation [173]. Crucially, while early Wnt-activating mutations in *LGR5*+ stem cells are sufficient for conventional adenoma formation, TSAs acquire different activating mutations, in *KRAS* and *BRAF*, that provide a more graduated, cumulative selective advantage (Section 1.5.6). It could thus be that a similar, 'non-conventional' tumourigenic pathway is optimally selected in treated VG lesions, in which the cell of origin is a non-*Lgr5* stem cell and constitutive Wnt activation may thus be less effective. An updated model of the impact of anti-Grem1 therapy on mutation acquisition in VG+AOM polyps, incorporating this theory and the findings presented here, is illustrated in Figure 5.23.

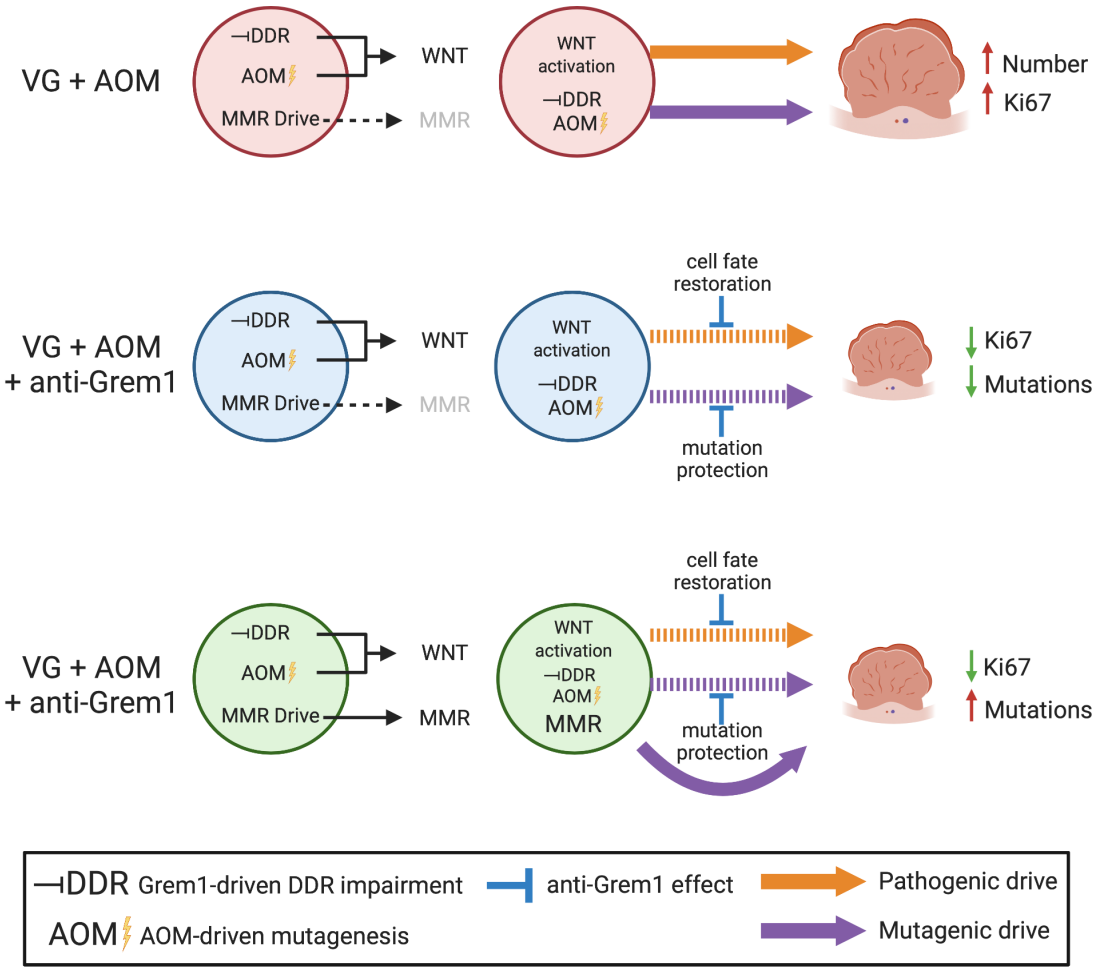


Figure 5.23: A model of the impact of anti-Grem1 therapy in *Vil1-Grem1* mice treated with AOM. In untreated VG AOM mice (red circle), the disease-intrinsic skew to MMR mutations within ectopic crypts is outcompeted by synergistic or competing mutagenic drives (arrows) from Grem1-driven DDR impairment and AOM mutagenesis, which result in ubiquitous acquisition of activating Wnt mutations and subsequent tumourigenesis. In VG+AOM+anti-Grem1 mice (blue circle), therapy restores cell fate determination and Grem1-driven impairment of DDR. These therapeutic outcomes thus overcome the pathogenic, proliferative drive of *Ctnnb1* activation and counteract AOM-induced DNA damage. Progression of these polyps is therefore stalled, and they harbour a detectable reduction in proliferative cells and mutation burden compared to untreated VG AOM polyps. In some VG AOM polyps (green circles), therapeutic neutralisation of Grem1-driven DDR impairment reduces its competitive mutagenic pressure and the combination of unimpeded disease-intrinsic skew to MMR deficiency with AOM mutagenesis results in the acquisition of MMR-associated mutations. In these polyps, therapy overcomes the pathogenic drive of Wnt activation and Grem1-driven DDR impairment, but is unable to neutralise the mutagenic drive and subsequent hypermutated phenotype associated with MMR-deficiency. These polyps thus evade therapeutic rescue and remain detectable, harbouring a reduction in proliferative cells but significantly increased mutation burdens.

The work presented here may, however, be limited by low mouse numbers limiting the number of polyps available for genomic analysis. Furthermore, the use of a single experimental endpoint inherently involves the sampling of polyps that have evaded therapeutic rescue, which may thus have introduced selection bias to these results. Given the rescue of colon phenotype by anti-Grem1 therapy, the identification and sampling of colon polyps may also have been prone to error. The functional significance of the missense mutations detected in DNA repair-associated genes also requires validation. Further work could therefore involve specific interrogation of methylation and copy-number variation in treated and untreated VG AOM polyps, in order to shed light on the mechanism by which anti-Grem1 therapy impacts upon genomic stability. Multi-regional sequencing of treated polyps would also enable the construction of phylogenetic trees, thus revealing the timing and type of activating mutations acquired in treated and untreated polyps. Furthermore, increased sampling of pre-neoplastic tissue would demonstrate the effect of therapy in the earlier stages of tumourigenesis and reveal the mutational landscape of remodelled epithelium. As mentioned previously, incorporating multi-regional sampling and increased sequencing depth would also shed light on the origin of the mutations detected. Whether 'resistance' mutations are monoclonal or polyclonal in origin is an important question that will shed light on the mechanism by which therapy acts as a selective pressure within treated polyps. Integrating these findings with those presented earlier, characterisation of miR-96 expression in anti-Grem1 treated epithelium could also confirm the proposed Grem1-miR-96 axis within ectopic crypts.

"Look harder" - Rafiki, The Lion King

6

The impact of therapy on mutation acquisition in 129SvEv.*Rag*$^{-/-}$ mice

Contents

6.1 Introduction 167
6.2 Results 168
6.2.1 Phenotypic characterisation 168
6.2.2 Genomic analysis 171
6.2.2.1 Mutation burden analysis of anti-IL-22-treated normal and polyp tissue 172
6.2.2.2 Characterisation of mutation types within anti-IL-22-treated normal and polyp tissue 175
6.2.2.3 Mutation signature analysis of anti-IL-22-treated normal and polyp tissue 177
6.2.3 Transcriptomic analysis 184
6.2.3.1 Characterisation of disease progression 184
6.2.3.2 Intestinal cell populations and signalling pathways 193
6.2.3.3 The effect of anti-IL-22 therapy on disease progression 199
6.2.3.4 Intestinal cell populations and signalling pathways in treated samples 206
6.2.4 Summary of findings 213
6.3 Discussion 214

6.1 Introduction

129SvEv.*Rag*$^{-/-}$ mice fail to express the functional RAG proteins that are required for somatic recombination of T-cell receptor (TCR) and immunoglobulin (Ig) genes. In the absence of these proteins, *TCR* and *Ig* genes cannot rearrange, resulting in a lack of mature T and B cells [174]. These mice are therefore immunodeficient and develop a chronic innate colitis upon infection with the mouse pathogen *Helicobacter hepaticus* (*Hh*)[175]. In the presence of the mutagen AOM, described previously in Chapter 3, this progresses to invasive adenocarcinoma [57, 58]. 129SvEv.*Rag*$^{-/-}$ mice exposed to both *Hh* and AOM therefore represent a model of colitis-associated cancer (CAC), as dysplasia in these animals has been shown to be driven by a combination of genetic susceptibility and environmental factors [175]. This model thus parallels human conditions such as Inflammatory Bowel Disease (IBD) where a chronically deranged inflammatory response results in an increased risk of colon cancer [176]. In 2013, Kirchberger *et al* sought to investigate the immune cells and cytokines that govern the transition from colitis to colon cancer within 129SvEv.*Rag*$^{-/-}$ mice exposed to *Hh* and AOM [57]. Their work identified a specific role for the cytokine IL-22 in the development and maintenance of invasive CAC. IL-22 mainly acts through binding the IL-22 receptor (IL-22R) on epithelial cells, thereby leading to activation of STAT3 signalling, which has been shown to be a key driver of carcinogenesis [177]. This led them to utilize an anti-IL-22 antibody to neutralise IL-22 within the intestinal environment, which almost completely abrogated STAT3 phosphorylation and epithelial proliferation, resulting in a resolution of high grade dysplasia and invasive CRC. However, despite characterising the immune and histological response to anti-IL-22 treatment, the mutational profile of treated and untreated *Hh*+AOM tumours was not explored. I was therefore keen to investigate the impact of abrogation of IL-22 signalling, and subsequent amelioration of colitis, on mutation susceptibility and acquisition. I explored this by recapitulating the mouse model used in the 2013 study and utilising exome and RNA sequencing to

characterise normal and tumour tissue from experimental 129SvEv.*Rag*$^{-/-}$ treated with anti-IL-22. Specific aims were:

- To recapitulate and characterise the intestinal phenotype of 129SvEv.*Rag*$^{-/-}$ mice exposed to *Hh* and AOM and treated with anti-IL-22 or isotype-mAb
- To explore the impact of abrogation of inflammatory drive on the genomic and transcriptomic profile of treated and untreated colonic tumours

6.2 Results

In order to explore the effect of anti-IL-22 therapy, 129SvEv.*Rag*$^{-/-}$ mice were fed 3× on consecutive days with *Hh* (1.0 × 108 CFU) by oral gavage (Section 2.2.4.3) and intraperitoneally injected with AOM once a week at a dose of 10 mg/kg for 5 weeks, starting 6 weeks after *Hh* infection [57, 130]. Treatment groups were mixed in cages to minimize cage effects and animals were treated at weeks 12–19 with either anti-IL-22 or an isotype mAb and analysed at week 21 (Section 2.2.4.4; Figure 6.1). At the experimental endpoint, week 21, mice were sacrificed and the intestinal tract collected prior to sampling of colonic polyps and adjacent normal tissue for DNA extraction (section 2.2.6.2). The colon was then swiss rolled and processed for immunohistochemical analysis of intestinal epithelial markers (Section 2.2.6). The objective of this experimental design was to expose *Hh*-driven, regenerative inflammation to AOM-driven mutagenesis and assess the impact of anti-IL-22 therapy upon any subsequent neoplasia.

6.2.1 Phenotypic characterisation

As shown in Figure 6.2, there was no significant difference in colon length between experimental groups. There was, however, a significant increase in colon crypt length between mice exposed to AOM-only (355 um) and *Hh*+AOM+isotype (*Hh*+AOM from here, 449 um). Anti-IL-22 treatment resulted in a statistically significant reduction in colon crypt length compared to all other experimental groups (205 um) (Figure 6.2; C). Analysis of blinded pathology scoring data

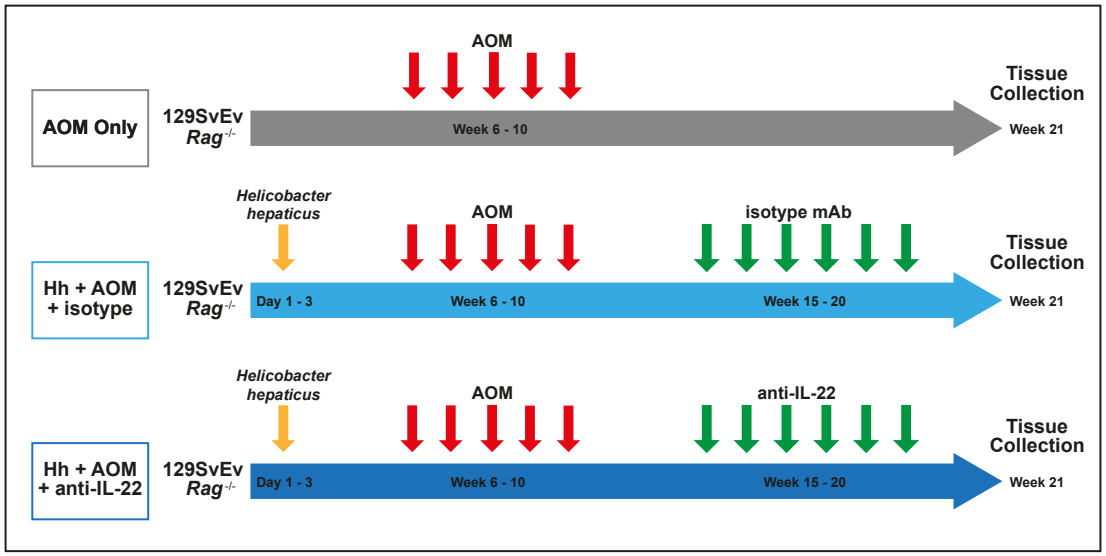


Figure 6.1: Treatment scheme for 129SvEv.*Rag*$^{-/-}$ mice exposed to AOM only or *Hh*+AOM and treated with an anti-IL-22 or isotype mAb.

confirmed the significant anti-IL-22-driven reduction in dysplasia and dysplastic lesions per colon in *Hh*+AOM mice. However, although the average number of colon adenocarcinomas was lower in anti-IL-22 treated mice, this did not reach statistical significance. Taken together, these findings indicate that *Hh*-driven inflammation results in a thickening of AOM-exposed 129SvEv.*Rag*$^{-/-}$ colonic epithelium that is rescued by anti-IL-22 therapy. Furthermore, there is a significant reduction in colonic dysplasia and associated lesions in *Hh*+AOM mice treated with anti-IL-22. Exposure to *Hh*+AOM also produced detectable regions of activated, nuclear β-catenin, which were absent within those colons treated with anti-IL-22 (Figure 6.2; B). This indicates that anti-IL-22 therapy may rescue regions of constitutively activated Wnt signalling caused by *Hh*+AOM exposure. In order to further characterise the anti-IL-22-driven reduction in colon phenotype and crypt length, I analysed the number of proliferative, apoptotic and progenitor cells within crypts from *Hh*+AOM or *Hh*+AOM+anti-IL-22 mice (Figure 6.3; A). The results indicate that anti-IL-22 therapy results in a statistically significant, 50% reduction in the percentage of proliferative cells per crypt in *Hh*+AOM exposed colons (30%) versus those treated with anti-IL-22 (15%), as indicated by Ki67 staining. This is accompanied by a 6% increase in the percentage of progenitor

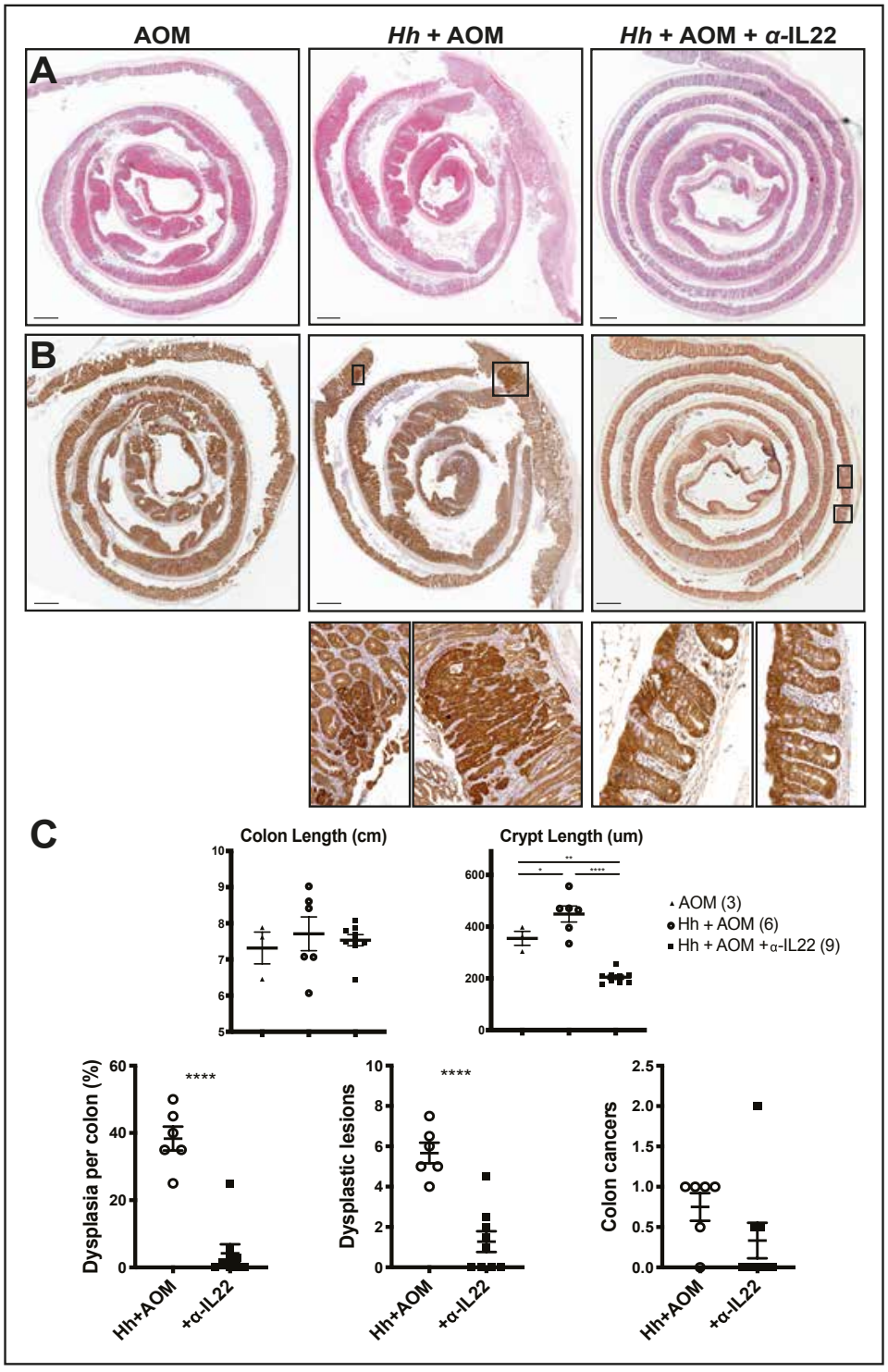


Figure 6.2: Characterisation of colon phenotype in 129SvEv.*Rag*$^{-/-}$ mice exposed to AOM only or *Hh*+AOM and treated with anti-IL-22 or isotype m-Ab. Representative images of A) hematoxylin and eosin (H&E) and B) β-catenin staining of whole intestine (middle) and focus (bottom) per indicated experimental condition. Scale bars represent 1 mm. C) Quantification of colon and colonic crypt length (top) and colon pathology scores (B) per mouse per experimental group. N as indicated, error bars represent SEM, *p < 0.05, **p < 0.005, ****p < 0.0001 as calculated by one-way ANOVA with Tukey post-test or student's t-test.

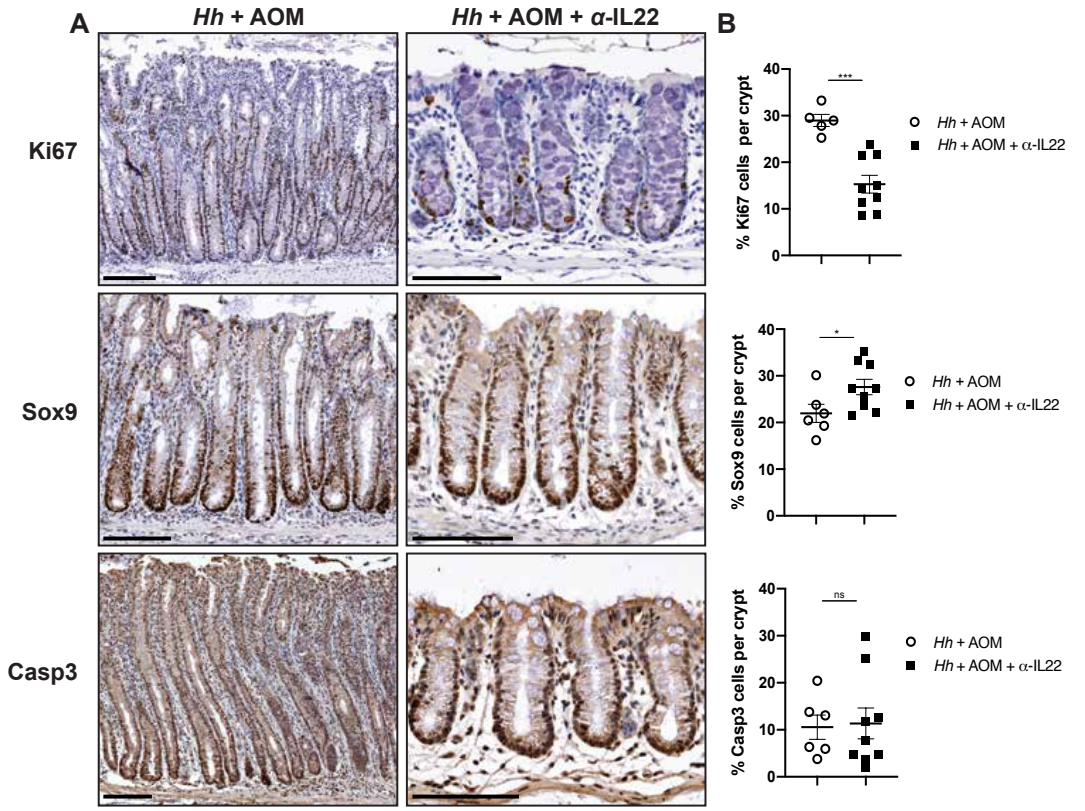


Figure 6.3: Characterisation of proliferative and progenitor cell number and apoptosis in 129SvEv.*Rag*$^{-/-}$+*Hh*+AOM mice treated with anti-IL-22 or isotype-mAb. A) Representative images of Ki67, Sox9 and Caspase-3 staining of colon tissue per experimental condition. Scale bars represent 100 μm. B) Quantification of the number of Ki67, Sox9 and Caspase-3 positive cells per colonic crypt. Symbols represent average per mouse per experimental condition. Error bars represent SEM. *p = <0.05, ***p = <0.0005, as calculated by unpaired t-test.

cells, indicated by Sox9 staining per crypt, as expected given the overall reduction in crypt length and proliferative cells. There was no significant difference in the percentage of apoptotic cells per crypt (11% in both groups)(Figure 6.3; B). Taken together, these results indicate that the reduction in crypt length brought on by anti-IL-22 therapy in mice exposed to *Hh* and AOM is driven by a halving of the proliferative cell compartment within treated crypts, as opposed to a decrease in progenitor cell number or an increase in apoptotic activity.

6.2.2 Genomic analysis

Having shown that anti-IL-22 therapy results in a reduction in intestinal polyp burden and dysplasia, accompanied by a reduction in the proliferative cell

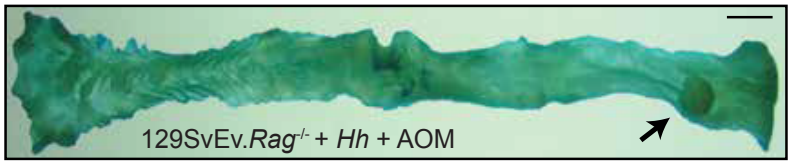


Figure 6.4: Representative image of an *Hh*+AOM colonic polyp. Colon stained with methylene blue for ease of polyp identification. Arrow indicates polyp, scale bar represents 5 mm.

compartment, I next sought to investigate whether this difference in intestinal phenotype was accompanied by a shift in mutation acquisition. I therefore extracted DNA (Section 2.4.1.1) from colonic polyps and adjacent normal tissue collected from each treatment group at the experimental end point (Figure 6.4). Sample quality control and Whole Exome Sequencing (WES), and pre-processing, variant calling and filtering of provided sequencing files were carried out as described previously by the Genomics Core at the WCHG and Ryan Schenck, respectively (Section 2.4.3.2; Section 2.4.3.2-6). I carried out all subsequent variant analysis using R and various publicly available resources.

6.2.2.1 Mutation burden analysis of anti-IL-22-treated normal and polyp tissue

I first sought to characterise the non-synonymous mutation burden within normal tissue and polyps, in order to characterise *Hh*+AOM-driven mutagenesis and determine whether anti-IL-22 treatment protected against it. There was a low number of non-synonymous mutations in normal samples from all three experimental groups, with fewer than 20 mutations per sample detected with the exception of one anti-IL-22 treated sample that contained 112 non-synonymous mutations (Figure 6.5). Surprisingly, this was paralleled in polyps, where the majority (13/17) of samples analysed contained fewer than 40 non-synonymous mutations (Figure 6.6). Of the remaining 4 polyps that harboured more than 40 mutations, 3 were *Hh*+AOM samples (71, 355 and 733 mutations respectively, comprising 50% of *Hh*+AOM polyps analysed) and 1 was anti-IL-22 treated (247 mutations, comprising 11% of *Hh*+AOM+anti-IL-22 polyps analysed). Interestingly, filtering total non-synonymous mutations using a set

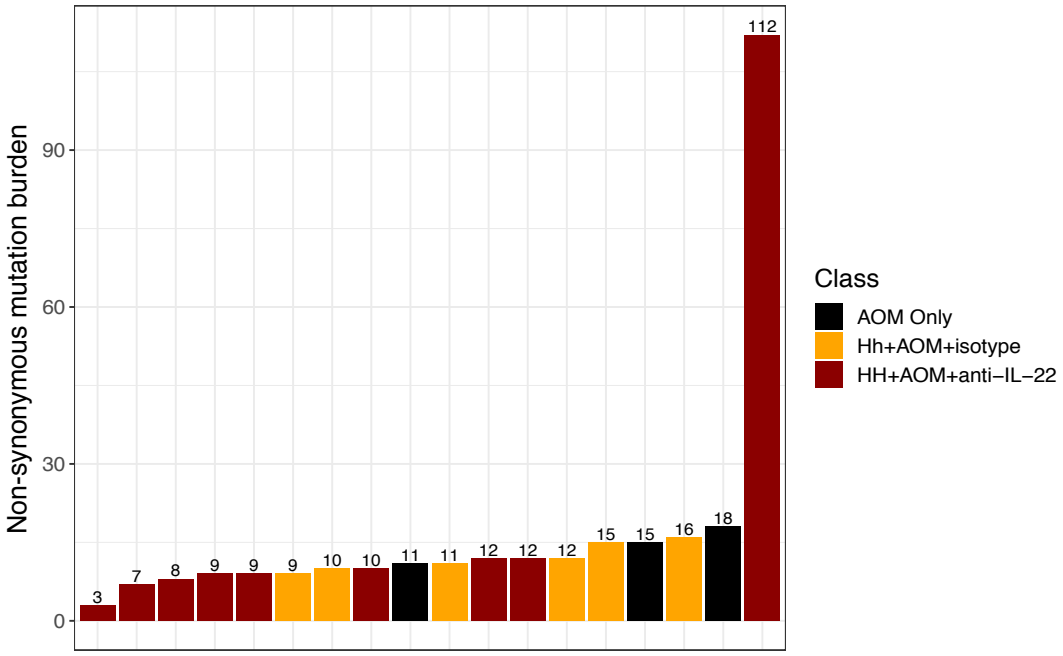


Figure 6.5: Non-synonymous mutation burden in adjacent normal colon samples from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Samples represented as individual bars, coloured according to experimental group.

of 113 putative CRC driver genes (described previously, Section 2.4.3.7) revealed that the three polyps exhibiting the highest non-synonymous mutation burdens all harboured multiple driver gene mutations (Figure 6.7). Taken together then, these findings indicate somewhat surprisingly that despite an expansion in the intestinal proliferative cell compartment in 129SvEv.*Rag*$^{-/-}$ mice exposed to *Hh*+AOM, this putative increase in the number of mutation susceptible cells does not result in a subsequent increase in mutation acquisition, as evidenced by low average mutation burden in normal tissue from all experimental groups (Figure 6.8). Similarly, polyps from all three experimental groups harboured low numbers of non-synonymous mutation burdens, indicating that tumorigenesis in these animals is not driven by increased mutation acquisition. Despite *Hh*+AOM polyps exhibiting an increase in average non-synonymous mutation burden (238) compared to *Hh*+AOM+anti-IL-22 polyps (50), this difference was not statistically significant. Furthermore, only 2/6 (33%) of *Hh*+AOM polyps and

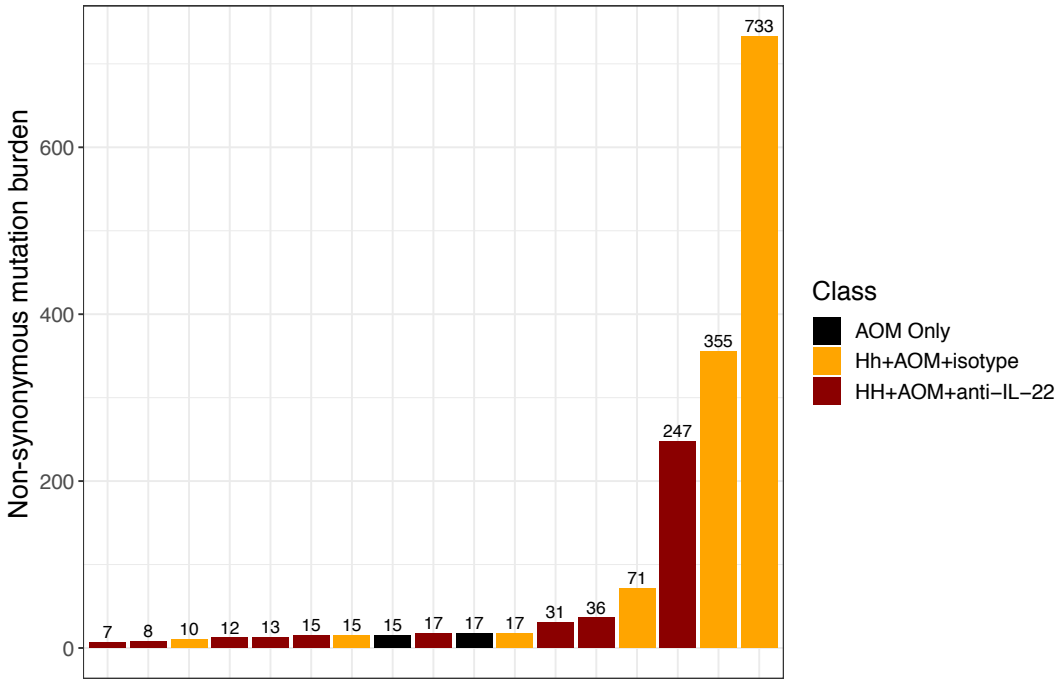


Figure 6.6: Non-synonymous mutation burden in polyp samples from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Samples represented as individual bars, coloured according to experimental group.

1/9 (11%) of *Hh*+AOM+anti-IL-22 polyps harboured mutations in CRC driver genes, with these three polyps comprising 100% of those harbouring >100 non-synonymous mutations. These results indicate a low level of mutation acquisition in both non-dysplastic colon tissue and colonic polyps in mice from all three experimental groups. The phenotypic differences observed between *Hh*+AOM and *Hh*+AOM+anti-IL-22 colon may therefore not be attributable to shifts in general or driver gene-specific mutation acquisition. Given the observation of areas of nuclear β-catenin staining indicative of activated Wnt signalling within colonic regions of mice from both experimental groups, this finding may indicate that, in this model, Wnt activation is not driven by somatic mutation acquisition in key Wnt pathway components such as *Apc* or *Ctnnb1* but instead may be a consequence of epigenetic or transcriptomic Wnt activation, possibilities explored in more detail later in this chapter.

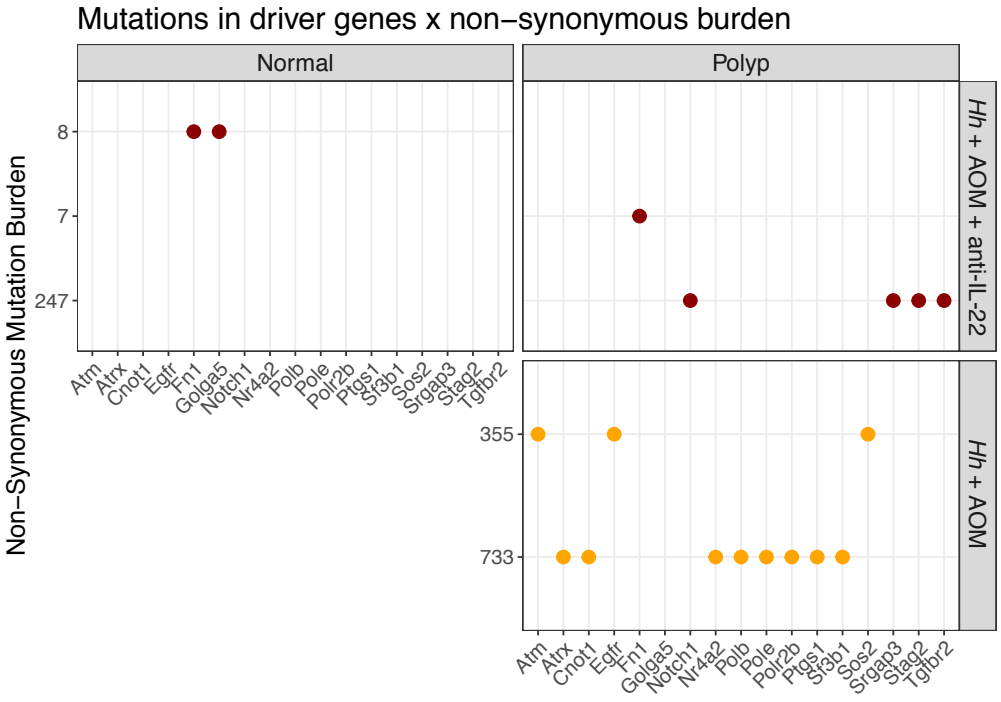


Figure 6.7: Characterisation of CRC driver gene mutations within normal and polyp samples from *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Sample type indicated by top panels, experimental group indicated by right panels. Vertical lines represent individual driver genes as indicated on the x-axis; horizontal lines represent samples, with non-sysnonymous mutation burden per sample indicated on the y-axis. Mutations represented by dots, coloured according to sample type.

6.2.2.2 Characterisation of mutation types within anti-IL-22-treated normal and polyp tissue

I next sought to investigate whether anti-IL-22 therapy results in a shift in the types of mutations present within both normal and polyp tissue from each treatment group. I therefore characterised the mutations within each sample according to impact, revealing significant variety in the percentage contribution of each mutation type within both normal (Figure 6.9) and polyp (Figure 6.10) samples, possibly due to the majority of samples harbouring a low number of mutations overall. While there was a small degree of variability in the average percentage contribution of each mutation type between treatment groups, there was very little difference between normal and polyp samples within groups (Figure 6.11). In normal tissue exposed only to AOM, the majority of mutations are made up by intron (~50%) variants, followed by missense (~25%)

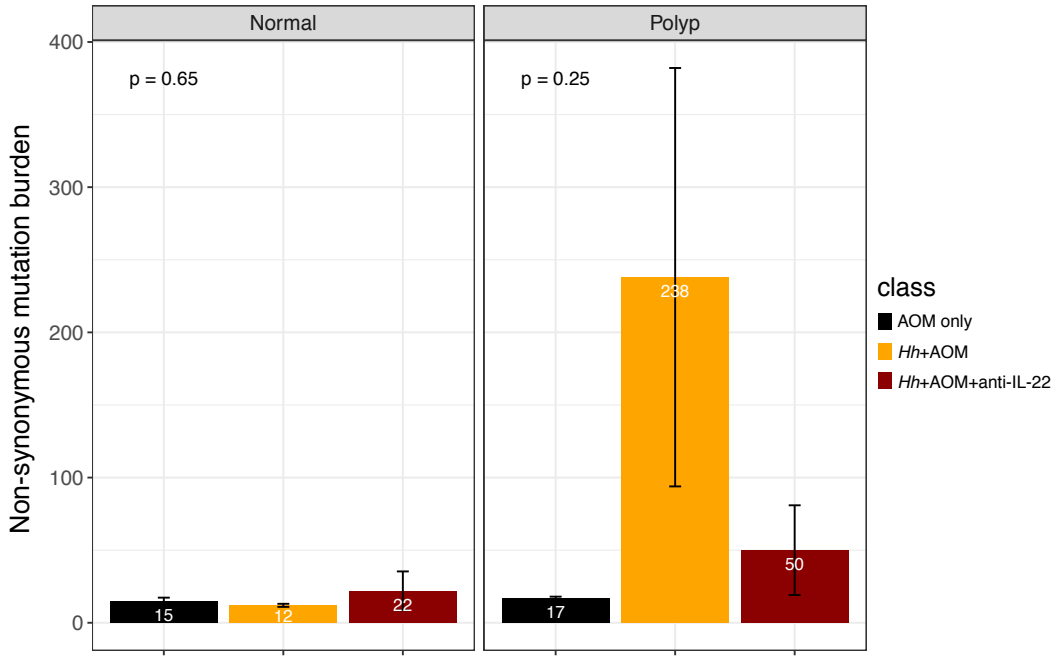


Figure 6.8: Average non-synonymous mutation burden in normal adjacent samples and polyps from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Error bars represent SEM; p-value as indicated, calculated by one-way ANOVA.

and non-genic (~12%) variants, with the remainder comprised of insertions and deletions, splice variants and synonymous mutations. Normal tissue from *Hh*+AOM mice also exhibited UTR variants, with an increase in intron variants and reduction in missense mutations and 'indels'. *Hh*+AOM+anti-IL-22 normal tissue harboured an increase in the average contribution of synonymous mutations. Polyps from animals exposed only to AOM exhibited the same mutation type spectrum as normal tissue from these animals, with 6 mutation types contributing all mutations detected. Polyps from animals exposed to *Hh*+AOM showed a greater variety of mutation types such as UTR, non-coding transcript and nonsense variants. Compared to polyps from the AOM-only experimental group, *Hh*+AOM polyps exhibit a relative doubling of the contribution of synonymous and missense variants, alongside a decrease in non-genic variants. Anti-IL-22 treatment results in a slight decrease in the average contribution of synonymous and missense variants, although the overall average mutation type contribution

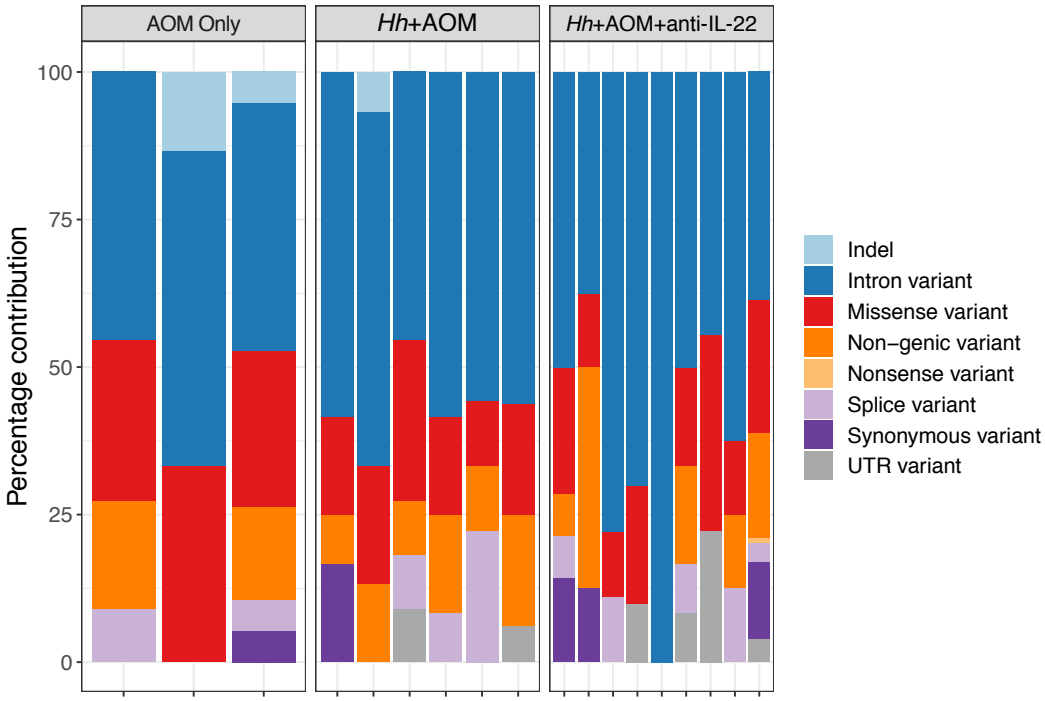


Figure 6.9: Percentage contribution of mutation types in adjacent normal samples from *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Top panels indicate experimental groups; bars represent individual samples, coloured according to the percentage contribution of each indicated mutation type.

remained largely unchanged between *Hh*+AOM+anti-IL-22 and *Hh*+AOM polyps. Taken together, these findings indicate that despite generally harbouring very few mutations overall, there are potentially damaging nonsense and missense mutations occurring within both normal and polyp tissue across all treatment groups. There is no evidence, however, of a therapy-driven shift in mutation types in *Hh*+AOM+anti-IL-22 versus *Hh*+AOM samples.

6.2.2.3 Mutation signature analysis of anti-IL-22-treated normal and polyp tissue

In order to further characterise mutation acquisition, I next sought to determine whether anti-IL-22 therapy produced a detectable shift in the mutational processes that are active within normal and polyp samples. Mutation signatures within normal (Figure 6.12) and polyp (Figure 6.13) samples were extracted using base substitutions and information on the sequence context of each mutation

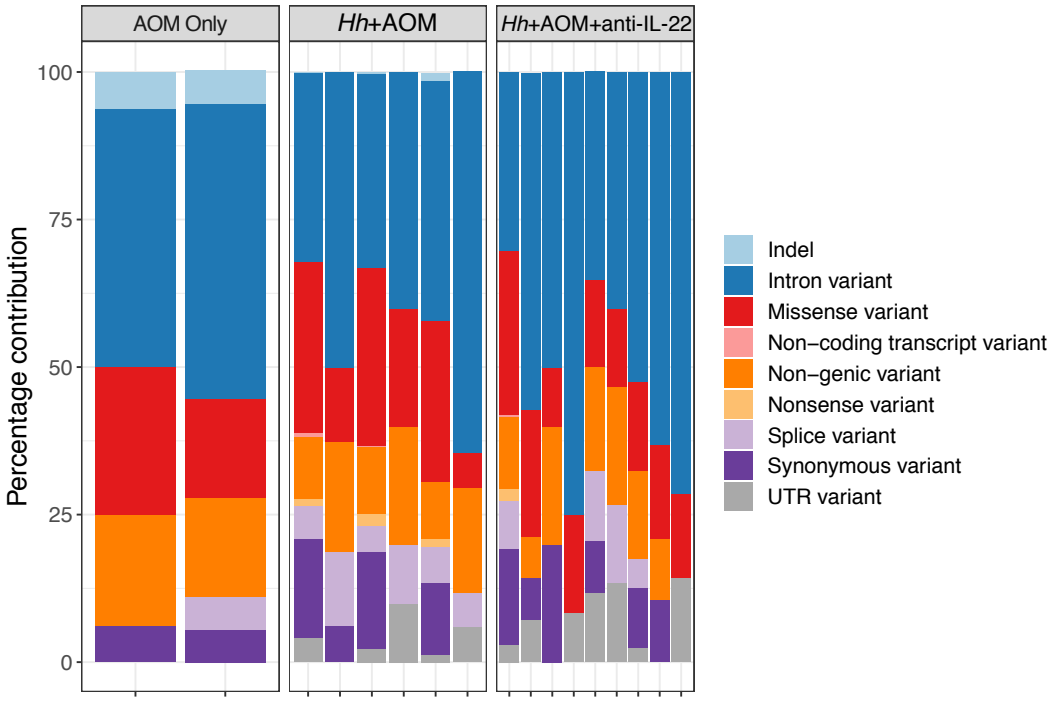


Figure 6.10: Percentage contribution of mutation types in polyps from AOM only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Top panels indicate experimental groups; bars represent individual polyps, coloured according to the percentage contribution of each indicated mutation type.

(Section 2.4.3.8). Applying this approach to the analysis of normal tissue samples from each treatment group revealed the presence of 8 distinct, validated mutational signatures that have been reported previously and documented in the COSMIC database (Figure 6.14). As would be expected, a clock-like mutational process (SBS1) was present in normal samples from all treatment groups. Interestingly, the alkylation-associated SBS11 previously suggested to be associated with AOM mutagenesis was diluted in AOM only normal samples but present those from both *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. This may indicate that in AOM-only normal samples, the inflammation-free period of time between AOM exposure and sample collection facilitates the processing of the bulky adducts caused by AOM by MMR and NER associated processes, thereby diluting the AOM-associated alkylation signal. In normal samples from mice harbouring *Hh*-driven inflammation, however, the alkylation signal remains

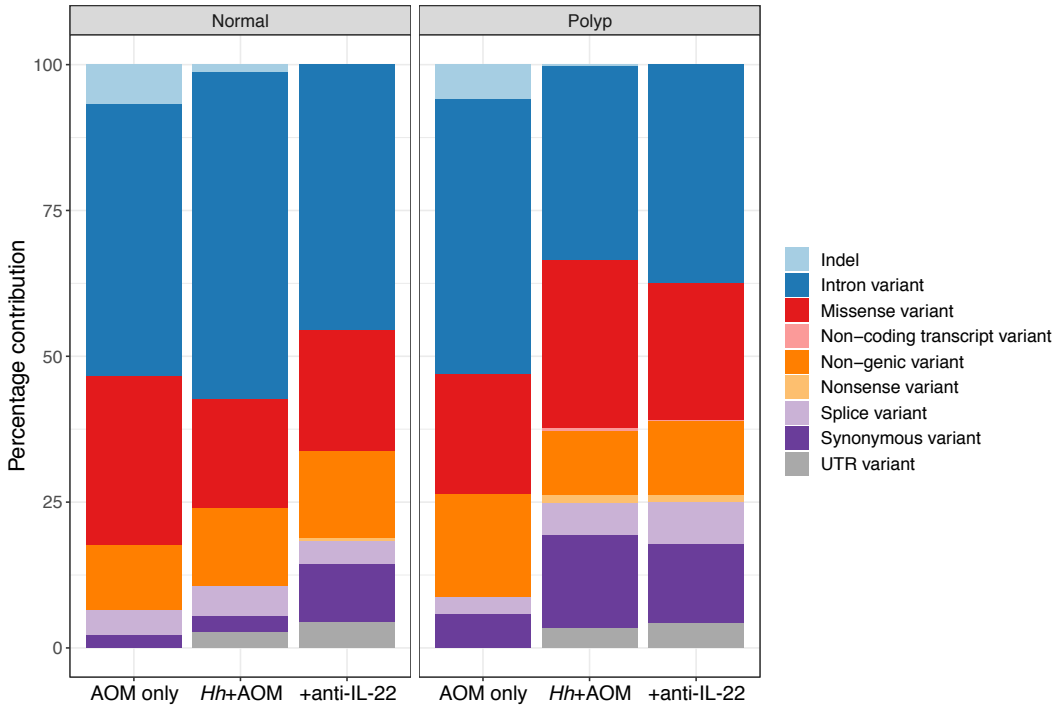


Figure 6.11: Average mutation type contribution in adjacent normal and polyp samples from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Sample type indicated by top panels; bars represent experimental groups, coloured according to the average percentage contribution of each indicated mutation type to all detected mutations per group.

detectable alongside agreater variety of signatures associated with genomic instability, perhaps indicating the greater variety of genotoxic stress caused by inflammation in these mice. Importantly, there was no difference in the mutational processes detected in normal samples from *Hh*+AOM+anti-IL-22 mice compared to those in *Hh*+AOM mice. Taken together, these data indicate that despite harbouring low numbers of mutations, in mice exposed to both *Hh* and AOM the majority of these are contributed by MMR, NER and alkylation-associated mutational processes that dilute clock-like mutation accumulation, thus indicating a degree of inflammation-associated genomic instability in these mice.

The majority of mutations within polyps from mice exposed only to AOM were contributed by two mutation signatures also detected within normal samples in this group, the clock-like SBS1 and NER-associated SBS15. The remainder were contributed by SBS7, associated with the processing of bulky DNA adducts

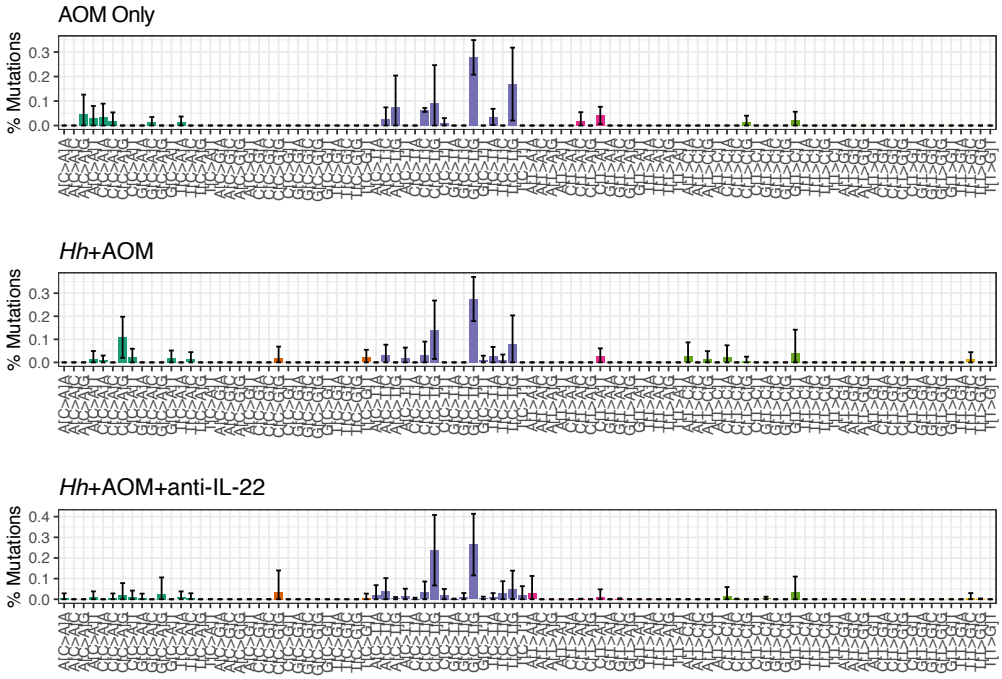


Figure 6.12: 96-class signature analysis of mutations in normal samples from AOM-only, *Hh*+AOM and *Hh*+AOM+anti-IL-22 mice. Represented as the average contribution of each substitution to all mutations detected within adjacent normal samples per indicated experimental group. Error bars represent SEM.

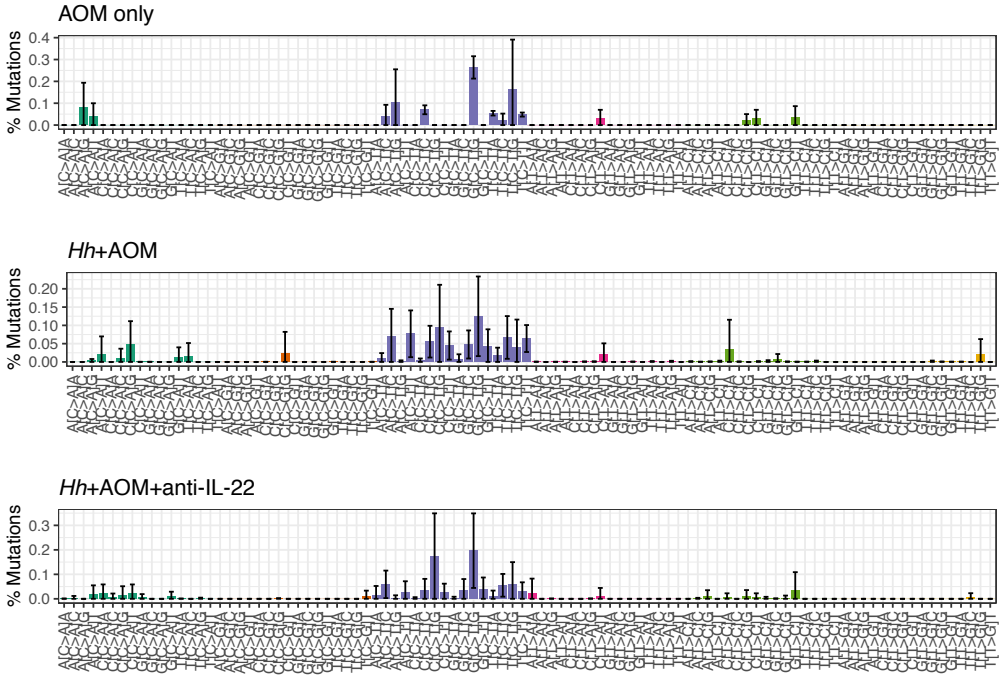


Figure 6.13: 96-class signature analysis of mutations in polyps. Represented as the average contribution of each substitution to all mutations detected within adjacent normal samples per indicated experimental group. Error bars represent SEM.

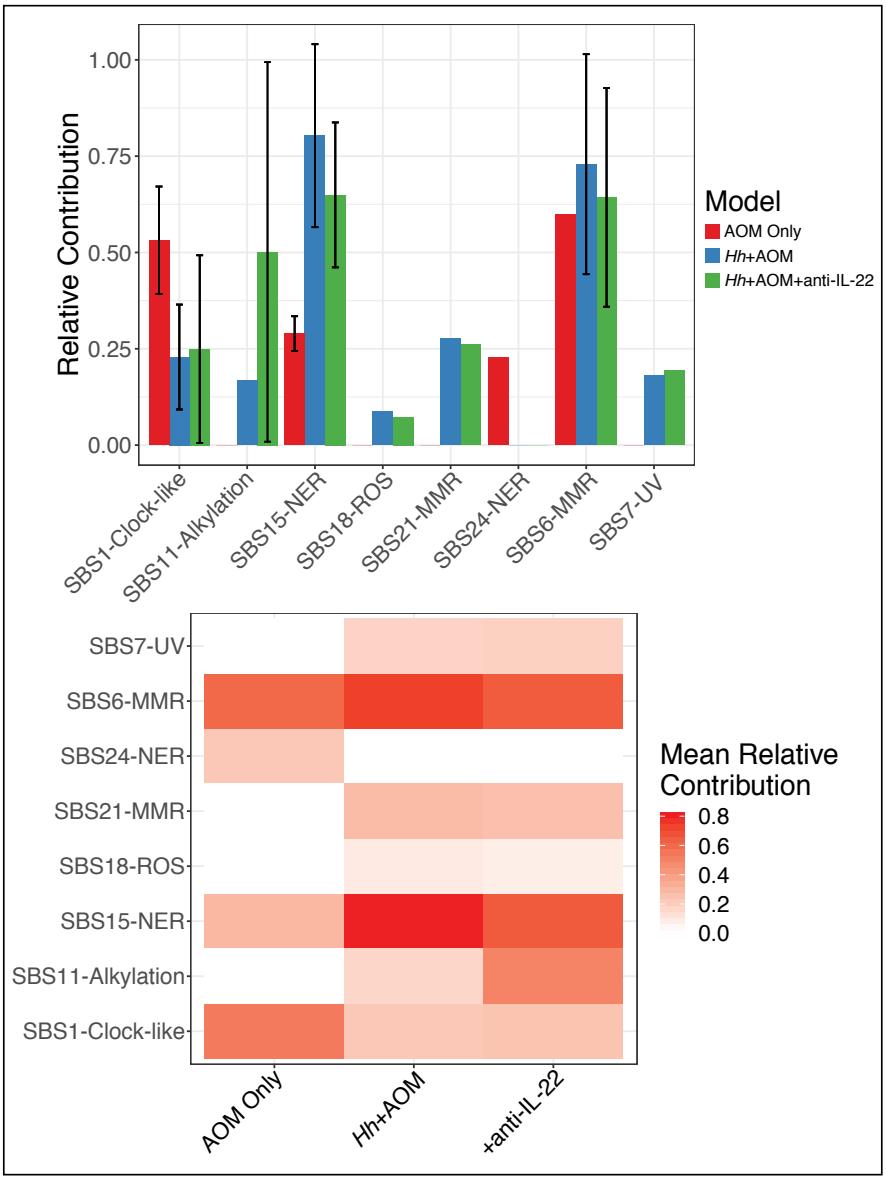


Figure 6.14: Relative contribution of mutational signatures in normal samples from each experimental group. The average contribution of each single base substitution (SBS) signature to all mutations detected in adjacent normal samples per experimental group, represented as a bar plot (top) coloured according to experimental group and a heatmap (bottom) coloured according to the mean relative contribution of each signature.

usually caused by exposure to UV radiation (Figure 6.15). This indicates that in polyps from these mice, mutations are largely contributed by the same processes active in non-dysplastic tissue. This is also true for polyps from mice exposed to *Hh*+AOM, with the majority of processes active in normal samples also detected in polyps. The AOM-associated SBS11 contributed the majority of mutations in polyps from *Hh*+AOM mice, with the remainder generated by a variety of processes relating to MMR (SBS6, SBS7, SBS14), NER (SBS8, SBS15) and the damage caused by reactive oxygen species (SBS18). Interestingly, treatment with anti-IL-22 caused a reduction in the proportion of mutations contributed by the AOM-associated signature compared to untreated polyps, with the remainder of mutations in *Hh*+AOM+anti-IL-22 polyps contributed by a greater variety of DNA repair-associated processes than in untreated *Hh*+AOM polyps. This may indicate that therapy facilitates a greater degree of repair of the DNA damage caused by AOM, such that mutations are generated more from these repair processes themselves. There were also unique signatures detected within untreated *Hh*+AOM polyps that were not present in *Hh*+AOM+anti-IL-22 polyps, and vice versa. SBS8 was unique to *Hh*+AOM polyps, while SBS4 and 21 were unique to *Hh*+AOM+anti-IL-22 polyps. This indicates that despite the processes that contribute the majority of mutations in normal tissue remaining active, further diversification of mutation generation occurs within polyps. Treatment with anti-IL-22 thus seems to result in a small but detectable shift in polyp mutational processes that is distinct from those active in polyps treated with isotype-mAb, along with a dilution of the AOM-associated mutational signature. Taken together, however, these findings indicate that the majority of mutational processes active in normal samples from all treatment groups also contribute the majority of mutations within polyps, with anti-IL-22 treatment having a minimal effect on the mutational processes active in *Hh*+AOM polyps.

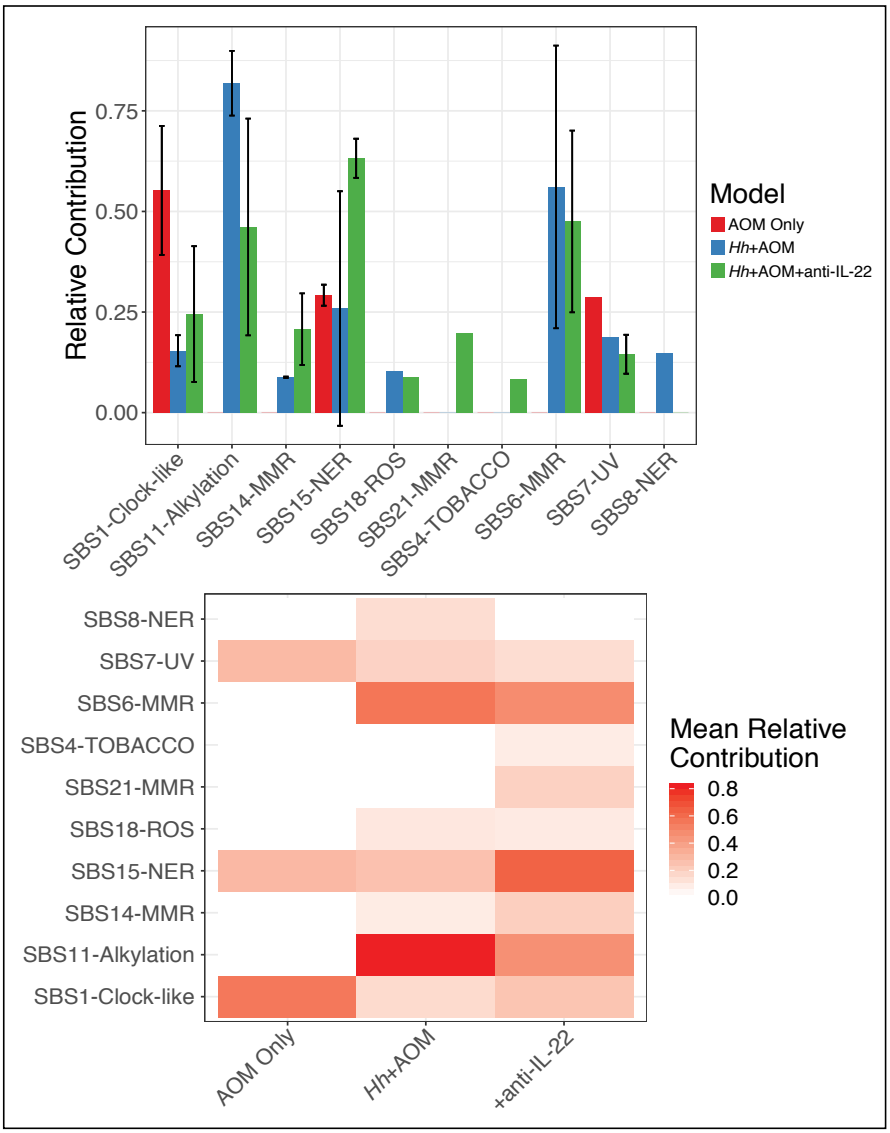


Figure 6.15: Relative contribution of mutational signatures in polyps from each experimental group. The average contribution of each single base substitution (SBS) signature to all mutations detected in polyps from experimental group, represented as a bar plot (top) coloured according to experimental group and a heatmap (bottom) coloured according to the mean relative contribution of each signature.

6.2.3 Transcriptomic analysis

Given that the findings presented thus far indicate that the phenotypic differences observed between isotype-treated and anti-IL-22-treated mice are not attributable to shifts in general or driver gene-specific mutation acquisition, I next sought to characterise the transcriptome within samples from each experimental group. RNA was extracted (Section 2.5.1) from normal and polyp samples for each treatment group, along with representative 'steady state' normal samples from untreated 129SvEv.*Rag*$^{-/-}$ mice, and submitted for RNA sequencing by the High-Throughput Genomics Group at the Wellcome Centre for Human Genetics, Oxford. Gene mapping and generation of read counts was performed by Matt Jackson, a bioinformatician within the Powrie Group, using the R packages bowtie and featureCounts (Section 2.5.2). I carried out all subsequent analysis using the DESeq2 package in R and various publicly available resources cited hereafter.

In order to stratify the various differential expression analyses I carried out, I characterised 'normal' samples from untreated mice as 'steady state', from isotype-treated mice as 'colitis' and from anti-IL-22 treated mice as 'treated colitis'. Similarly, polyp samples from isotype-treated and anti-IL-22 treated mice were deemed 'colitic polyp' and 'treated polyp', respectively. Given that the purpose of this analysis was to address my aim of exploring the impact of anti-IL-22 therapy on the transcriptomic profile of treated colitis and polyp samples, I first characterised disease pathogenesis in this model, represented by the transcriptomic progression from steady state to colitis and then colitis to colitic polyp. I then go on to contrast the transcriptomic profile of untreated samples with that of anti-IL-22-treated samples, in order to determine the effect of therapy upon colitis and colitic polyps, respectively.

6.2.3.1 Characterisation of disease progression

I first sought to characterise disease pathogenesis in this model by performing differential expression (DE) analysis on expression data from steady state, colitis and colitic polyp samples. Principal component analysis of normalized expression

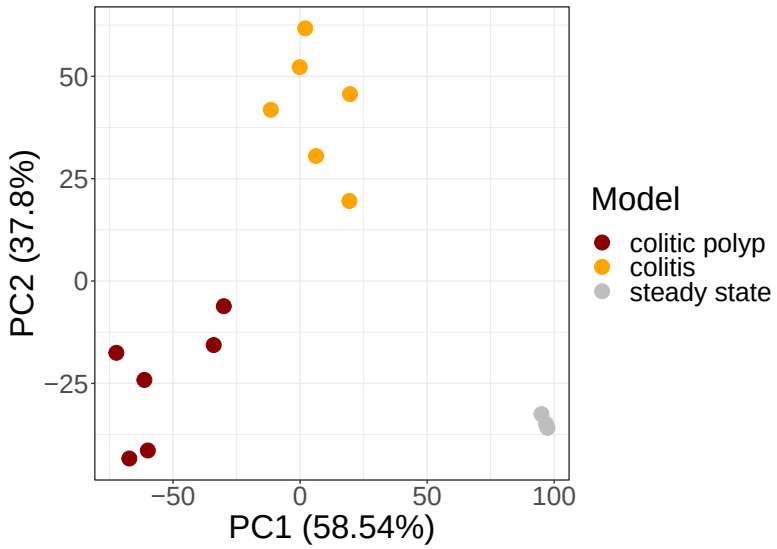


Figure 6.16: Principal component analysis plot showing clustering of steady state, colitis and colitic polyp samples.

data showed that these samples clustered strongly according to disease state (Figure 6.16), indicating that the progression from steady state to colitis and colitic polyp is driven by a strong degree of transcriptomic variance between each state. Importantly, in contrast to the results of the genomic analyses presented earlier, there appears to be a distinct transcriptomic difference between colitis and polyp samples.

Steady state to colitis

Visualisation of the genes differentially expressed in colitis compared to steady state samples demonstrates that the progression to colitis in this model is associated with a significant degree of transcriptomic change (Figure 6.17), with 1469 genes significantly upregulated (padj <0.05, log2foldchange >1.5) and 1983 significantly downregulated (padj <0.05, log2foldchange <-1.5).

In order to investigate the biological relevance of the genes differentially expressed in the progression to colitis, I next carried out an enrichment analysis to identify statistically enriched pathways in each gene list. To do this, I inputted the lists of genes significantly up or down regulated in the progression to colitis ("UP in colitis" and "DOWN in colitis", respectively) to the Metascape

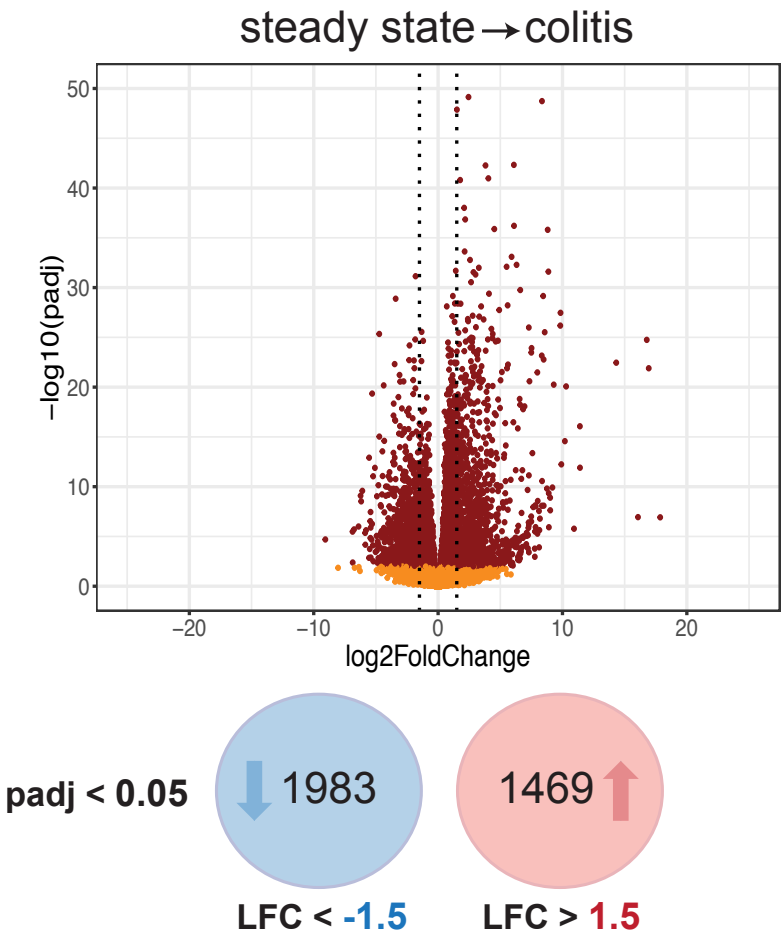


Figure 6.17: Volcano plot of genes differentially expressed in colitis samples compared to steady state samples, showing multiple testing-corrected p-values (padj) and log2FoldChange in expression. Points in yellow represent genes with padj>0.05; dotted lines signify log2FC 1.5 and -1.5; circles represent the number of genes significantly (padj<0.05) up and down-regulated (LFC > 1.5 and < -1.5, respectively).

enrichment analysis tool. In order to identify the most significant biological processes impacted in the progression to colitis in this model, I focused on the top 100 most statistically significant enriched pathways across both input lists (Figure 6.18). Of these, the majority of pathways enriched in upregulated genes correspond to processes involved in the activation of immune and inflammatory responses, such as cytokine production and the bacterial pathogen response, along with cell death and apoptotic processes. Pathways enriched in downregulated genes represented a variety of biological processes, with several involved in the regulation of cell and signalling pathways such as extracellular organisation and transmembrane signalling pathways.

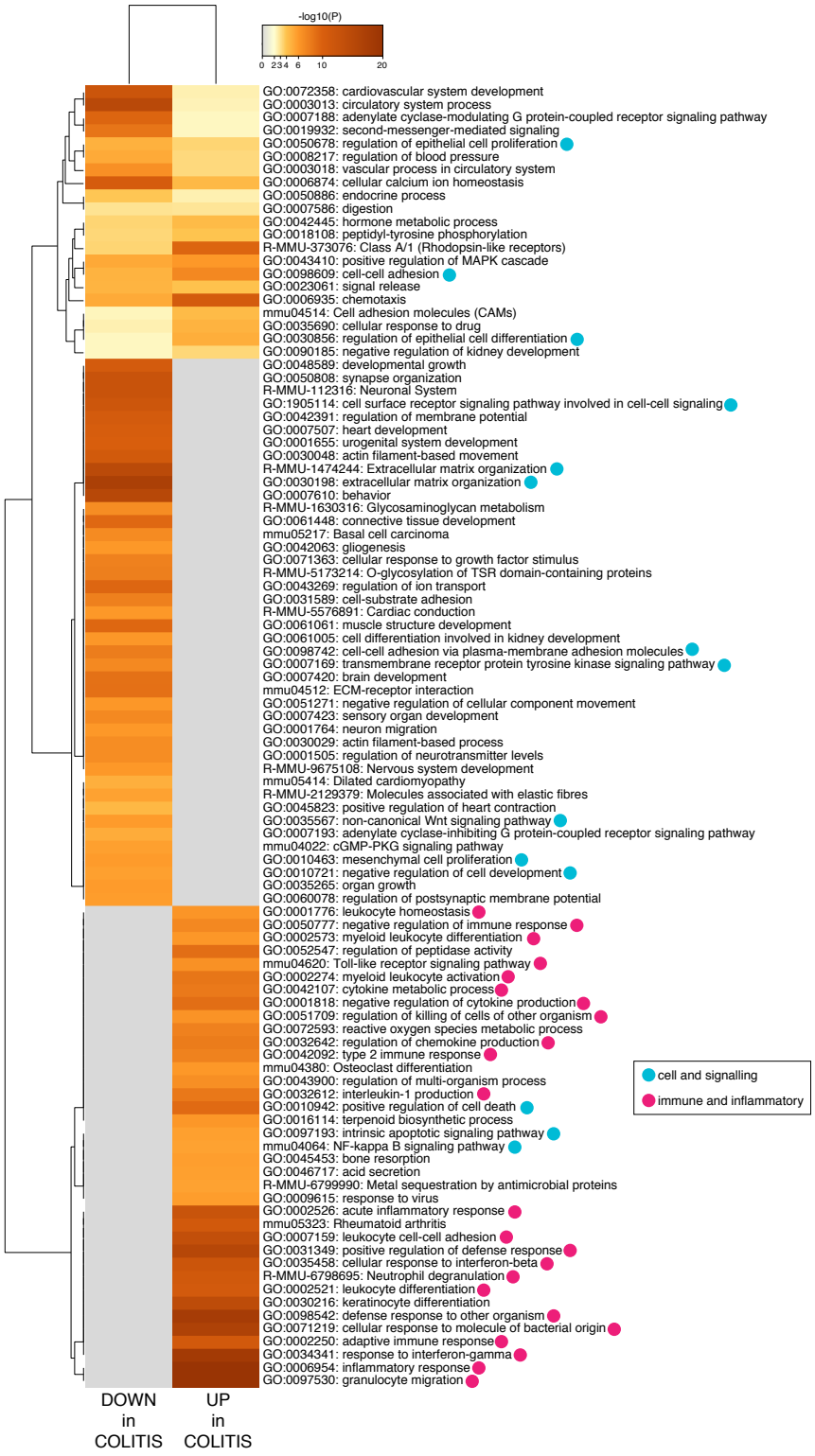


Figure 6.18: Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in colitis samples (compared to steady state), coloured according to significance. Dots indicate biological relevance of the indicated pathways.

Interestingly, pathways corresponding to the regulation of cellular proliferation and differentiation were enriched in genes both up and downregulated in colitis. Taken together, these findings indicate that the progression from steady state to colitis in this model is associated with a significant program of transcriptomic change involving the activation of immune and inflammation-associated pathways and differential regulation of key intestinal processes, such as proliferation and differentiation. The specific impact of these pathways upon intestinal cell populations and homeostatic signalling pathways is explored later in this section.

Colitis to colitic polyp

I next sought to investigate the transcriptomic changes enacted in the development of colitic polyps. I therefore performed DE analysis on expression data from colitis and colitic polyp samples. Visualisation of the differentially expressed genes indicates that the progression from colitis to colitic polyp in this model is associated with a stronger degree of upregulation (Figure 6.19), with 1517 genes significantly upregulated (padj<0.05, LFC > 1.5) compared to 735 significantly downregulated (padj <0.05, LFC < -1.5).

Enrichment analysis further revealed that the majority of the top 100 most significantly enriched processes were represented by genes upregulated in polyps (Figure 6.20). The majority of these processes correspond to pathways involved in immune and inflammatory responses, with a significant number additionally implicated in morphological and developmental pathways such as tissue remodelling. Contrastingly, the processes significantly enriched in downregulated genes largely correspond to metabolic pathways. Similar to colitis samples, pathways associated with cellular processes such as proliferation and differentiation were enriched in genes both up and downregulated in polyps. Together, these findings indicate that the progression from colitis to colitic polyp in this model is associated with a stronger degree of upregulated gene expression, with genes involved in morphogenetic and developmental pathways upregulated and those involved in metabolic processes downregulated.

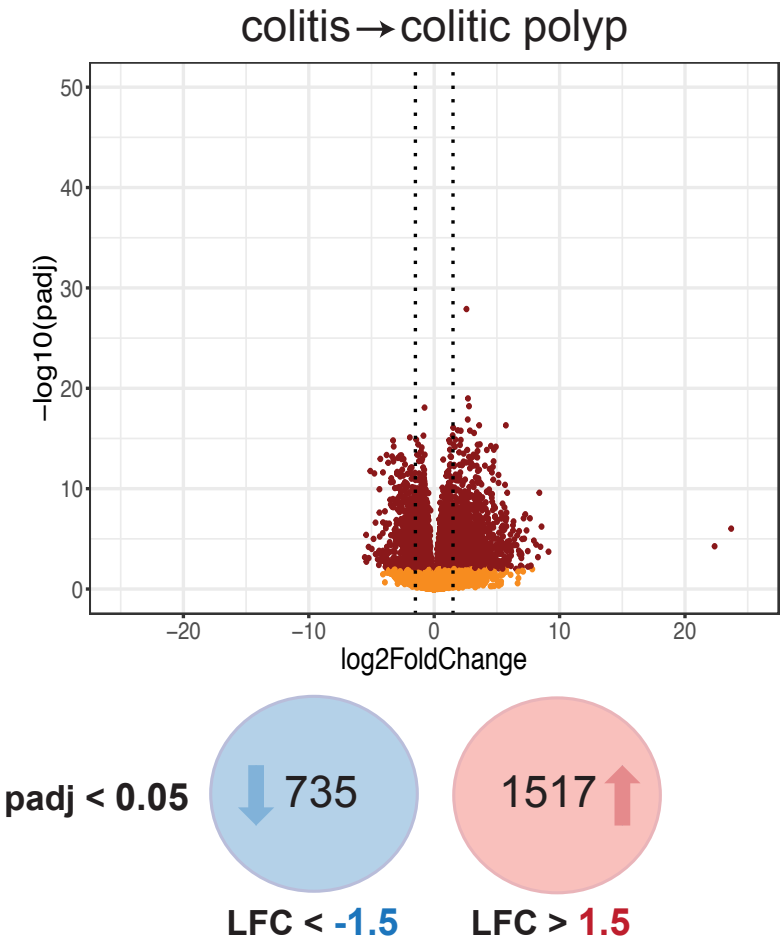


Figure 6.19: Volcano plot of genes differentially expressed in colitic polyps compared to colitis samples, showing multiple testing-corrected p-values (padj) and log2FoldChange in expression. Points in yellow represent genes with padj>0.05; dotted lines signify log2FC 1.5 and -1.5; circles represent the number of genes significantly (padj<0.05) up and down-regulated (LFC > 1.5 and < -1.5, respectively).

Gene set enrichment analysis

The results presented to here have demonstrated that disease progression in this model is largely underpinned by an upregulation of genes involved in activation of the immune and inflammatory response. This is in line with previous work demonstrating that disease pathogenesis in this model is sustained through the production of IL-22 by group 3 innate lymphoid cells (ILC3s), with subsequent activation of epithelial Stat3 signalling resulting in increased intestinal proliferation [57]. However, I have also shown that genes involved in key

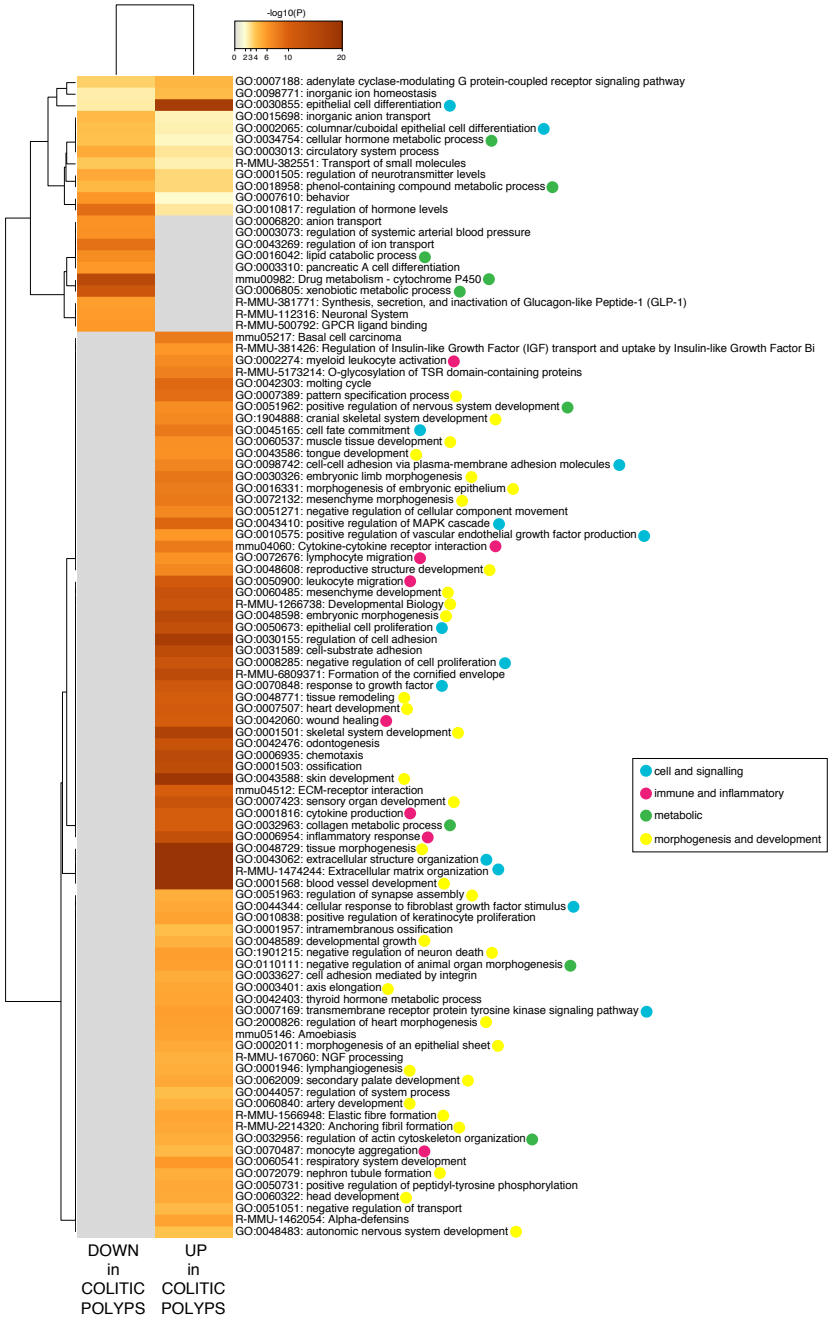


Figure 6.20: Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in colitic polyps (compared to colitis), coloured according to significance. Dots indicate biological relevance of the indicated pathways.

biological processes that govern intestinal homeostasis, such as proliferation and differentiation, are differentially expressed to different degrees within colitis and colitic polyp samples. Since the work of Kirchberger *et al* in 2013, numerous studies have implicated IL-22 signalling in a wide variety of these intestinal processes, including proliferation [178], differentiation [48] and the DNA damage response [49]. I therefore next sought to determine whether any of these IL-22-associated processes are specifically implicated in the progression from steady state to colitis and colitic polyp in this model. I thus carried out gene set enrichment analysis (GSEA) utilising established gene sets from the Gene Ontology (GO) and Molecular Signatures(MSigDB) databases. I selected 25 gene sets from these databases that correspond to key intestinal pathways and processes linked to IL-22 signalling, such as Wnt, Notch, DNA repair and apoptotic pathways (Appendix Table 1).

I first characterised the processes enriched in the progression from steady state to colitis in this model. GSEA of genes differentially expressed in colitis showed marked enrichment of gene signatures associated with IL-17 production and the IL-22-associated JAK-STAT signalling pathway (Figure 6.21). This confirms the strong immune-driven aspect of disease pathogenesis in this model. Interestingly, genes differentially expressed in colitis also exhibited significant negative enrichment (NES -1.73, padj 0.014) of canonical Wnt signalling and positive enrichment of gene signatures associated with the DNA damage response. Rashi *et al* characterised the first of these signatures, "Rashi response to ionizing radiation 6", as late responding genes activated in ATM-deficient, but not wild type, tissues, while the second signature enriched here, "Rashi response to ionizing radiation 2", correspond to late ATM-dependent genes induced by ionizing radiation treatment. This indicates that the progression from steady state to colitis in this model is also underpinned by enrichment of 'non-canonical' or Atm-independent DNA damage response genes in signature 6 that may activate the late-responding DNA repair genes of signature 2.

Applying the same enrichment analysis to genes differentially expressed in colitic polyps compared to colitis revealed the processes enriched in the

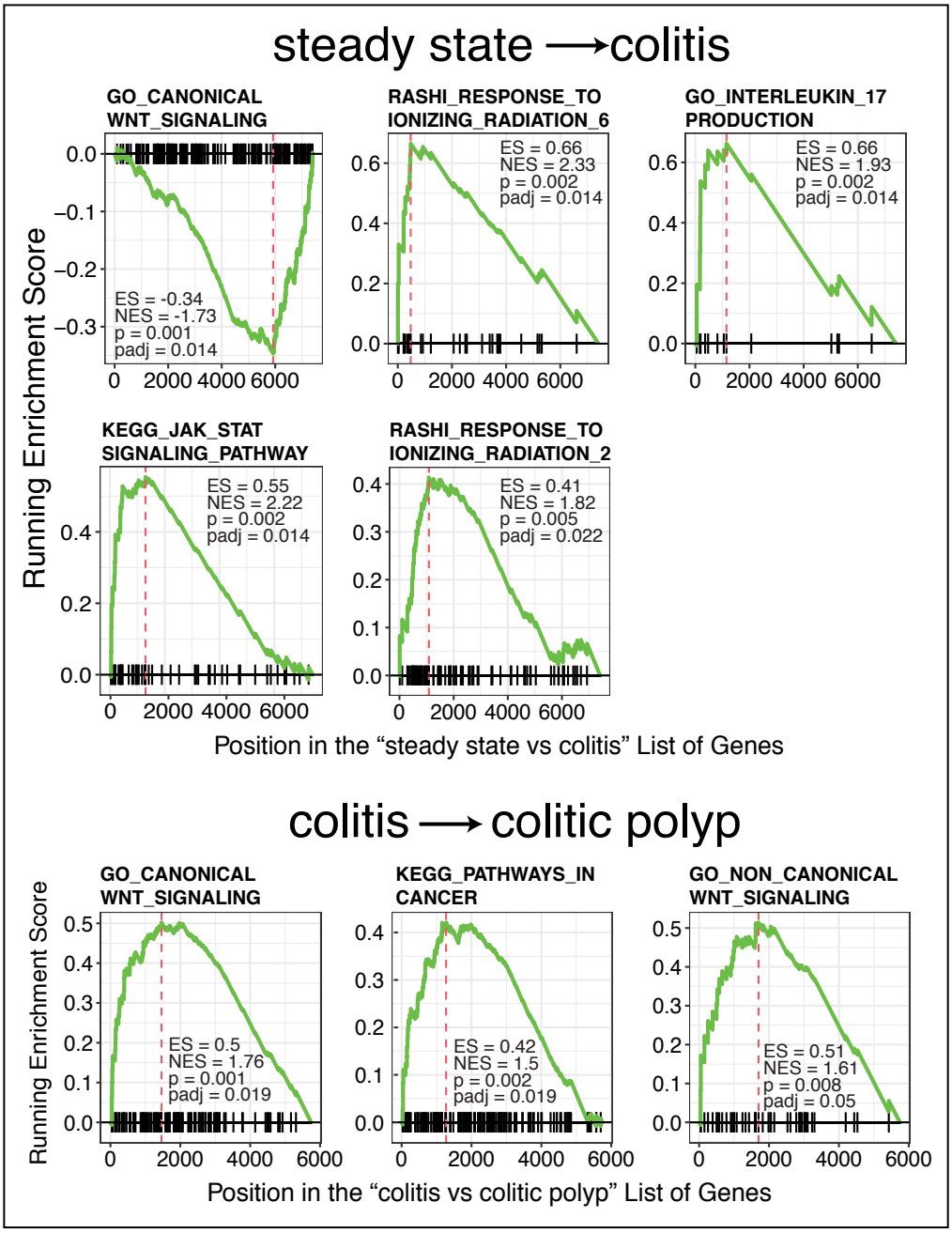


Figure 6.21: Gene set enrichment analysis plots showing enriched processes within colitis (compared to steady state) and colitic polyp (compared to colitis) samples. Enrichment score (ES), Normalised enrichment score (NES), nominal and multiple testing-adjusted p-values as calculated by the GSEA function (clusterProfiler package, R).

progression from colitis to polyp in this model. Strikingly, the negative enrichment of canonical Wnt signalling in colitic samples was reversed in polyps, alongside strong positive enrichment of non-canonical Wnt signalling. Polyp formation is also underpinned by strong positive enrichment (NES 1.5, padj 0.019) of cancer-associated pathways.

Together, these findings demonstrate that the progression from steady state to colitis in this model involves enrichment of inflammatory processes and putative 'non-canonical' DNA repair genes and downregulation of genes involved in Wnt signalling. Further progression from colitis to to colitic polyp, however, involves strong positive enrichment of canonical and non-canonical Wnt signalling and cancer-associated pathways.

6.2.3.2 Intestinal cell populations and signalling pathways

The GSEA results confirmed differential enrichment of key intestinal pathways underpins colitis and polyp formation in this model. I next sought to further characterise disease progression within the specific context of these processes, by analysing markers of key intestinal cell populations and signalling pathways in colitis and colitic polyp samples.

Il-22

I first confirmed a statistically significant increase in *IL-22* and *Stat3* expression in colitis and colitic polyps, as compared to steady state samples (Figure 6.22). Interestingly, this increase in IL-22 signalling does not appear to be driven by increased expression of IL-22 receptors (*Il22ra1* and *Il22ra2*). This finding is in line with those of Kirchberger *et al*, who demonstrated that increased IL-22 signalling in this model is instead driven by increased production of IL-22 itself by infection-driven invasion of ILC3s.

Novel work has since suggested that the environmental sensor aryl hydrocarbon receptor (Ahr) mediates IL-22 production in ILC3 cells, and that an Ahr-IL-22-Atm axis is involved in regulating DDR within the intestine [49]. In

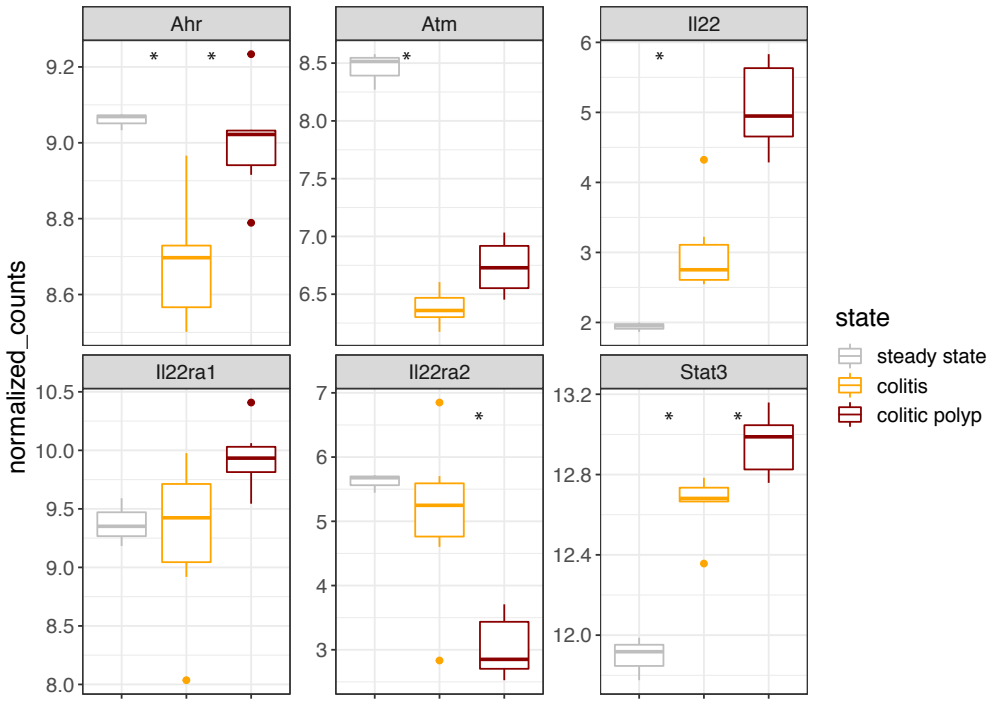


Figure 6.22: Normalized expression of genes involved in IL-22 signalling within steady state, colitis and colitic polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

contrast to this work, in which decreased *Ahr* expression was associated with decreased *IL-22* and *Atm* expression, colitis samples from this model exhibit significantly decreased expression of *Ahr* and *Atm* compared to steady state tissue. This may be a result of the difference in pathogenic insult, such that dietary toxins result in an upregulation of *Ahr* but bacterial infection results in reduced *Ahr* expression in order to facilitate an antimicrobial response. These findings thus suggest that *Ahr* expression is not the primary determinant of *Il-22* production within the colon, as colitis in this model is associated with increased expression of *IL-22* and *Stat3* and reduced expression of *Ahr* and *Atm*. Interestingly, there is a restoration of *Ahr* expression to steady state levels in polyps, while *IL-22* and *Stat3* expression are further increased. This may indicate that the anti-microbial reduction in *Ahr* expression in colitis is overtaken by an increased, *Stat3*-associated proliferative drive in polyps. Importantly, the significant reduction in Atm expression within colitis and polyp samples provides

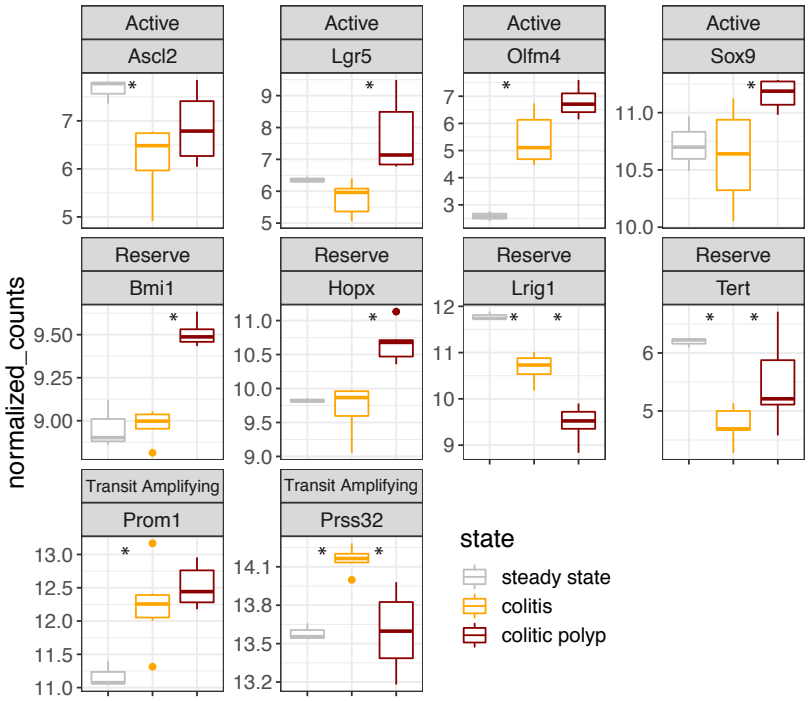


Figure 6.23: Normalized expression of intestinal stem and proliferating cell markers within steady state, colitis and colitic polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

a mechanistic basis for my finding that Atm-independent DNA damage response pathways are enriched in these samples.

Stem and proliferative cells

I next sought to characterise stem, proliferating and differentiating cells within colitis and colitic polyp samples, given that IL-22 has been proposed to play a role in proliferation and differentiation within the intestine through Wnt and Notch signalling [178]. In colitis samples, the expression of active and reserve stem cell markers was either unchanged (*Lgr5*, *Sox9*, *Bmi1*, *Hopx*) or reduced (*Ascl2*, *Lrig1*, *Tert*), with the exception of *Olfm4*, which was increased in colitic samples (Figure 6.23). In comparison to colitic samples, however, the expression of these markers in polyps was either unchanged (*Ascl2*, *Olfm4*) or increased (*Lgr5*, *Sox9*, *Bmi1*, *Hopx*).

The expression of proliferative cell markers *Prom1* and *Prss32* was increased in colitis samples and unchanged (*Prom1*) or reduced (*Prss32*) in polyps. These results indicate that colitis is associated with a reduction in stem cells and an

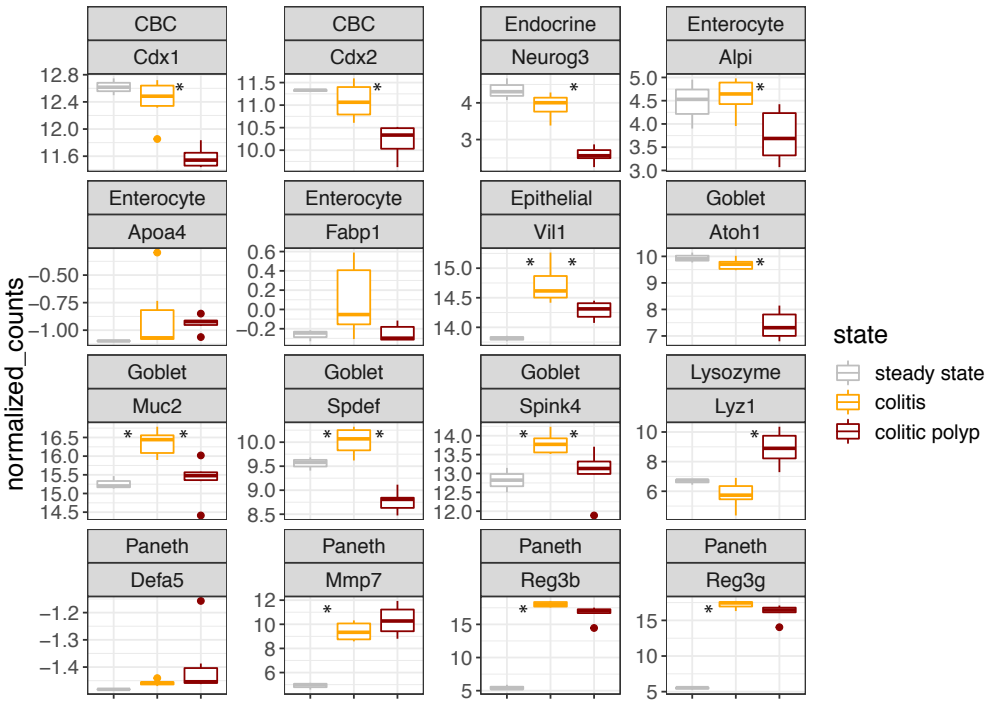


Figure 6.24: Normalized expression of markers of intestinal epithelial cell types within steady state, colitis and colitic polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

increase in proliferative cells that is partially abrogated or diluted upon colitic polyp formation. Characterisation of markers of specific intestinal cell types revealed that colitis results in increased expression of anti-microbial goblet (*Muc2*, *Spdef*, *Spink4*) and Paneth (*Mmp7*, *Reg3b*, *Reg3g*) cell markers (Figure 6.24), as expected given the inflammatory response underpinning colitis in this model. In polyps, this increase in goblet cells is abrogated while increased expression of Paneth cell markers is maintained. Furthermore, the expression of markers of crypt base columnar, endocrine and enterocyte cells was unchanged in colitis but reduced in polyps. Taken together, these results indicate that colitis in this model is associated with a reduction in stem cells and an increase in proliferative, goblet and Paneth cells. Relative to colitis samples, however, polyps harbour an increase in stem cells and a reduction in differentiated cells.

Wnt signalling

I have thus established that colitis in this model is driven by a putative reduction

in stem cells, alongside an increase in proliferative cells, that is partially abrogated in polyps. I next characterised the Wnt and Notch signalling pathways, which govern intestinal proliferation and differentiation, in order to determine the mechanistic basis for these observations and the GSEA results presented earlier. Strikingly, colitis samples harboured either unchanged or significantly reduced expression of both non-canonical and canonical Wnt ligands, along with all frizzled and Lrp receptors analysed (Figure 6.25). Confirming earlier GSEA results, this resulted in a statistically significant reduction in expression of the Wnt target genes *Axin2* and *Ephb3* in colitis compared to steady state samples.

In comparison to colitis samples, colitic polyps exhibited statistically significant increases in expression of the non-canonical Wnt ligands *Wnt11*, *Wnt16*, *Wnt5a*, *Wnt6* and *Wnt7b*. Similarly, expression of the frizzled receptors *Fzd1*, *Fzd10*, *Fzd3*, *Fzd6*, *Fzd7* and *Lrp5* was significantly upregulated in polyps compared to colitis samples. This resulted in a statistically significant increase in expression of all three Wnt target genes analysed in polyps compared to colitis samples. These findings indicate that the negative enrichment of Wnt signalling in colitis samples demonstrated earlier is driven by reduced expression of Wnt ligands and receptors. In polyps, however, increased expression of non-canonical Wnt ligands and Fzd receptors results in increased Wnt signalling relative to colitis samples. This supports results presented earlier in this section that demonstrated an enrichment of non-canonical Wnt signalling in polyps.

Notch signalling

Given the well documented interplay between Wnt and Notch signalling in maintaining gradients of proliferation and differentiation within the intestine, I next characterised the expression of key Notch signalling components within steady state, colitis and colitic polyp samples. The expression of all Notch signalling ligands was either unchanged (*Jag1*, *Jag2*) or reduced (*Dll1*) in colitis compared to steady state samples (Figure 6.26). Strikingly, all three Notch receptors exhibited statistically significant reductions in expression in colitis

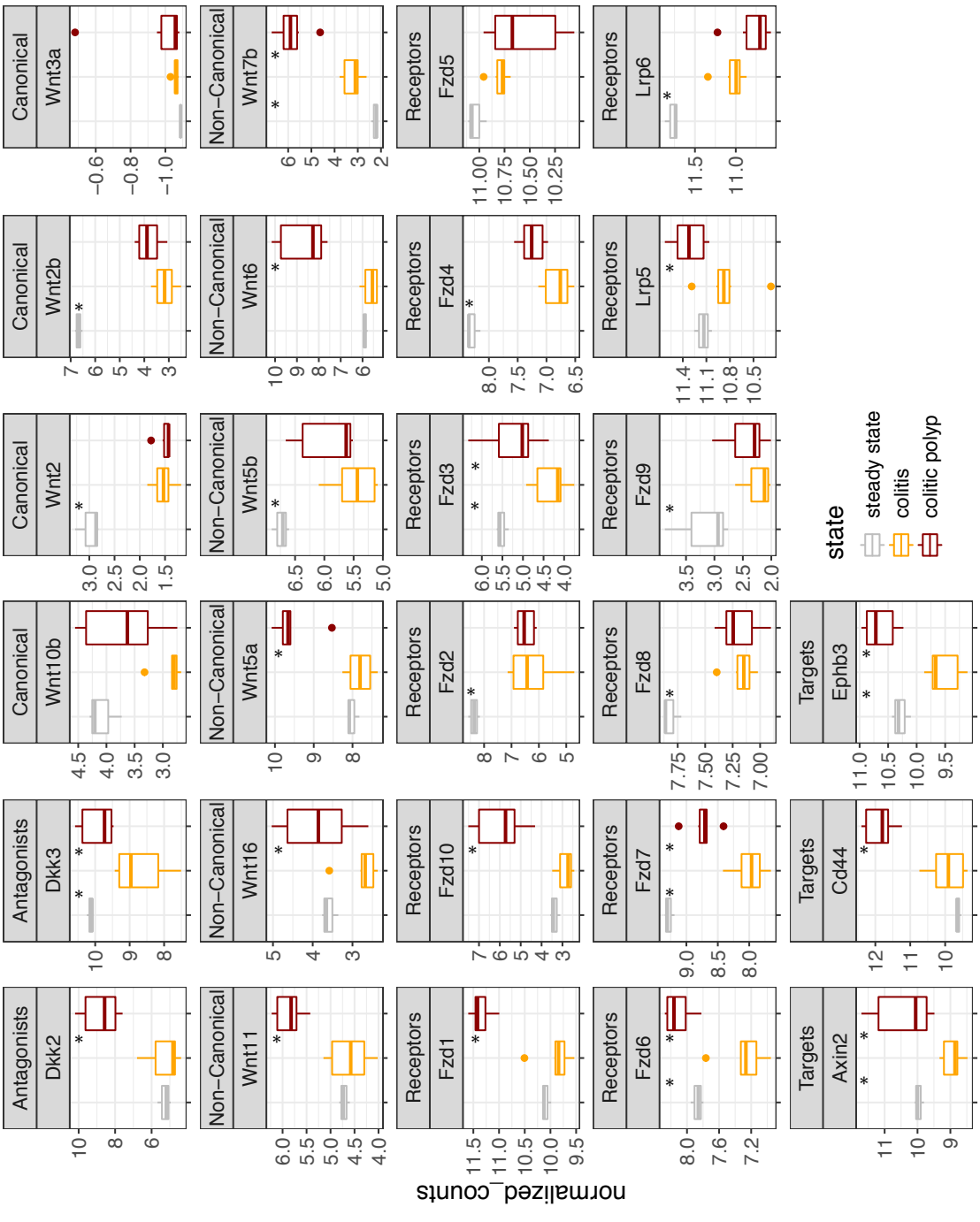


Figure 6.25: Normalized expression of genes involved in Wnt signalling within steady state, colitis and colitic polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

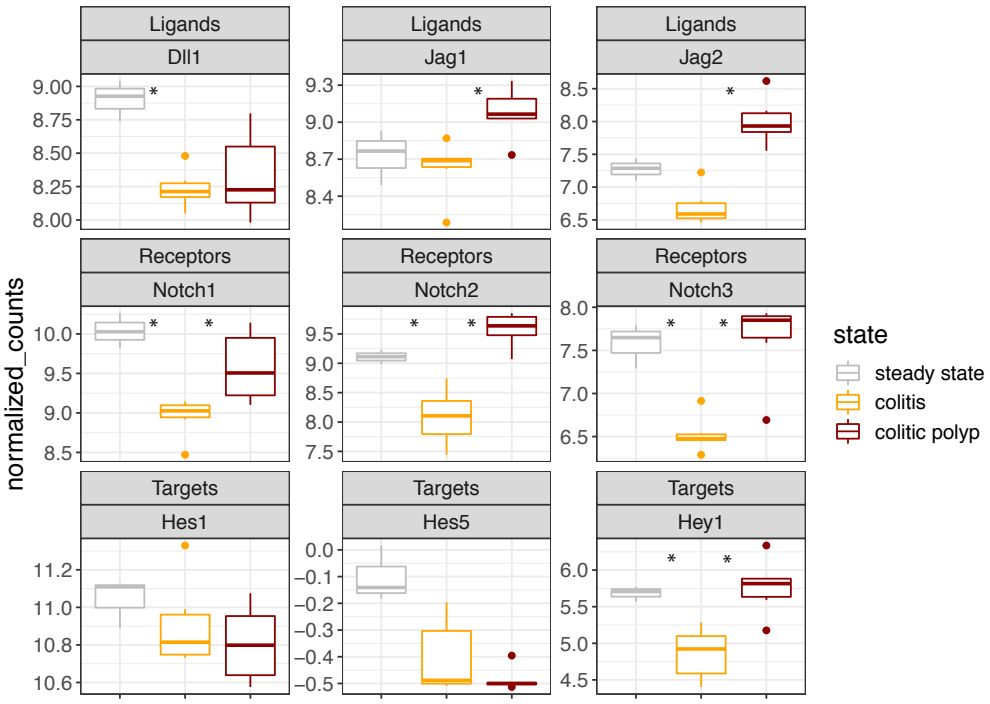


Figure 6.26: Normalized expression of genes involved in Notch signalling within steady state, colitis and colitic polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

samples, resulting in a reduction in the expression of all three Notch signalling targets analysed.

In polyps, however, the expression of Notch components was largely rescued compared to colitis samples, with statistically significant increases in expression observed for the ligands *Jag1* and *Jag2*, all three Notch receptors and the Notch target gene *Hey1*. These results indicate that colitis in this model is associated with a reduction in Notch signalling that is driven by reduced expression of Notch receptors, while the progression from colitis to polyp results in increased Notch activity as a result of upregulation of Notch ligands and receptors.

6.2.3.3 The effect of anti-IL-22 therapy on disease progression

Having characterised disease pathogenesis in colitis and colitic polyp samples, I next sought to determine the impact of anti-iL-22 therapy on disease progression. Principal component analysis of normalized expression data from steady state,

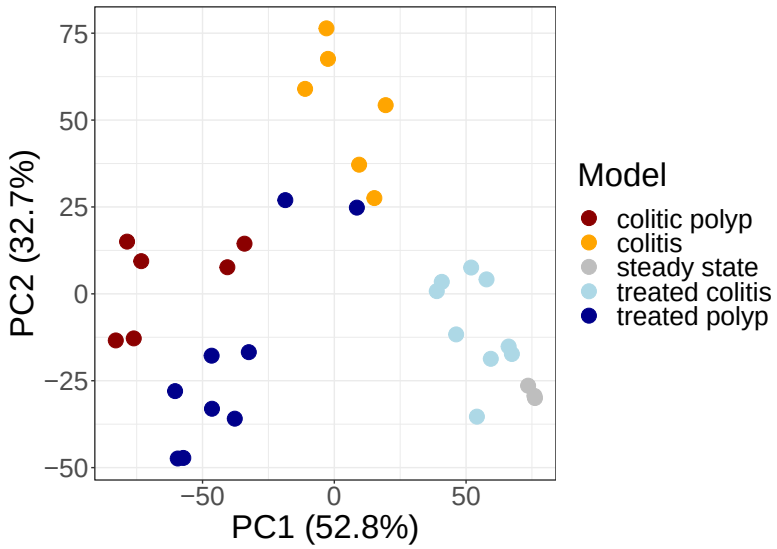


Figure 6.27: Principal component analysis plot showing clustering of steady state, colitis, colitic polyp, treated colitis and treated polyp samples.

colitis, colitic polyp, anti-IL-22-treated colitis (treated colitis from here) and anti-IL-22-treated colitic polyp (treated polyp from here) samples showed that, as before, untreated samples clustered strongly according to disease state (Figure 6.27).

Treated colitis samples, however, appear to cluster more closely to steady state samples than untreated colitis samples. Contrastingly, treated polyp samples largely clustered with untreated polyps. This indicates that anti-IL-22 treatment in colitis samples results in a smaller degree of transcriptomic variance relative to steady state samples. Importantly, this effect was not observed in treated polyps.

The effect of anti-IL-22 therapy on gene expression

Visualisation of the genes differentially expressed in treated colitis compared to untreated colitis samples demonstrates that anti-IL-22 therapy in colitis is associated with a stronger degree of upregulation (Figure 6.28), with 1713 genes significantly upregulated (padj<0.05, LFC > 1.5) compared to 529 significantly downregulated (padj <0.05, LFC < -1.5), although downregulated genes exhibited a greater range of negative LFC.

Visualisation of the genes differentially expressed in treated polyps compared to untreated polyps demonstrates that anti-IL-22 therapy does not result in a

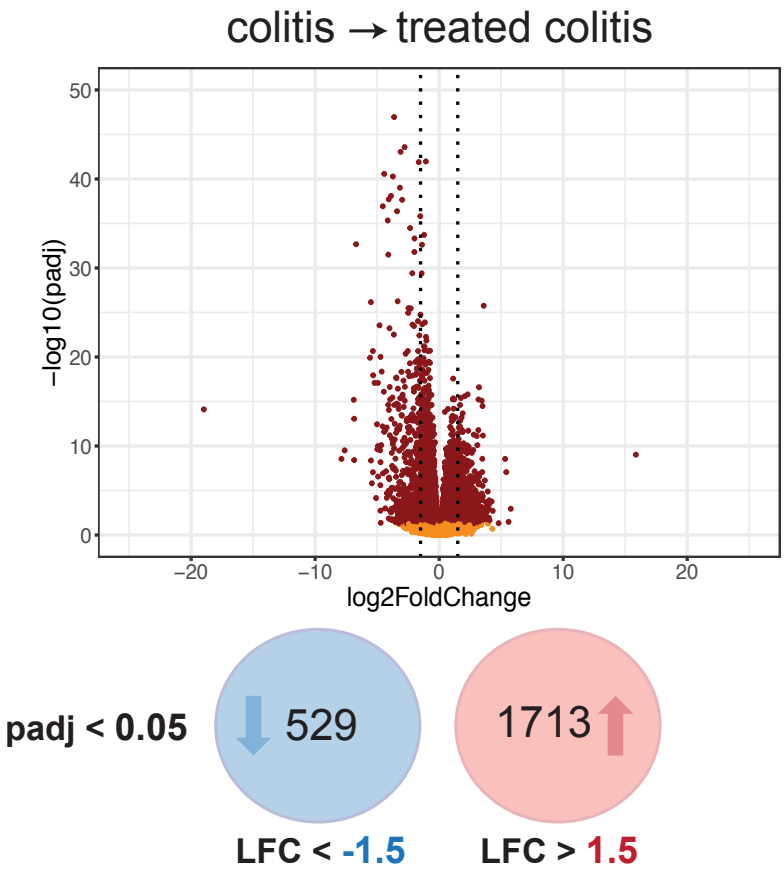


Figure 6.28: Volcano plot of genes differentially expressed in anti-IL-22-treated colitis compared to untreated colitis, showing multiple testing-corrected p-values (padj) and log2FoldChange in expression. Points in yellow represent genes with padj>0.05; dotted lines signify log2FC 1.5 and -1.5; circles represent the number of genes significantly (padj<0.05) up and down-regulated (LFC > 1.5 and < -1.5, respectively).

significant degree of transcriptomic change in polyps (Figure 6.29), with only 212 genes significantly upregulated (padj<0.05, LFC > 1.5) and 250 significantly downregulated (padj <0.05, LFC < -1.5).

Anti-IL-22 therapy thus appears to have a different impact on gene expression in colitis samples compared to polyps. To explore this further, I compared the transcriptome of untreated and treated colitis and polyp samples to steady state samples. Comparison to the same baseline allowed for direct comparison of the differentially expressed genes (DEGs) within each state.

While there was a marked difference between the number of DEGs in untreated colitis compared to treated colitis, this was not the case between treated and untreated polyps (Figure 6.30). Untreated colitis samples harboured a total of

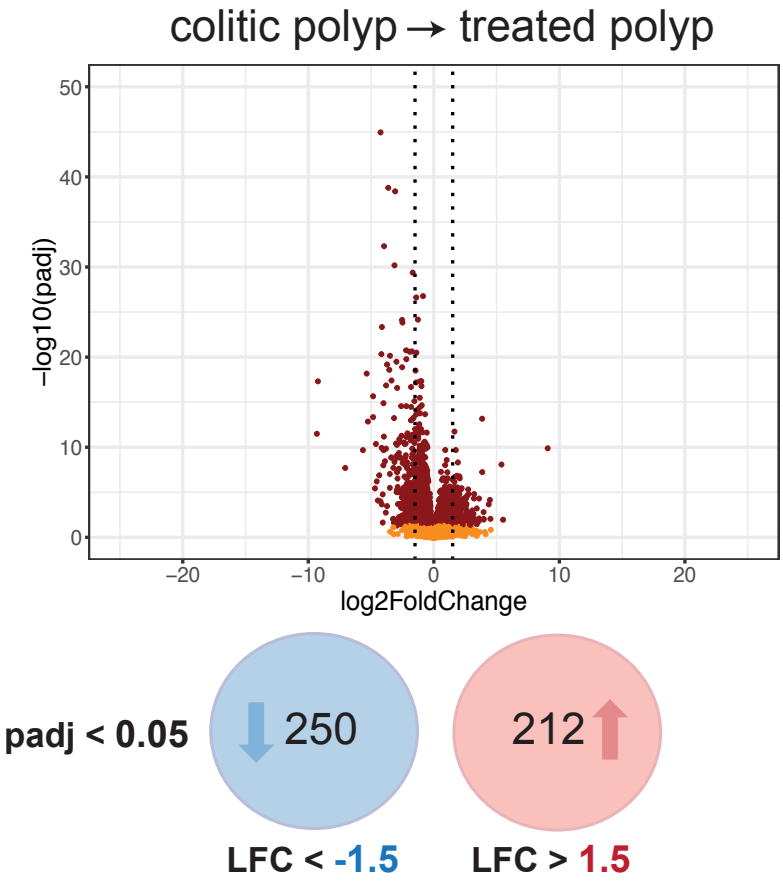


Figure 6.29: Volcano plot of genes differentially expressed in anti-IL-22-treated polyps compared to untreated polyps, showing multiple testing-corrected p-values (padj) and log2FoldChange in expression. Points in yellow represent genes with padj>0.05; dotted lines signify log2FC 1.5 and -1.5; circles represent the number of genes significantly (padj<0.05) up and down-regulated (LFC > 1.5 and < -1.5, respectively).

3452 significantly differentially expressed genes relative to steady state samples, while only 1298 genes were differentially expressed between treated colitis and steady state samples. Contrastingly, untreated and treated polyps harboured similar numbers of differentially expressed genes relative to steady state samples (4936 and 4161, respectively).

In order to determine whether the genes differentially expressed in untreated colitis and polyp samples remained so in treated samples, I next carried out an overlap analysis of the significantly differentially expressed gene lists per comparison (Figure 6.30; C). This revealed that, of 1983 genes significantly downregulated in colitis samples, only 339 (17%) remained so in treated colitis,

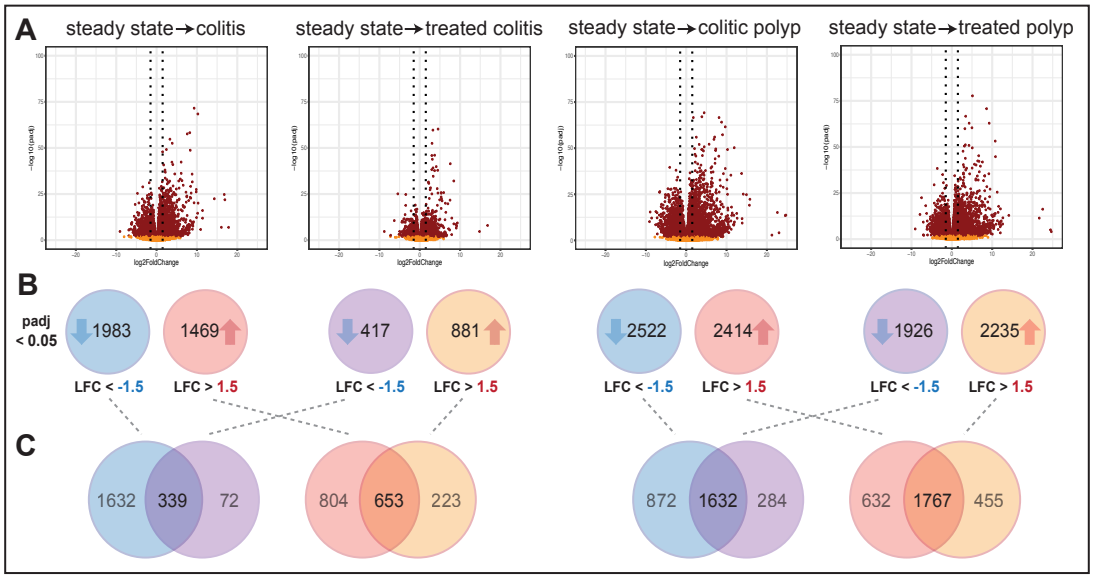


Figure 6.30: Characterisation of the transcriptomic effect of anti-IL-22 therapy in colitis and colitic polyp samples. A) Volcano plot of differentially expressed genes in untreated and anti-IL-22-treated colitis and colitic polyp samples (compared to steady state), with multiple testing-corrected p-values (padj) on the y-axis and log2FoldChange in expression on the x-axis. Points in yellow represent genes with padj>0.05; dotted lines signify log2FC 1.5 and -1.5; B) Coloured circles represent the number of genes significantly (padj<0.05) up and down-regulated (LFC > 1.5 and < -1.5, respectively) in each differential expression analysis. C) Overlap analysis of the significant differentially expressed genes in untreated and anti-IL-22-treated colitis (left) and colitic polyp (right) samples.

and of the 1469 genes significantly upregulated, 653 (44%) remained so following treatment. In polyps, however, of 2522 and 2414 genes significantly down and upregulated in untreated polyps, 1632 (65%) and 1767 (73%) remained so in treated polyps, respectively. These results indicate that anti-IL-22 therapy rescues the majority of genes differentially expressed in the progression to colitis in this model. However, this is not the case in polyps, such that the majority of genes differentially expressed in untreated polyps remain so in treated polyps.

The effect of anti-IL-22 therapy on biological processes

I next sought to characterise the impact of anti-IL-22 therapy on biological processes within treated and untreated samples. Enrichment analysis of genes differentially expressed in colitis samples revealed that the majority of the top 100 most significantly enriched processes were detected in genes upregulated in treated colitis (Figure 6.31). Processes enriched in upregulated genes were

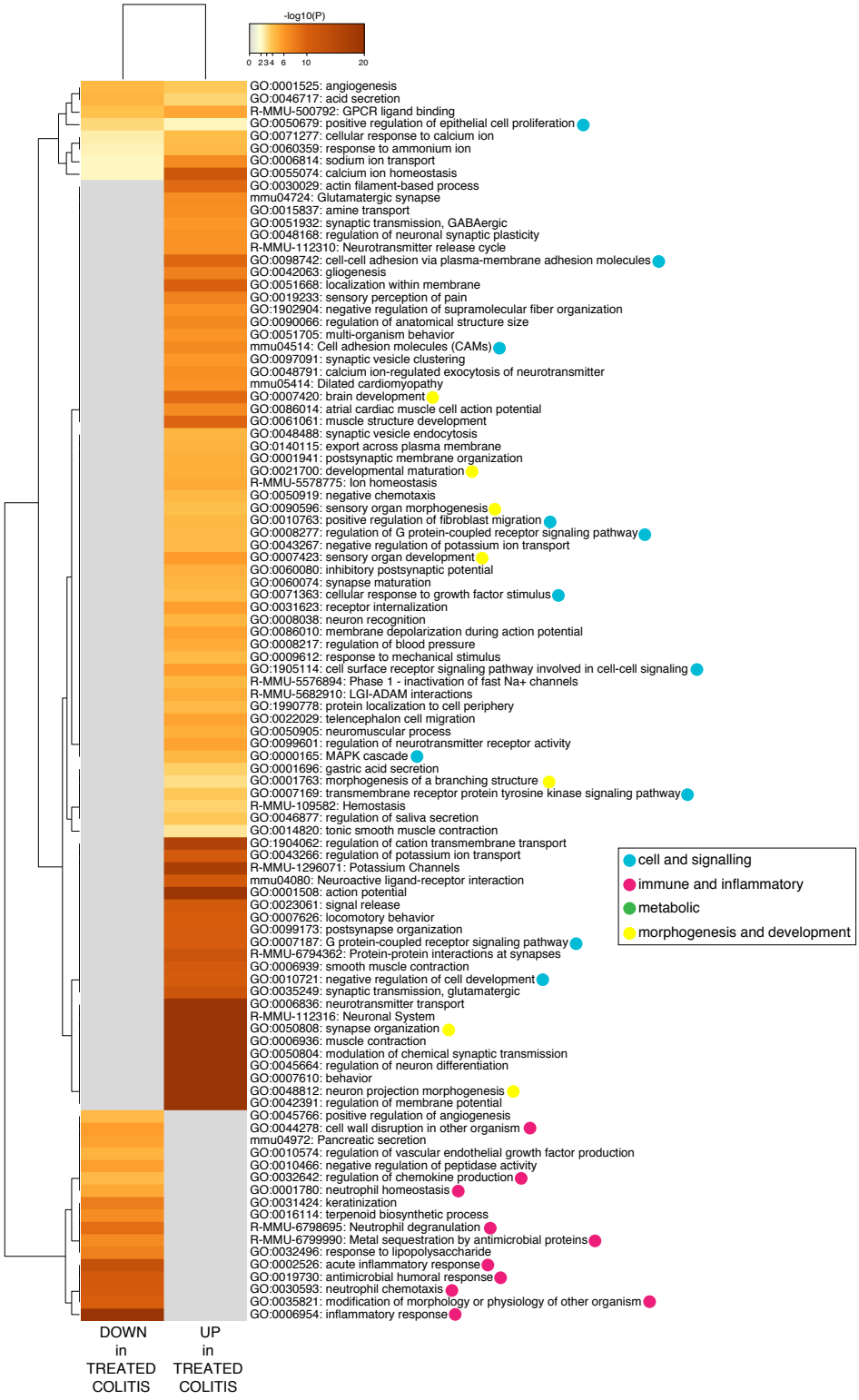


Figure 6.31: Heatmap of the top 100 pathways enriched in lists of differentially expressed genes in treated colitis compared to untreated colitis, coloured according to significance. Dots indicate biological relevance of the indicated pathways.

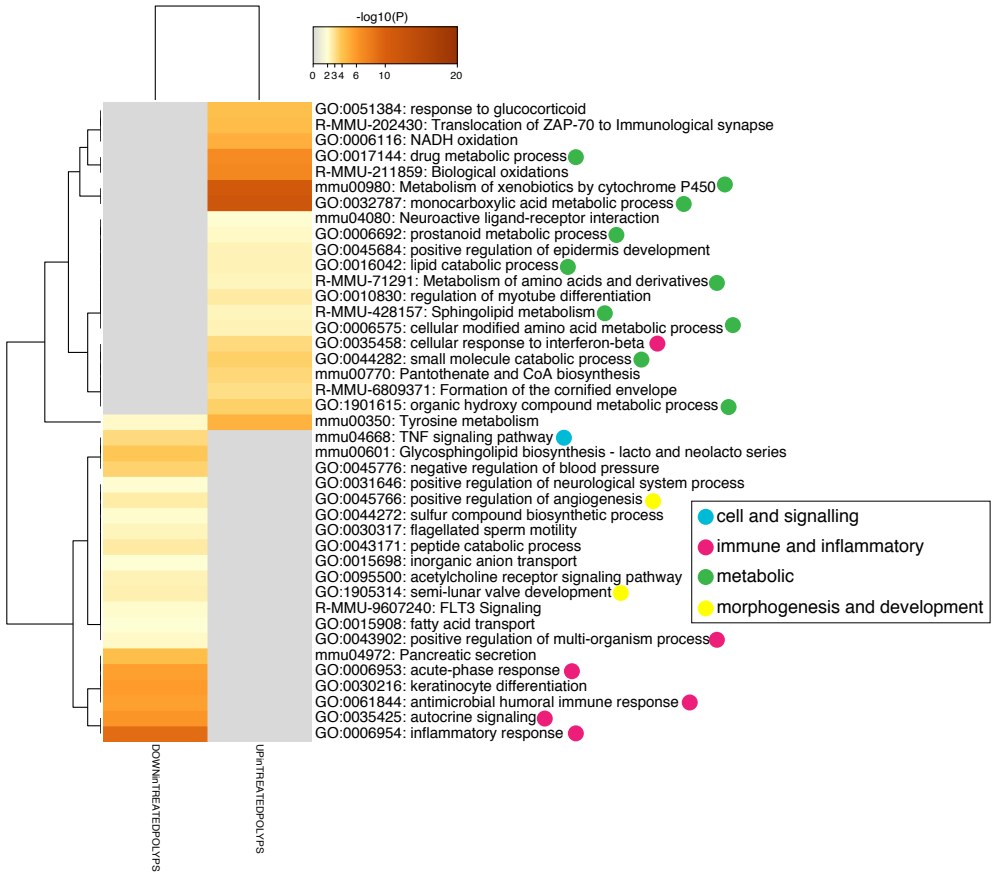


Figure 6.32: Heatmap of the top 100 pathways enriched in the lists of genes differentially expressed in anti-IL-22-treated polyps compared to untreated polyps, coloured according to significance. Dots indicate biological relevance of the indicated pathways.

associated with cellular and morphogenetic pathways, while downregulated genes correspond largely to immune and inflammatory processes. This is in line with previous work from Kirchberger *et al* in which neutralisation of IL-22 abrogated the inflammatory drive in treated tissue [57]. Enrichment analysis of genes differentially expressed in polyps revealed that only 21 pathways were enriched in genes upregulated in treated polyps and 20 in genes downregulated in treated polyps (Figure 6.32).

Processes enriched in upregulated genes correspond to metabolic pathways, while those enriched in downregulated genes correspond to the immune response and morphogenesis. This indicates that while anti-IL-22 treatment also targets inflammatory pathways in polyps, it does not result in a comprehensive program of transcriptomic change.

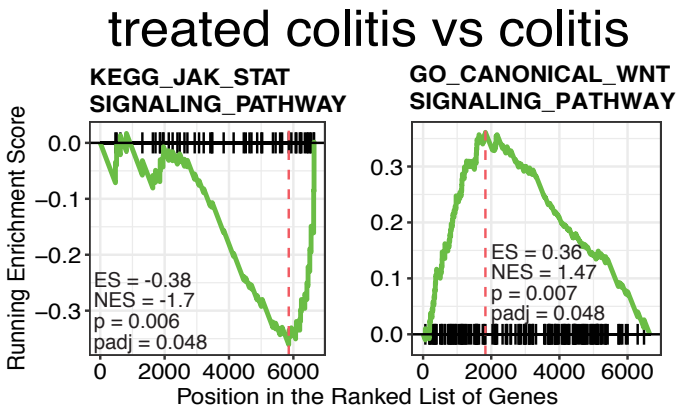


Figure 6.33: GSEA of differentially expressed genes in anti-IL-22 treated colitis compared to untreated colitis. Enrichment score (ES), Normalised enrichment score (NES), nominal and multiple testing-adjusted p-values as calculated by the GSEA function (clusterProfiler package, R).

In order to specifically characterise the effect of anti-IL-22 therapy the intestinal pathways of interest described earlier, I next carried out GSEA of the genes differentially expressed in treated compared to untreated samples. Strikingly, treating colitis samples with anti-IL-22 results in negative enrichment of IL-22-associated Jak-Stat signalling and positive enrichment of canonical Wnt signalling (Figure 6.33). This indicates that the restorative phenotypic and transciptomic effect of anti-Il-22 therapy observed in colitis samples occurs via rescue of dysregulated Jak-Stat and Wnt signalling, which earlier results have shown to drive colitis in this model. Importantly, GSEA analysis of the genes differentially expressed in treated polyps compared to untreated polyps revealed no significantly enriched pathways. This indicates that anti-IL-22 therapy has minimal impact on the processes enriched in polyps in this model.

6.2.3.4 Intestinal cell populations and signalling pathways in treated samples

It may thus be that the processes enriched in the progression from colitis to colitic polyp, characterised earlier in this section, are sufficient to protect polyps against the transcriptomic effect of anti-IL-22 treatment observed in treated colitis samples. Alternatively, the changes that drive polyp formation, unlike colitis, may be imprinted through somatic (epi)genetic alterations. In order to explore

this and further contrast the impact of therapy in colitis and polyps, I next characterised the expression of key intestinal cell populations and signalling pathway markers in treated colitis and polyp samples. In order to characterise any changes in expression within the context of overall disease progression in this model, I visualised the expression of each marker within treated samples alongside steady state and untreated samples.

Il-22

The increase in *IL-22* expression in untreated colitis and polyp samples compared to steady state samples is unchanged by anti-IL-22 therapy (Figure 6.34). The increase in IL-22 signalling, represented by *Stat3* expression, observed in untreated colitis and polyp samples, however, is abrogated in treated samples. This occurs to a greater degree in treated colitis samples, in which *Stat3* expression is returned almost to steady state levels. Treatment with anti-IL-22 does not produce any change in expression in IL-22 receptors in colitis samples, although expression of *IL22ra2* is partially restored in treated compared to untreated polyps. Increased *IL22ra2* expression may thus underpin the incomplete rescue of *Stat3* expression observed in treated polyp samples compared to treated colitis samples.

Interestingly, decreased expression of the environmental sensor *Ahr* in colitis was rescued in treated colitis samples, indicating that abrogation of the inflammatory signal may restore expression of environmental sensing mechanisms. Strikingly, however, expression of the key DNA damage response gene *Atm* was unchanged by anti-IL-22 therapy in both colitis and polyp samples, such that all treated and untreated samples exhibited significantly reduced levels of *Atm* expression relative to steady state. Together, these findings indicate that IL-22 signalling activity and *Ahr* expression is restored in treated colitis samples, and partially so in treated polyp samples, while *Atm* expression is reduced in this model regardless of therapy.

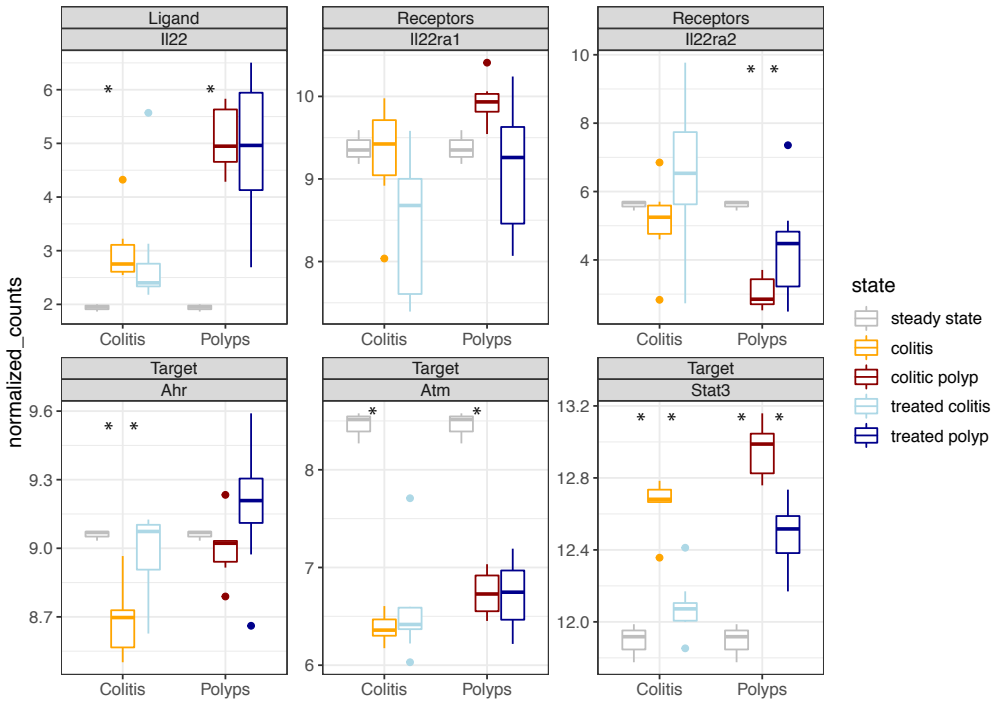


Figure 6.34: Normalized expression of genes involved in Il-22 signalling within steady state, colitis, colitic polyp, anti-IL-22-treated colitis and anti-IL-22-treated polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

Stem, proliferative and differentiated cells

I next characterised the impact of anti-IL-22 therapy on stem and proliferative cell populations. Strikingly, treatment with anti-IL-22 rescued expression of all stem cell markers downregulated in untreated colitis samples (*Ascl2*, *Lrig1*, *Tert*), and abrogated increased expression of the proliferative cell markers *Prom1* and *Prss32* (Figure 6.35). While treated polyps also exhibited a reduction in the expression of *Prom1*, the expression of stem cell markers upregulated in untreated polyps (*Lgr5*, *Sox9*, *Bmi1*, *Hopx*) remained unchanged. This indicates that anti-IL-22 therapy rescues reduced expression of stem and proliferative cell markers in colitis, while untreated and treated polyps that have escaped therapeutic rescue continue to exhibit an increase in stem cell markers.

Analysis of the expression of markers of differentiated intestinal cells revealed that anti-IL-22 therapy restores steady state expression of goblet and Paneth cell markers in both treated colitis and treated polyps (Figure 6.36). There was no

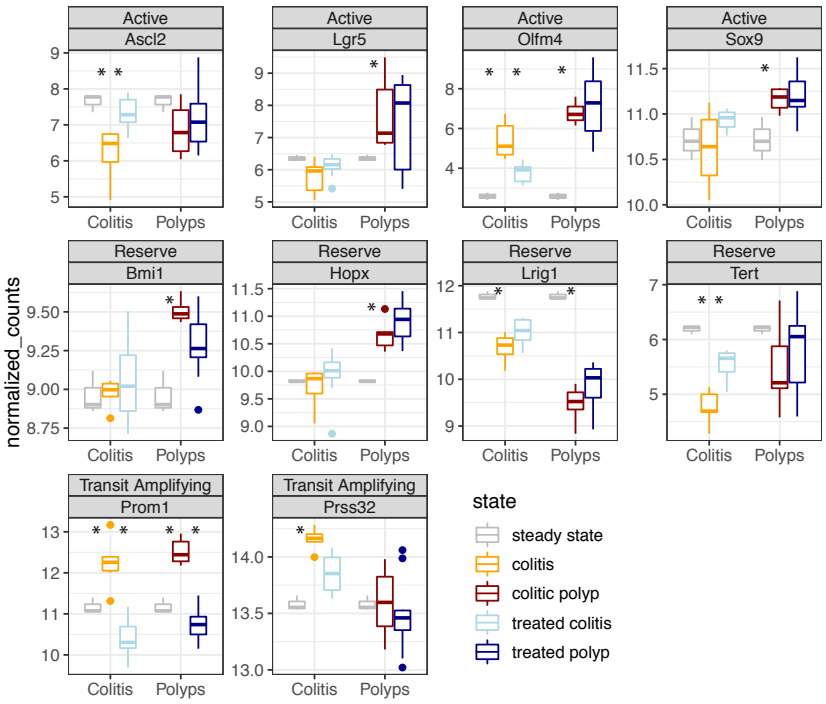


Figure 6.35: Normalized expression of intestinal stem and proliferating cell markers within steady state, colitis, colitic polyp, anti-IL-22-treated colitis and anti-IL-22-treated polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

difference in the impact of therapy in colitis and polyp samples, with anti-IL-22 restoring the expression of goblet and Paneth cell markers in all treated samples.

Wnt signalling

In treated colitis samples, therapy restored expression of the Wnt antagonist *Dkk3*, the non-canonical ligand *Wnt5b* and the receptors *Fzd2*, *Fzd3*, *Fzd4*, *Fzd7* and *Fzd8* (Figure 6.37). This resulted in the restoration of steady state levels of Wnt activity in treated colitis samples, represented by increased expression of the canonical Wnt target genes *Axin2* and *Ephb3*.

Strikingly, in contrast to the changes observed in colitis, anti-IL-22 therapy produced no changes in the expression of any Wnt pathway component in treated polyps compared to untreated polyps, with the exception of the canonical Wnt receptor *Fzd1*, which was downregulated in all treated compared to untreated samples. I have shown previously that colitis in this model is driven by reduced

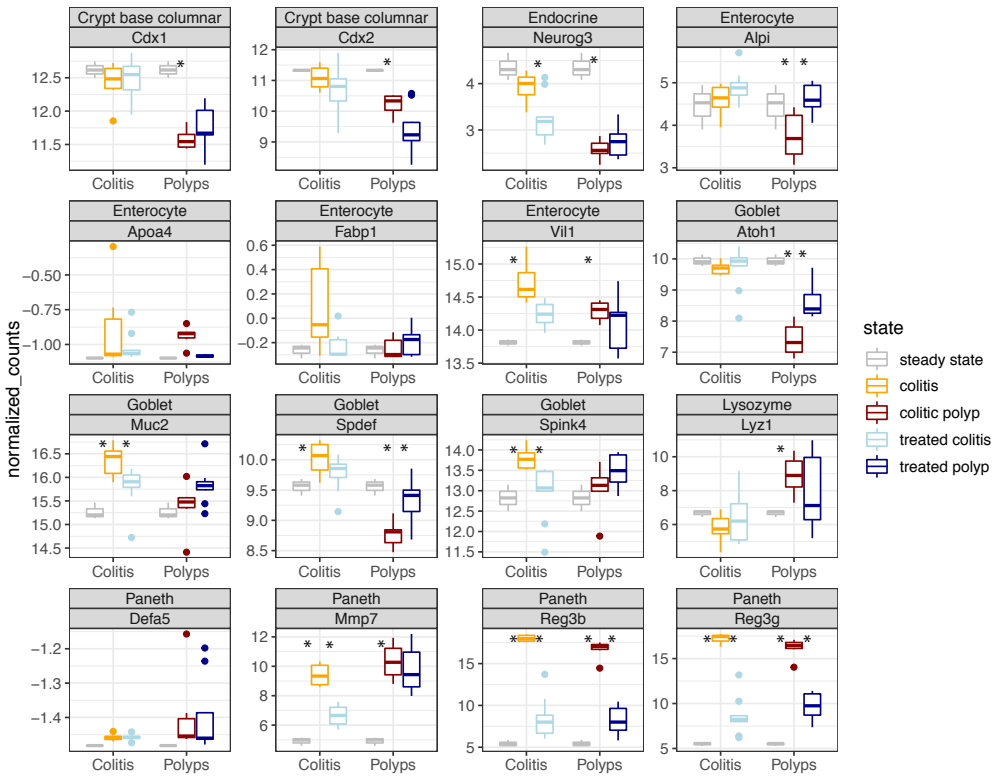


Figure 6.36: Normalized expression of markers of intestinal epithelial cell types within steady state, colitis, colitic polyp, anti-IL-22-treated colitis and anti-IL-22-treated polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

Wnt activity, which is markedly increased in the progression to polyps through increased expression of non-canonical Wnt ligands and receptors. It may thus be that the increase in non-canonical Wnt activity within untreated polyps enables escape from therapeutic restoration of Wnt activity, which appears to act largely upon canonical Wnt pathway components. This may explain that, while expression of the Wnt target genes *Axin2* and *Cd44* was unchanged in treated polyps compared to untreated polyps, only the canonical Wnt target *Ephb3* was reduced in treated polyps. Anti-IL-22 therapy may therefore rescue the reduction in expression and subsequent activity of canonical Wnt signalling components in colitis; but is unable to rescue or indeed results in increased non-canonical Wnt activity in polyps.

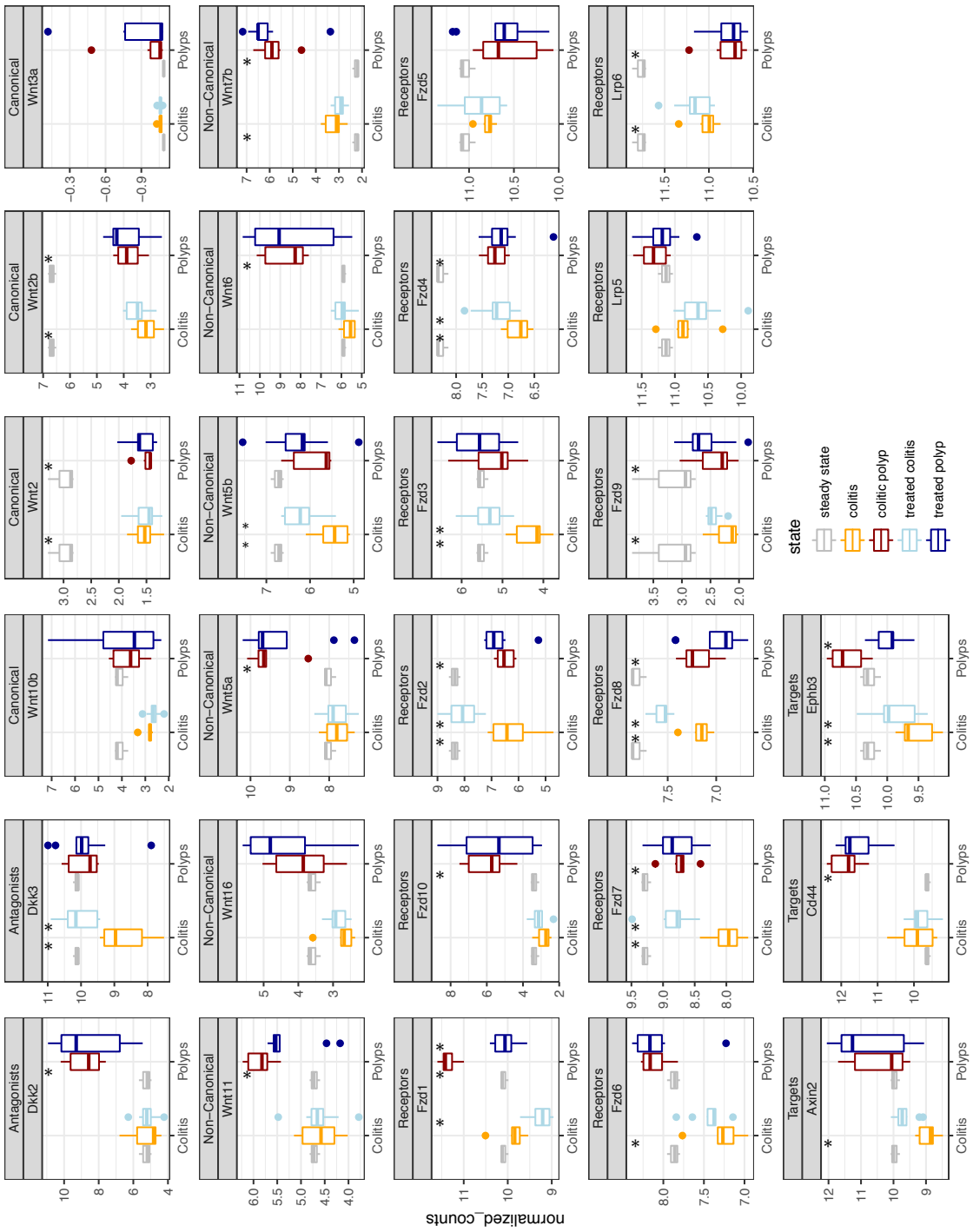


Figure 6.37: Normalized expression of genes involved in Wnt signalling within steady state, colitis, colitic polyp, anti-IL-22-treated colitis and anti-IL-22-treated polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DESeq2 package, R).

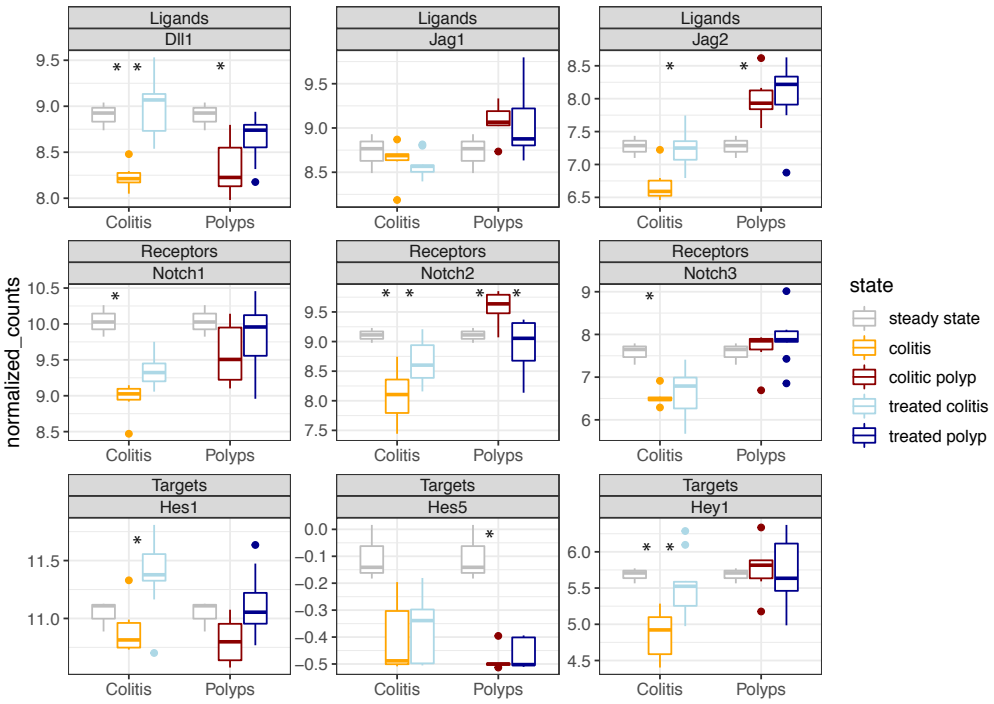


Figure 6.38: Normalized expression of genes involved in Notch signalling within steady state, colitis, colitic polyp, anti-IL-22-treated colitis and anti-IL-22-treated polyp samples. *multiple testing-corrected p-value < 0.05 as calculated by the DESeq function (DEseq2 package, R).

Notch signalling

Characterisation of the impact of anti-IL-22 treatment in treated colitis and polyp samples revealed therapeutic rescue of reduced expression of Notch ligands *Dll1* and *Jag2* in treated colitis samples, resulting in increased expression of the Notch targets *Hes1* and *Hey1* in treated compared to untreated colitis (Figure 6.38). Despite an upregulation of the *Notch2* receptor in treated polyps compared to untreated polyps, anti-IL-22 therapy did not impact the expression of Notch pathway components in polyps. Given that previous results indicated that colitis in this model is associated with a reduction in Notch signalling, while polyps harbour increased Notch activity, it may thus be that increased Notch signalling in untreated colitic polyps enables escape from therapeutic rescue.

Summary: The effect of anti-IL-22 on disease progression

The results presented here confirm the limited impact of anti-IL-22 therapy

within treated polyps compared to treated colitis samples. In treated colitis samples, anti-IL-22 treatment results in a restoration of the expression of stem and proliferative cell markers, along with increased Wnt and Notch signalling activity. Treated polyps, however, continue to exhibit the increased expression of stem cell markers and specific Wnt and Notch signalling components observed in untreated polyps. Colitis in this model is therefore associated with transcriptomic changes that are able to be reversed by anti-IL-22 therapy. This is not, however, the case in polyps, which exhibit minimal transcriptomic changes following anti-IL-22 therapy. Interestingly, I have also shown that polyp formation in this model is associated with increased expression of stem cell and non-canonical Wnt pathway components. These may therefore provide a putative proliferative drive in the absence of driver gene mutations within polyps in this model.

6.2.4 Summary of findings

In view of the aims set out at the beginning of this chapter, the key findings presented here are below.

- Anti-IL-22 therapy in 129SvEv.*Rag*$^{-/-}$ mice exposed to Hh+AOM rescues colon phenotype through a reduction of the proliferative cell compartment within treated crypts
- Treatment with anti-IL-22 does not result in a shift in the number or types of mutation present within normal or polyp samples from HH+AOM mice
- Exposure to Hh+AOM results in an increase in DNA repair-associated mutational processes in adjacent normal and polyp samples, indicating that disease progression in this model is associated with increased genomic instability.
- Treatment with anti-IL-22 does not impact upon the mutational processes active within adjacent normal samples but results in a diversification of the mutational processes active within treated polyps. This dilutes the

AOM-associated SBS11 signal in anti-IL-22-treated compared to untreated Hh+AOM polyps

- Disease progression in this model is underpinned by increased production of inflammatory cytokines such as Il-22 and IL-17, resulting in increased Jak-Stat signalling activity in all samples
- The progression from steady state to colitis is underpinned by decreased Wnt signalling and an increase in 'non-canonical', Atm-deficient DNA repair pathways
- Treatment with anti-IL-22 abrogates the transcriptional changes enacted in the progression to colitis, with a majority of the genes differentially expressed in colitis samples not so in treated colitis samples. This rescues dysregulated Wnt and Jak-stat signalling and restores expression of most stem cell, Wnt and Notch pathway markers
- Progression from colitis to colitic polyp involves increased expression of stem cell markers, non-canonical Wnt ligands and Notch receptors that results in increased Wnt activity and the upregulation of cancer-associated pathways
- Anti-IL-22 treatment has minimal effect on the differentially expressed genes and pathways that drive polyp formation in this model

6.3 Discussion

The results presented here have demonstrated that, in the absence of somatic mutations within CRC driver genes, disease pathogenesis in this model is instead evidenced by changes in the transcriptome. I have shown that while colitis and polyps both harbour increased activation of inflammatory responses, progression to polyp formation in this model is associated with increased non-canonical Wnt signalling and an enrichment of cancer-associated pathways that is unaffected

by anti-IL-22 therapy. In terms of my original hypothesis regarding the impact of therapy upon somatic mutation acquisition, I have shown that in an immune-deficient model of colitis-associated tumourigenesis, disease pathogenesis is not driven by the accumulation of somatic driver gene mutations. It may thus be that pathogenesis in this model is driven by genomic instability in the form of copy number variation or epigenetic alterations. However, disease progression remains evident in changes to the transcriptome that, in colitis, are rescued by therapy to produce a detectable reduction in disease phenotype severity. This has implications for our understanding of therapeutic mechanisms and their impact upon disease-state driven differences in the impact of mutagenesis. It could thus be that an IL-22-driven increase in non-canonical DNA repair pathways results in increased genomic instability through error-prone repair of DNA damage. This could result in increased CNV or epigenetic alterations that drive tumourigenesis in this model, with the selective pressure of anti-IL-22 therapy then selecting for epigenetic alterations that increase non-canonical Wnt signalling and cancer-associated pathways.

The double-edged role of IL-22 in intestinal homeostasis and disease is well documented, with IL-22 signalling and subsequent activation of Stat3 signalling underpinning proliferation and the regenerative response of intestinal epithelial cells; but increased IL-22 levels also associated with IBD and chronic inflammation. An intriguing theory potentially explaining the lack of detectable AOM-driven mutagenesis within this model centres upon recent work published by Gronke *et al* demonstrating a role for IL-22 signalling in DDR pathway activation. Their work showed that IL-22 is required for effective initiation of DDR following DNA damage, with stem cells deprived of IL-22 signals prior to exposure to AOM escaping DDR-initiated apoptosis, harbouring more mutations and being more likely to give rise to colon cancer [49]. They went on to demonstrate the role of the environmental sensor Ahr in an IL-22-DDR feedback loop, with the sensing of dietary toxins by AHR resulting in production of IL-22 that facilitates intact DDR responses to the genotoxic contents of the gut lumen. Given that the treatment

scheme utilised here involved the initiation of chronic colitis, shown to result in increased levels of IL-22 [57], prior to mutagenesis, it may thus be that elevated inflammation-associated IL-22 also conferred protection against AOM-induced mutagenesis. This is supported by neither isotype nor anti-IL-22 treated mice harbouring notable mutation burden in normal or polyp samples, given that inflammation and mutagenesis took place prior to antibody administration in both experimental groups. This could be explored by implementing earlier collection timepoints to the experimental protocol and initiating antibody therapy earlier in the treatment protocol.

However, the Ahr-IL-22-Atm axis proposed by Gronke *et al* is in conflict with the findings presented here, in which increased IL-22 expression is associated with marked reductions in both *Ahr* and *Atm* expression. While *Ahr* expression is responsive to therapeutic rescue by anti-IL-22, *Atm* expression is not. Given that Gronke *et al* specifically focused upon *Lgr5*+ stem cells, it may thus be that different stem cell populations within the intestine harbour differential sensitivity to inflammatory cytokines. This raises interesting questions about the role of active and reserve stem cell populations within an inflammatory setting, whereby specific populations of cells vary in their responsiveness to inflammatory cytokines. This could be explored further by isolating cells expressing specific stem cell markers and characterising their response to inflammation and therapy. This has clinical relevance given that IBD is associated with increased IL-22 expression, so it may be that patients with inflammation-driven intestinal disease vary in their susceptibility to therapeutic rescue based on the specific populations of cells that are expanded. This is supported by the findings presented in Chapter 5, whereby treated ectopic crypts in VG mice appear to take a non-conventional tumourigenic pathway involving the acquisition of mutations in DNA repair genes. It may thus be that in immune-deficient Rag mice under the pathogenic pressure of *Hh* infection, stem cells that are differentially responsive to IL-22 transcriptomically upregulate non-canonical DNA repair mechanisms and Wnt signalling in order to provide a proliferative drive and impair (epi)genetic stability.

The findings presented within this chapter may have been limited by caveats limiting the analytical approach utilised. Firstly, the sampling of colonic polyps may have been prone to error due to presence of inflammatory nodules. However, I have significant experience in the identification and sampling of colon polyps and have presented a representative image to demonstrate the distinct macroscopic appearance of polyps within this model. The clustering of samples deemed to be polyps following principle component analysis further supports the successful identification of these samples as true colonic polyps. The main caveat to the results presented here, however, is that RNA and DNA were not isolated from the same polyps, so the observations from the genomic analysis presented here may not apply to the samples that underwent transcriptomic analysis. This is therefore a major limitation of these findings that could be rectified in future experiments through simultaneous isolation of RNA and DNA from collected samples. The results of the genomic and transcriptomic analyses presented here should thus be considered separately, such that while disease progression does not appear to be driven by the accumulation of somatic driver gene mutations within the polyp samples genomically analysed, it remains the case that disease progression and the effect of anti-IL-22 therapy are evident as changes in gene expression. Another potential caveat that represents an interesting avenue for further work is the role of the microbiome in disease progression within this model. *Hh*-driven colitis is also dependent upon 'clean' mice, and many of the pathways implicated in these findings are involved in the response to bacterial infection, so it could be that the microbiome is another component of disease states that influences disease progression in these mice. Further work is underway to collect faecal samples alongside intestinal samples to investigate changes in the microbiome that may impact upon disease state and subsequent susceptibility to mutagenesis and response to therapy. Finally, it is important to note the limited resolution of whole exome sequencing depth, both in terms of clonal resolution and the subsequent impact upon VAF and mutation signature analysis. It may be that both disease state and therapy elicit genetic changes that occur beneath the

threshold for detection employed in our analyses. While the results presented here remain biologically relevant in terms of clonally expanded mutations or polyclonal mutations arising simultaneously, future work will benefit from added resolution that enables detection of small clone sizes and sheds light on early mutation dynamics within these models.

7 Discussion

7.1 Summary of key findings

The work presented within this thesis sought to explore somatic mutation acquisition as a window to intestinal disease pathogenesis. In Chapter 3, I demonstrated the capacity of normal human colonic epithelium to 'tolerate' the acquisition of somatic mutations within CRC driver genes and showed that inflamed epithelium exhibits increased mutation acquisition. Using mouse models of spontaneous and chronic intestinal disease, I went on to show that changes in underlying disease state influence mutation acquisition and provided novel evidence of an AOM-associated mutational signature in mice. In Chapter 4, I provided evidence linking constitutively deranged BMP signalling to impaired DDR pathways in VG mice and proposed that a novel Grem1-miR-96 axis influences genomic instability in ectopic crypts. In Chapter 5, I went on to demonstrate that therapeutic restoration of morphogen gradients in VG mice can rescue disease phenotype and confer protection against mutation acquisition. Crucially, however, I have also shown that the selective pressures introduced by anti-Grem1 therapy can result in a putative skew to DNA-repair associated mutation acquisition and a putative 'hypermutated' phenotype. In Chapter 6, I provided evidence that disease progression in an immune-deficient model of

inflammation-associated cancer occurs in the absence of mutation accumulation, but remains detectable through dysregulated gene expression. Importantly, abrogation of the inflammatory drive in this model rescued disease phenotype in colitis but not polyps.

7.2 Discussion

7.2.1 Updating the genotype-phenotype model: persistence is key

Cancer has long been suspected to be a disease of somatic acquisition in a fixed stem cell. In recent years, this theory has been updated to reflect increased understanding of the importance of additional factors influencing disease pathogenesis, such as the cell of origin and microenvironment. It is now generally accepted that the endlessly variable interactions between these factors contributes the vast heterogeneity of diseases such as CRC.
At the beginning of this thesis, I described a model in which intestinal disease phenotype is influenced by genotype, divided into genetic predisposition and somatic mutation, and environment, divided into cell-extrinsic factors, such as the signalling environment, and cell-intrinsic factors, such as the cell of origin. The results presented here suggest that it may now be pertinent to update this model to include the notion of persistence. In order for a mutated cell to persist it not only has to acquire a mutation but also survive long enough for it to have an effect - if the environment facilitates this then it is equally as culpable as the mutation itself. Based on the findings of this thesis, I have thus proposed a model in which disease progression is influenced by mutation persistence. Within this model, the persistence of a mutated cell is determined by the acquisition of a mutation, influenced by the level of genomic stability, and its environment, influenced by intact signalling pathways and intestinal architecture (Figure 7.1).

Crucially, the factors that inform mutation persistence are interrelated and can thus enhance or inhibit each other to varying degrees. Mutation persistence during homeostasis is therefore low, given that a mutated cell would have to

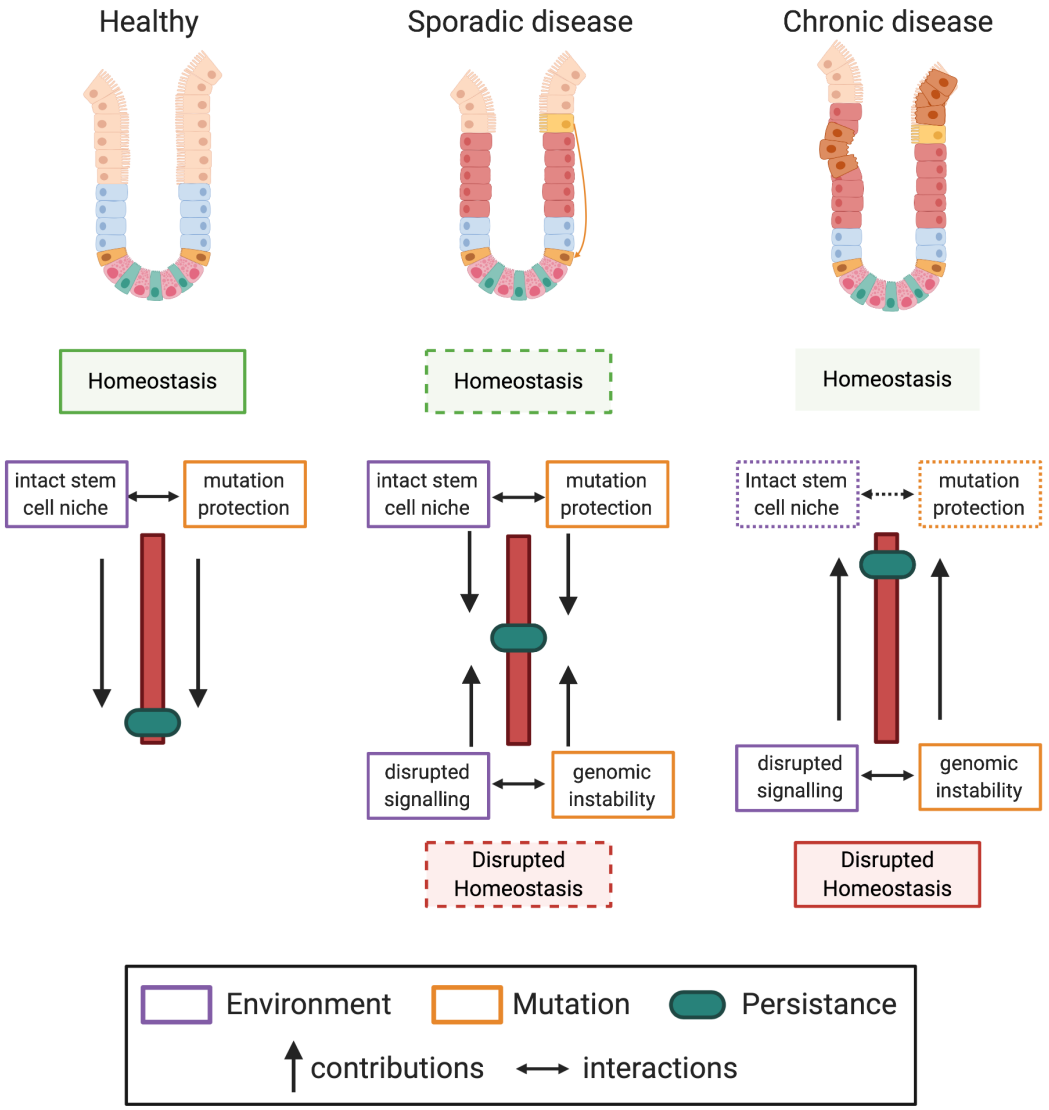


Figure 7.1: Model depicting the factors that influence the persistence of mutated cells. Persistence within this model is determined by the variable contributions of mutation acquisition (influenced by factors in orange boxes) and the environment (influenced by factors in purple boxes). Vertical arrows represent the contribution of mutation and environment to the persistence of mutated cells. Horizontal arrows represent the interaction between factors that influence the contribution of mutation and environment. The persistence of mutated cells within each disease state is represented by the teal rheostat.

overcome strictly regulated signalling pathways and intact mutation protection mechanisms to persist. In the event of sporadic inflammation or intestinal injury, transient disruption of homeostasis to facilitate regeneration and plasticity increases mutation persistence. This is evidenced by findings in Chapter 3, which demonstrated increased mutation acquisition in biopsies of inflamed colon. In the event of chronic intestinal disease, homeostatic maintenance of signalling pathway gradients and mutation protection mechanisms is constitutively compromised, resulting in in high mutation persistence. This is evidenced by increased mutation accumulation in VG mice and, in humans, increased cancer risk of patients with chronic intestinal disease.

7.2.2 The influence of persistence on disease progression

If the interaction between mutation and environment can influence tumour evolution, this may explain the wide variation in phenotypic outcomes observed in CRC. Importantly, it also provides a mechanistic basis for the capacity of tumours to adapt and evolve under environmental and therapeutic pressures and explains why tumours harbouring the same mutation might behave differently. The prospect of tumours coevolving with the microenvironment further illustrates our departure from an understanding of CRC that is underpinned by the occurrence of epithelial somatic mutation. Work presented in this thesis has confirmed that disease pathogenesis is far more multifactorial, and that variation within each factor contributes to disease heterogeneity. The notion of mutation persistence is therefore central to our understanding of the most appropriate therapeutic interventions for specific intestinal diseases. In order to improve therapies, we need to not only able to understand the co-evolution of somatically mutant epithelium with its environment but also the key points of this interaction that are able to be clinically assessed, often from a single fixed time point (such as tumour sampling at diagnosis). The field of personalised medicine is already moving towards a more holistic understanding of the factors that underpin intestinal neoplasia. Maley *et al* have proposed a classification system for CRCs

that incorporates an evolutionary index, stratified into mutation diversity and change over time, and an ecological index, stratified into hazards and resources [179]. Similar to the persistence model I have described here, their ultimate goal is to better target specific therapeutic strategies to the neoplasms, or indeed disease states, most susceptible to them. The gold standard of personalised CRC therapy would then involve the development of a range of therapies that work together to probe the key weaknesses identified.

7.2.3 Exploiting mutation persistence for novel therapeutic strategies

Excitingly, personalised cancer therapies appear to be within reach for patients harbouring mismatch repair-deficient and microsatellite instability-high CRC. Increased mutation accumulation in these cancers leads to the generation of neoantigens that can be recognised by the immune system, but the environmental aspect that facilitates mutation persistence here is the avoidance of immune detection by these tumours due to expression of immune checkpoint inhibitors that downregulate the immune response [180]. Exploiting our understanding of mutation persistence in these cancers has thus enabled development of a combined therapeutic strategy employing therapeutic enhancement of the immune response following radiation, so that increased radiation-induced mutation turns a previously 'immune-cold' tumour to an 'immune-hot' tumour that activates the innate response.

An alternative approach that utilises therapy as a selective pressure to exploit the persistence rheostat involves the use of PARP1 inhibitors in BRCA1-deficient breast and prostate cancer. Within these lesions, increased mutation persistence is facilitated by an inability to accurately repair harmful DNA damage via HR, resulting in the use of error-prone, PARP1-directed repair [181]. This is exploited through the combination of chemotherapy, to induce DNA damage, and PARP1 inhibition, to inhibit PARP1-directed repair. This strategy therefore forces mutated cells down an evolutionary dead end, given that tumours harbouring

extremely high mutation rates are at increased risk of deleterious mutations and the generation of non-viable daughter cells [179].

The interaction between mutation and environment can also result in therapeutic interventions that target the environment also inherently impacting upon mutation acquisition, as evidenced by results presented in Chapter 5. 'Resistance' mutations, which may be deleterious or neutral in the absence of therapy, could thus become selectively advantageous for neoplastic cells exposed to therapy. This further illustrates the importance of a multifactorial approach, not only towards understanding disease progression but also in the identification of potential vulnerabilities targeted by therapy.

7.2.4 Conclusion

In this thesis, I have demonstrated that intestinal disease is underpinned by a variable complex of interacting factors that contribute to disease progression in a context-dependent manner. I have utilised mouse models of spontaneous and chronic intestinal disease to illustrate the influence of differences in underlying disease state on mutation acquisition. The findings presented here also shed light upon the crosstalk between morphogenetic signalling pathways and mutation protection mechanisms, which represents an important aspect of future therapeutic strategies aimed at targeting these factors within intestinal disease. Overall, I have demonstrated that somatic mutation acquisition within the intestine varies and reflects the influence of the microenvironment, cell of origin and mutation protection mechanisms upon disease pathogenesis. The work presented here thus contributes to a more holistic and accurate understanding of the mechanisms underpinning the vast heterogeneity of intestinal disease. It is my hope that it will move us closer to realising the once distant ambition of personalised therapy in colorectal cancer.

"Because, look, even the missing are there, the gone and taken are with them in the shade pools of the peppermints by the beautiful, the beautiful the river."

- Tim Winton, *Cloudstreet*

Appendices

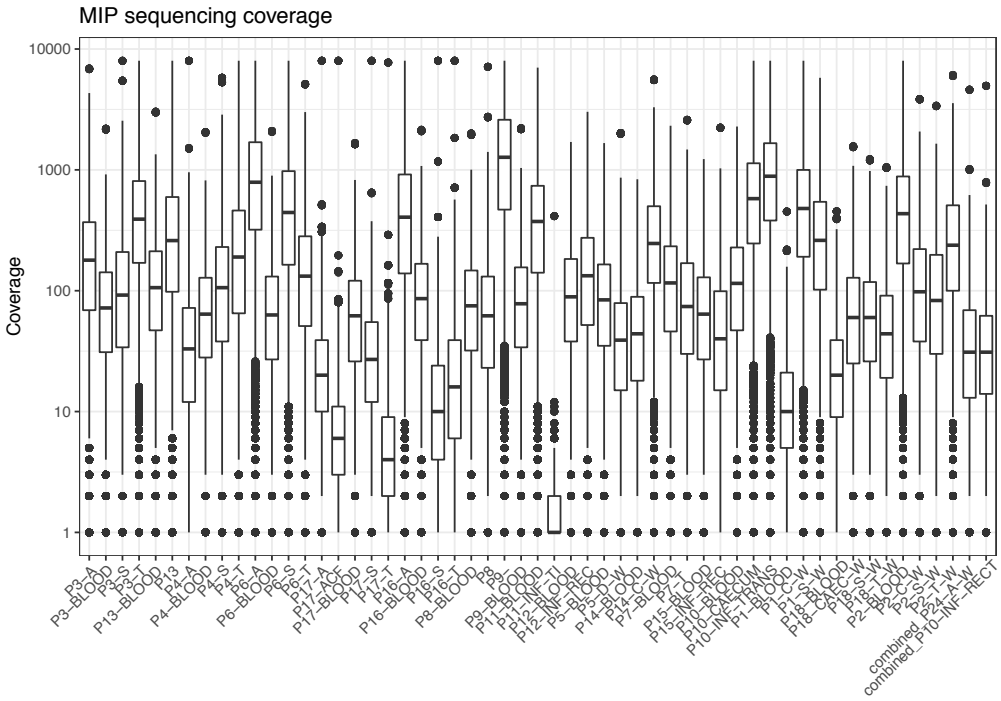


Figure 2: Bar plot showing MIP sequencing read depth coverage per sample. Samples with mean coverage < 10x were omitted from downstream analysis.

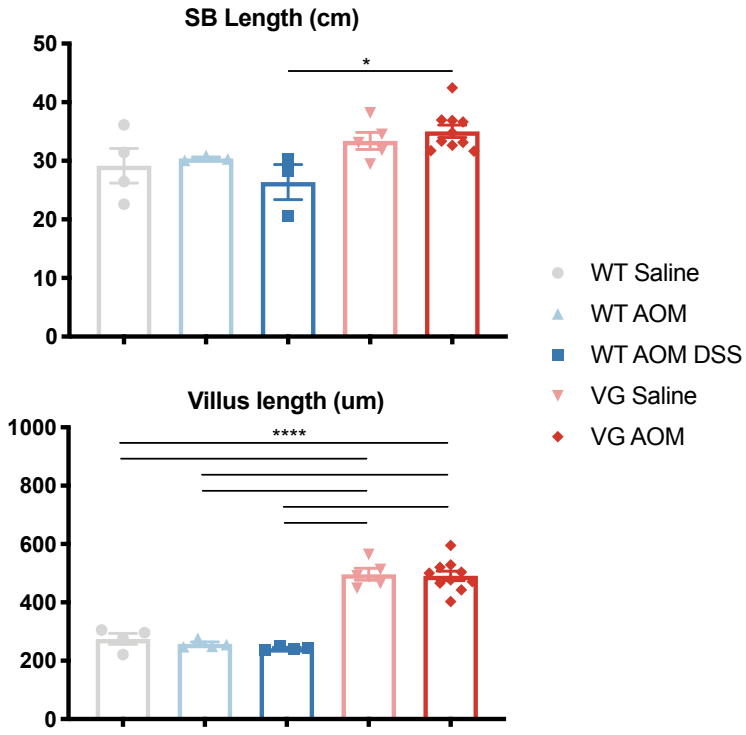


Figure 3: Average small bowel (SB) and villus length of mice within each experimental group. Points represent individual mice (top) and average villus length per mouse (bottom), coloured according to experimental group. *p<0.05, ****p<0.0001 as calculated by one-way ANOVA with Tukey post-test.)

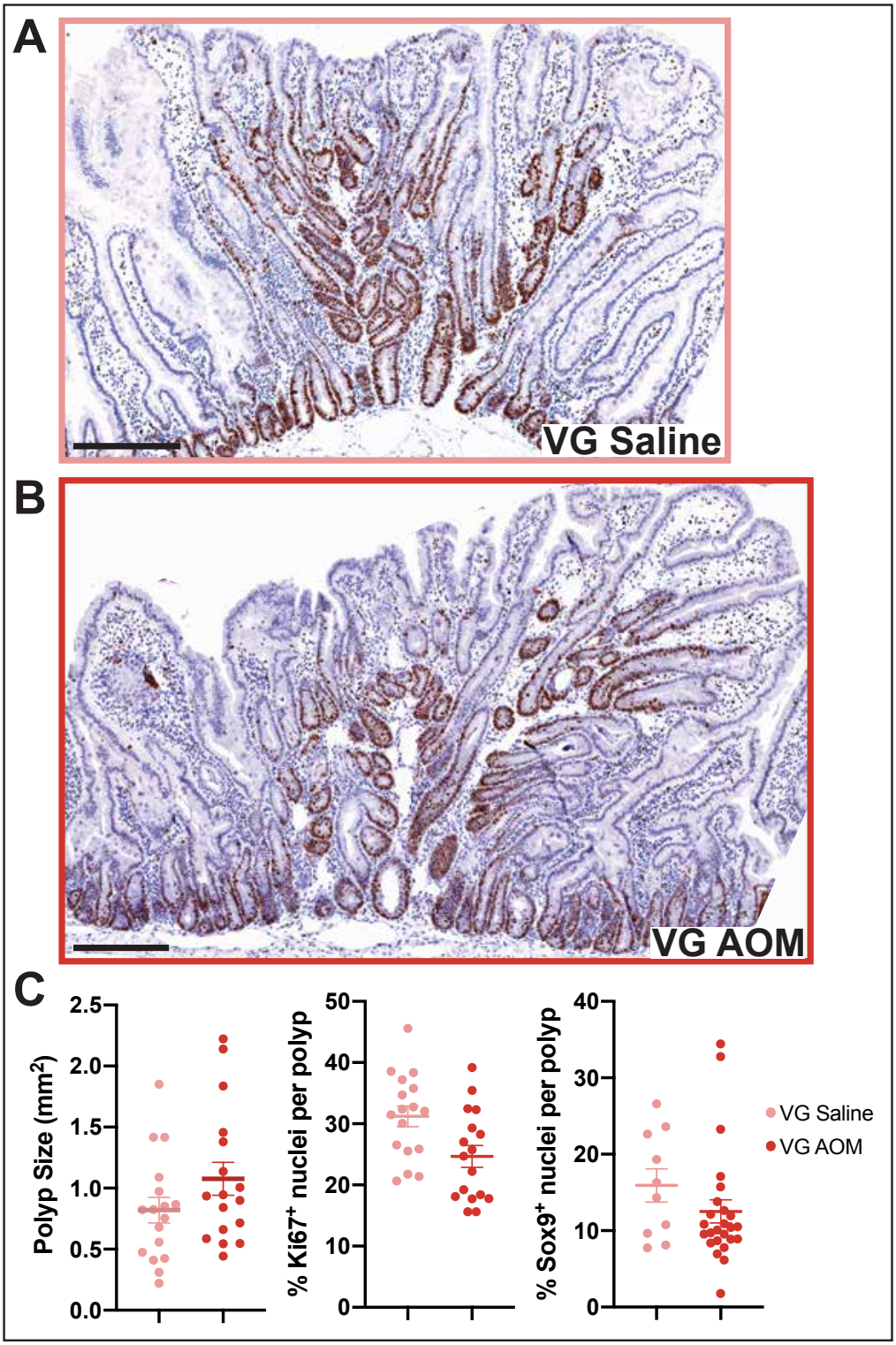


Figure 4: Characterisation of small bowel polyps in *Vil1-Grem1* mice. Representative images of IHC Ki67 staining in small bowel polyps from VG Saline (A) and VG AOM (B) mice. C) Analysis of polyp size and the percentage of proliferative and progenitor cells per polyp, represented by Ki67+ and Sox9+ nuclei per polyp per mouse. Individual points represent polyps, coloured according to experimnental group.

Gene Ontology
CELLULAR RESPONSE TO DNA DAMAGE STIMULUS
CANONICAL WNT SIGNALING PATHWAY
NON CANONICAL WNT SIGNALING PATHWAY
NOTCH SIGNALING PATHWAY
DNA REPAIR
RESPONSE TO IONIZING RADIATION
INTERLEUKIN 23 MEDIATED SIGNALING PATHWAY
INTERLEUKIN 17 PRODUCTION
INTERLEUKIN 23 PRODUCTION
Kyoto Encyclopedia of Genes and Genomes
NUCLEOTIDE EXCISION REPAIR
NOTCH SIGNALING PATHWAY
HOMOLOGOUS RECOMBINATION
NON HOMOLOGOUS END JOINING
BASE EXCISION REPAIR
APOPTOSIS
PATHWAYS IN CANCER
COLORECTAL CANCER
MISMATCH REPAIR
JAK STAT SIGNALING PATHWAY
Molecular Signatures Database (v7.1)
RASHI RESPONSE TO IONIZING RADIATION 6
RASHI RESPONSE TO IONIZING RADIATION 5
RASHI RESPONSE TO IONIZING RADIATION 4
RASHI RESPONSE TO IONIZING RADIATION 3
RASHI RESPONSE TO IONIZING RADIATION 2
RASHI RESPONSE TO IONIZING RADIATION 1

Table 1: Gene sets utilised for gene set enrichment analysis, stratified according to database.

References

[1] H. Gehart and H. Clevers. "Tales from the crypt: new insights into intestinal stem cells". *Nature Reviews Gastroenterology & Hepatology* 16.1 (2019), pp. 19–34.

[2] D. J. Huels and O. J. Sansom. "Stem vs non-stem cell origin of colorectal cancer". *British Journal of Cancer* 113.1 (2015), pp. 1–5.

[3] S. Takashima, D. Gold, and V. Hartenstein. "Stem cells and lineages of the intestine: a developmental and evolutionary perspective". *Development genes and evolution* 223.1-2 (2013), pp. 85–102.

[4] S. Roth, P. Franken, A. Sacchetti, A. Kremer, K. Anderson, O. Sansom, and R. Fodde. "Paneth Cells in Intestinal Homeostasis and Tissue Injury". *PLoS ONE* 7.6 (2012). Ed. by S. R. Singh, e38965.

[5] T. Sato, J. H. van Es, H. J. Snippert, D. E. Stange, R. G. Vries, M. van den Born, N. Barker, N. F. Shroyer, M. van de Wetering, and H. Clevers. "Paneth cells constitute the niche for Lgr5 stem cells in intestinal crypts". *Nature* 469.7330 (2011), pp. 415–418.

[6] A. S. Darwich, U. Aslam, D. M. Ashcroft, and A. Rostami-Hodjegan. "Meta-analysis of the turnover of intestinal epithelia in preclinical animal species and humans". *Drug Metabolism and Disposition* 42.12 (2014), pp. 2016–2022.

[7] H. Cheng and C. P. Leblond. "Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. V. Unitarian Theory of the origin of the four epithelial cell types." *The American journal of anatomy* 141.4 (1974), pp. 537–561.

[8] D. J. Winton, M. A. Blount, and B. A. J. Ponder. "A clonal marker induced by mutation in mouse intestinal epithelium". *Nature* 333.6172 (1988), pp. 463–466.

[9] M. Bjerknes and H. Cheng. "Clonal analysis of mouse intestinal epithelial progenitors". *Gastroenterology* 116.1 (1999), pp. 7–14.

[10] N. Barker, J. H. Van Es, J. Kuipers, P. Kujala, M. Van Den Born, M. Cozijnsen, A. Haegebarth, J. Korving, H. Begthel, P. J. Peters, and H. Clevers. "Identification of stem cells in small intestine and colon by marker gene Lgr5". *Nature* 449.7165 (2007), pp. 1003–1007.

[11] C. S. Potten, W. J. Hume, P. Reid, and J. Cairns. "The segregation of DNA in epithelial stem cells". *Cell* 15.3 (1978), pp. 899–906.

[12] C. S. Potten, G. Owen, and D. Booth. "Intestinal stem cells protect their genome by selective segregation of template DNA strands". *Journal of cell science* 115.11 (2002), pp. 2381–2388.

[13] A. E. Powell, Y. Wang, Y. Li, E. J. Poulin, A. L. Means, M. K. Washington, J. N. Higginbotham, A. Juchheim, N. Prasad, S. E. Levy, Y. Guo, Y. Shyr, B. J. Aronow, K. M. Haigis, J. L. Franklin, and R. J. Coffey. "The pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor". *Cell* 149.1 (2012), pp. 146–158.

[14] N. Takeda, R. Jain, M. R. LeBoeuf, Q. Wang, M. M. Lu, and J. A. Epstein. "Interconversion between intestinal stem cell populations in distinct niches". *Science (New York, N.Y.)* 334.6061 (2011), pp. 1420–1424.

[15] E. Sangiorgi and M. R. Capecchi. "Bmi1 is expressed in vivo in intestinal stem cells". *Nature genetics* 40.7 (2008), pp. 915–920.

[16] R. K. Montgomery, D. L. Carlone, C. A. Richmond, L. Farilla, M. E. G. Kranendonk, D. E. Henderson, N. Y. Baffour-Awuah, D. M. Ambruzs, L. K. Fogli, S. Algra, and D. T. Breault. "Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells." *Proceedings of the National Academy of Sciences of the United States of America* 108.1 (2011), pp. 179–184.

[17] S. J. A. Buczacki, H. I. Zecchini, A. M. Nicholson, R. Russell, L. Vermeulen, R. Kemp, and D. J. Winton. "Intestinal label-retaining cells are secretory precursors expressing Lgr5". *Nature* 495.7439 (2013), pp. 65–69.

[18] J. H. van Es, T. Sato, M. van de Wetering, A. Lyubimova, A. N. Yee Nee, A. Gregorieff, N. Sasaki, L. Zeinstra, M. van den Born, J. Korving, A. C. M. Martens, N. Barker, A. van Oudenaarden, and H. Clevers. "Dll1+ secretory progenitor cells revert to stem cells upon crypt damage". *Nature Cell Biology* 14.10 (2012), pp. 1099–1104.

[19] M. van de Wetering, E. Sancho, C. Verweij, W. de Lau, I. Oving, A. Hurlstone, K. van der Horn, E. Batlle, D. Coudreuse, A.-P. Haramis, M. Tjon-Pon-Fong, P. Moerer, M. van den Born, G. Soete, S. Pals, M. Eilers, R. Medema, and H. Clevers. "The β-Catenin/TCF-4 Complex Imposes a Crypt Progenitor Phenotype on Colorectal Cancer Cells". *Cell* 111.2 (2002), pp. 241–250.

[20] J. P. Medema and L. Vermeulen. "Microenvironmental regulation of stem cells in intestinal homeostasis and cancer". *Nature* 474.7351 (2011), pp. 318–326.

[21] R. Aoki, M. Shoshkes-Carmel, N. Gao, S. Shin, C. L. May, M. L. Golson, A. M. Zahm, M. Ray, C. L. Wiser, C. V. E. Wright, and K. H. Kaestner.

"Foxl1-Expressing Mesenchymal Cells Constitute the Intestinal Stem Cell Niche". *Cellular and Molecular Gastroenterology and Hepatology* 2.2 (2016), pp. 175–188.

[22] T. Valenta, B. Degirmenci, A. E. Moor, P. Herr, D. Zimmerli, M. B. Moor, G. Hausmann, C. Cantù, M. Aguet, and K. Basler. "Wnt ligands secreted by subepithelial mesenchymal cells are essential for the survival of intestinal stem cells and gut homeostasis". *Cell reports* 15.5 (2016), pp. 911–918.

[23] A. Glinka, C. Dolde, N. Kirsch, Y.-L. Huang, O. Kazanskaya, D. Ingelfinger, M. Boutros, C.-M. Cruciat, and C. Niehrs. "LGR4 and LGR5 are R-spondin receptors mediating Wnt/β-catenin and Wnt/PCP signalling". *EMBO reports* 12.10 (2011), pp. 1055–1061.

[24] K. S. Carmon, X. Gong, Q. Lin, A. Thomas, and Q. Liu. "R-spondins function as ligands of the orphan receptors LGR4 and LGR5 to regulate Wnt/beta-catenin signaling". *Proceedings of the National Academy of Sciences of the United States of America* 108.28 (2011), pp. 11452–11457.

[25] T. Reya and H. Clevers. "Wnt signalling in stem cells and cancer". *Nature* 434.7035 (2005), pp. 843–850.

[26] H. Ruffner, J. Sprunger, O. Charlat, J. Leighton-Davies, B. Grosshans, A. Salathe, S. Zietzling, V. Beck, M. Therier, A. Isken, Y. Xie, Y. Zhang, H. Hao, X. Shi, D. Liu, Q. Song, I. Clay, G. Hintzen, J. Tchorz, L. C. Bouchez, G. Michaud, P. Finan, V. E. Myer, T. Bouwmeester, J. Porter, M. Hild, F. Bassilana, C. N. Parker, and F. Cong. "R-Spondin Potentiates Wnt/β-Catenin Signaling through Orphan Receptors LGR4 and LGR5". *PLOS ONE* 7.7 (2012), e40976.

[27] V. Korinek, N. Barker, P. Moerer, E. van Donselaar, G. Huls, P. J. Peters, and H. Clevers. "Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4". *Nature Genetics* 19.4 (1998), pp. 379–383.

[28] X. C. He, J. Zhang, W.-G. Tong, O. Tawfik, J. Ross, D. H. Scoville, Q. Tian, X. Zeng, X. He, L. M. Wiedemann, Y. Mishina, and L. Li. "BMP signaling inhibits intestinal stem cell self-renewal through suppression of Wnt–β-catenin signaling". *Nature Genetics* 36.10 (2004), pp. 1117–1121.

[29] A.-P. G. Haramis, H. Begthel, M. van den Born, J. van Es, S. Jonkheer, G. J. A. Offerhaus, and H. Clevers. "De novo crypt formation and juvenile polyposis on BMP inhibition in mouse intestine." *Science (New York, N.Y.)* 303.5664 (2004), pp. 1684–1686.

[30] R. H. Church, A. Krishnakumar, A. Urbanek, S. Geschwindner, J. Meneely, A. Bianchi, B. Basta, S. Monaghan, C. Elliot, M. Strömstedt, N. Ferguson, F. Martin, and D. P. Brazil. "Gremlin1 preferentially binds to bone morphogenetic protein-2 (BMP-2) and BMP-4 over BMP-7". *Biochemical Journal* 466.1 (2015), pp. 55–68.

[31] M. Spit, B.-K. Koo, and M. M. Maurice. "Tales from the crypt: intestinal niche signals in tissue renewal, plasticity and cancer". *Open Biology* 8.9 (2018), p. 180120.

[32] B. Costanza, I. Umelo, J. Bellier, V. Castronovo, and A. Turtoi. "Stromal Modulators of TGF-β in Cancer". *Journal of Clinical Medicine* 6.1 (2017), p. 7.

[33] P. Cammareri, D. F. Vincent, M. C. Hodder, R. A. Ridgway, C. Murgia, M. Nobis, A. D. Campbell, J. Varga, D. J. Huels, C. Subramani, K. L. H. Prescott, C. Nixon, A. Hedley, S. T. Barry, F. R. Greten, G. J. Inman, and O. J. Sansom. "TGFβ pathway limits dedifferentiation following WNT and MAPK pathway activation to suppress intestinal tumourigenesis". *Cell Death & Differentiation* 24.10 (2017), pp. 1681–1693.

[34] S. Ihara, Y. Hirata, and K. Koike. "TGF-β in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota". *Journal of Gastroenterology* 52.7 (2017), pp. 777–787.

[35] J. Kempski, L. Brockmann, N. Gagliani, and S. Huber. "TH17 Cell and Epithelial Cell Crosstalk during Inflammatory Bowel Disease and Carcinogenesis". *Frontiers in Immunology* 8 (2017), p. 1373.

[36] R. Kopan. "Notch Signaling". *Cold Spring Harbor Perspectives in Biology* 4.10 (2012), a011213.

[37] E. R. Andersson, R. Sandberg, and U. Lendahl. "Notch signaling: simplicity in design, versatility in function". *Development* 138.17 (2011), pp. 3593–3612.

[38] S. Fre, M. Huyghe, P. Mourikis, S. Robine, D. Louvard, and S. Artavanis-Tsakonas. "Notch signals control the fate of immature progenitor cells in the intestine". *Nature* 435.7044 (2005), pp. 964–968.

[39] V. Zecchini, R. Domaschenz, D. Winton, and P. Jones. "Notch signaling regulates the differentiation of post-mitotic intestinal epithelial cells". *Genes & development* 19.14 (2005), pp. 1686–1691.

[40] K. L. VanDussen, A. J. Carulli, T. M. Keeley, S. R. Patel, B. J. Puthoff, S. T. Magness, I. T. Tran, I. Maillard, C. Siebel, A. Kolterud, A. S. Grosse, D. L. Gumucio, S. A. Ernst, Y.-H. Tsai, P. J. Dempsey, and L. C. Samuelson. "Notch signaling modulates proliferation and differentiation of intestinal crypt base columnar stem cells". *Development* 139.3 (2012), pp. 488–497.

[41] M. E. Keir, T. Yi, T. T. Lu, and N. Ghilardi. "The role of IL-22 in intestinal health and disease". *Journal of Experimental Medicine* 217.3 (2020), e20192195.

[42] C. B. Westphalen, S. Asfaha, Y. Hayakawa, Y. Takemoto, D. J. Lukin, A. H. Nuber, A. Brandtner, W. Setlik, H. Remotti, A. Muley, X. Chen, R. May, C. W. Houchen, J. G. Fox, M. D. Gershon, M. Quante, and T. C. Wang. "Long-lived intestinal tuft

cells serve as colon cancer–initiating cells". *Journal of Clinical Investigation* 124.3 (2014), pp. 1283–1295.

[43] C. Andrews, M. H. McLean, and S. K. Durum. "Cytokine Tuning of Intestinal Epithelial Function". *Frontiers in Immunology* 9 (2018), p. 1270.

[44] J. Qiu, J. J. Heller, X. Guo, Z.-m. E. Chen, K. Fish, Y.-X. Fu, and L. Zhou. "The Aryl Hydrocarbon Receptor Regulates Gut Immunity through Modulation of Innate Lymphoid Cells". *Immunity* 36.1 (2012), pp. 92–104.

[45] A. Metidji, S. Omenetti, S. Crotta, Y. Li, E. Nye, E. Ross, V. Li, M. R. Maradana, C. Schiering, and B. Stockinger. "The Environmental Sensor AHR Protects from Inflammatory Damage by Maintaining Intestinal Stem Cell Homeostasis and Barrier Integrity". *Immunity* 49.2 (2018), pp. 353–362.

[46] S. Brand, F. Beigel, T. Olszak, K. Zitzmann, S. T. Eichhorst, J.-m. Otte, H. Diepolder, A. Marquardt, W. Jagla, A. Popp, S. Leclair, K. Herrmann, J. Seiderer, T. Ochsenkühn, B. Göke, C. J. Auernhammer, and J. Dambacher. "IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration". *American Journal of Physiology-Gastrointestinal and Liver Physiology* 290.4 (2006), G827–G838.

[47] E. R. Kim and D. K. Chang. "Colorectal cancer in inflammatory bowel disease: the risk, pathogenesis, prevention and diagnosis". *World journal of gastroenterology* 20.29 (2014), pp. 9872–9881.

[48] O. B. Parks, D. A. Pociask, Z. Hodzic, J. K. Kolls, and M. Good. *Interleukin-22 Signaling in the Regulation of Intestinal Health and Disease*. 2016.

[49] K. Gronke, P. P. Hernández, J. Zimmermann, C. S. N. Klose, M. Kofoed-Branzk, F. Guendel, M. Witkowski, C. Tizian, L. Amann, F. Schumacher, H. Glatt, A. Triantafyllopoulou, and A. Diefenbach. "Interleukin-22 protects intestinal stem cells against genotoxic stress". *Nature* 566.7743 (2019), pp. 249–253.

[50] E. Jaeger, S. Leedham, A. Lewis, S. Segditsas, M. Becker, P. R. Cuadrado, H. Davis, K. Kaur, K. Heinimann, K. Howarth, J. East, J. Taylor, H. Thomas, and I. Tomlinson. "Hereditary mixed polyposis syndrome is caused by a 40-kb upstream duplication that leads to increased and ectopic expression of the BMP antagonist GREM1". *Nature Genetics* 44.6 (2012), pp. 699–703.

[51] H. Davis, S. Irshad, M. Bansal, H. Rafferty, T. Boitsova, C. Bardella, E. Jaeger, A. Lewis, L. Freeman-Mills, F. C. Giner, P. Rodenas-Cuadrado, S. Mallappa, S. Clark, H. Thomas, R. Jeffery, R. Poulsom, M. Rodriguez-Justo, M. Novelli, R. Chetty, A. Silver, O. J. Sansom, F. R. Greten, L. M. Wang, J. E. East, I. Tomlinson, and S. J. Leedham. "Aberrant epithelial GREM1 expression initiates colonic tumorigenesis from cells outside the stem cell niche". *Nature Medicine* 21.1 (2015), pp. 62–70.

[52] B. Chassaing, J. D. Aitken, M. Malleshappa, and M. Vijay-Kumar. "Dextran Sulfate Sodium (DSS)-Induced Colitis in Mice". *Current Protocols in Immunology* 104.1 (2014), pp. 1–14.

[53] C. Neufert, C. Becker, and M. F. Neurath. "An inducible mouse model of colon carcinogenesis for the analysis of sporadic and inflammation-driven tumor progression". *Nature Protocols* 2.8 (2007), pp. 1998–2004.

[54] A. I. Thaker, A. Shaker, M. S. Rao, and M. A. Ciorba. "Modeling colitis-associated cancer with azoxymethane (AOM) and dextran sulfate sodium (DSS)". *Journal of visualized experiments : JoVE* 67 (2012), p. 4100.

[55] J. G. Fox, Z. Ge, M. T. Whary, S. E. Erdman, and B. H. Horwitz. "Helicobacter hepaticus infection in mice: models for understanding lower bowel inflammation and cancer". *Mucosal immunology* 4.1 (2011), pp. 22–30.

[56] M. C. Kullberg, J. M. Ward, P. L. Gorelick, P. Caspar, S. Hieny, A. Cheever, D. Jankovic, and A. Sher. "Helicobacter hepaticus triggers colitis in specific-pathogen-free interleukin-10 (IL-10)-deficient mice through an IL-12- and gamma interferon-dependent mechanism." *Infection and immunity* 66.11 (1998), pp. 5157–5166.

[57] S. Kirchberger, D. J. Royston, O. Boulard, E. Thornton, F. Franchini, R. L. Szabady, O. Harrison, and F. Powrie. "Innate lymphoid cells sustain colon cancer through production of interleukin-22 in a mouse model". *The Journal of Experimental Medicine* 210.5 (2013), pp. 917–931.

[58] M. De Robertis, E. Massi, M. L. Poeta, S. Carotti, S. Morini, L. Cecchetelli, E. Signori, and V. M. Fazio. "The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies". *Journal of carcinogenesis* 10 (2011), p. 9.

[59] C. Lopez-Garcia, A. M. Klein, B. D. Simons, and D. J. Winton. "Intestinal stem cell replacement follows a pattern of neutral drift." *Science (New York, N.Y.)* 330.6005 (2010), pp. 822–825.

[60] H. J. Snippert, L. G. Van Der Flier, T. Sato, J. H. Van Es, M. Van Den Born, C. Kroon-Veenboer, N. Barker, A. M. Klein, J. Van Rheenen, and B. D. Simons. "Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells". *Cell* 143.1 (2010), pp. 134–144.

[61] L. Ritsma, S. I. J. Ellenbroek, A. Zomer, H. J. Snippert, F. J. de Sauvage, B. D. Simons, H. Clevers, and J. van Rheenen. "Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging". *Nature* 507.7492 (2014), pp. 362–365.

[62] H. J. Snippert, A. G. Schepers, J. H. van Es, B. D. Simons, and H. Clevers. "Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion". *EMBO reports* 15.1 (2014), pp. 62–69.

[63] L. Vermeulen, E. Morrissey, M. van der Heijden, A. M. Nicholson, A. Sottoriva, S. Buczacki, R. Kemp, S. Tavare, and D. J. Winton. "Defining Stem Cell Dynamics in Models of Intestinal Tumor Initiation". *Science* 342.6161 (2013), pp. 995–998.

[64] J. D'Orazio, S. Jarrett, A. Amaro-Ortiz, and T. Scott. "UV radiation and the skin". *International journal of molecular sciences* 14.6 (2013), pp. 12222–12248.

[65] T. Sugimura. "Nutrition and dietary carcinogens ". *Carcinogenesis* 21.3 (2000), pp. 387–395.

[66] P. Reynolds, S. Cooper, M. Lomax, and P. O'Neill. "Disruption of PARP1 function inhibits base excision repair of a sub-set of DNA lesions." *Nucleic acids research* 43.8 (2015), pp. 4028–4038.

[67] L. Nemzow, A. Lubin, L. Zhang, and F. Gong. "XPC: Going where no DNA damage sensor has gone before". *DNA Repair* 36 (2015), pp. 19–27.

[68] D. Gunz, M. T. Hess, and H. Naegeli. "Recognition of DNA Adducts by Human Nucleotide Excision Repair: EVIDENCE FOR A THERMODYNAMIC PROBING MECHANISM ". *Journal of Biological Chemistry* 271.41 (1996), pp. 25089–25098.

[69] A. Kinner, W. Wu, C. Staudt, and G. Iliakis. "Gamma-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin". *Nucleic acids research* 36.17 (2008), pp. 5678–5694.

[70] M. Podhorecka, A. Skladanowski, and P. Bozko. "H2AX phosphorylation: Its role in DNA damage response and cancer therapy". *Journal of Nucleic Acids* 2010 (2010).

[71] W.-D. Heyer, K. T. Ehmsen, and J. Liu. "Regulation of Homologous Recombination in Eukaryotes". *Annual Review of Genetics* 44.1 (2010), pp. 113–139.

[72] T. Sugiyama, E. M. Zaitseva, and S. C. Kowalczykowski. "A Single-stranded DNA-binding Protein Is Needed for Efficient Presynaptic Complex Formation by the Saccharomyces cerevisiae Rad51 Protein". *Journal of Biological Chemistry* 272.12 (1997), pp. 7940–7945.

[73] D. R. Whelan, W. T. C. Lee, Y. Yin, D. M. Ofri, K. Bermudez-Hernandez, S. Keegan, D. Fenyo, and E. Rothenberg. "Spatiotemporal dynamics of homologous recombination repair at single collapsed replication forks". *Nature Communications* 9.1 (2018), p. 3882.

[74] A. Nussenzweig and M. C. Nussenzweig. "A Backup DNA Repair Pathway Moves to the Forefront". *Cell* 131.2 (2007), pp. 223–225.

[75] C. W. Anderson and T. H. Carter. "The DNA-activated protein kinase—DNA-PK". *Molecular Analysis of DNA Rearrangements in the Immune System*. Springer, 1996, pp. 91–111.

[76] L. Deriano and D. B. Roth. "Modernizing the Nonhomologous End-Joining Repertoire: Alternative and Classical NHEJ Share the Stage". *Annual Review of Genetics* 47.1 (2013), pp. 433–455.

[77] M. R. Lieber. "The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway". *Annual review of biochemistry* 79 (2010), pp. 181–211.

[78] P. Frit, N. Barboule, Y. Yuan, D. Gomez, and P. Calsou. "Alternative end-joining pathway(s): Bricolage at DNA breaks". *DNA Repair* 17 (2014), pp. 81–97.

[79] L. Li, C. Robert, and F. V. "The Role of Error-Prone Alternative Non-Homologous End-Joining in Genomic Instability in Cancer". *DNA Repair and Human Health*. Ed. by C. Robert. Rijeka: InTech, 2011, Ch. 4.

[80] R. Dueva and G. Iliakis. "Alternative pathways of non-homologous end joining (NHEJ) in genomic instability and cancer". *Translational cancer research* 2.3 (2013).

[81] S. Sharma, S. M. Javadekar, M. Pandey, M. Srivastava, R. Kumari, and S. C. Raghavan. "Homology and enzymatic requirements of microhomology-dependent alternative end joining". *Cell Death & Disease* 6.3 (2015), e1697.

[82] S. B. Thyme and A. F. Schier. "Polq-Mediated End Joining Is Essential for Surviving DNA Double-Strand Breaks during Early Zebrafish Development". *Cell reports* 15.4 (2016), pp. 707–714.

[83] P. A. Mateos-Gomez, T. Kent, S. K. Deng, S. McDevitt, E. Kashkina, T. M. Hoang, R. T. Pomerantz, and A. Sfeir. "The helicase domain of Polθ counteracts RPA to promote alt-NHEJ". *Nature structural & molecular biology* 24.12 (2017), pp. 1116–1123.

[84] A. Sallmyr and A. E. Tomkinson. "Repair of DNA double-strand breaks by mammalian alternative end-joining pathways". *Journal of Biological Chemistry* 293.27 (2018), pp. 10536–10549.

[85] H. H. Y. Chang, N. R. Pannunzio, N. Adachi, and M. R. Lieber. "Non-homologous DNA end joining and alternative pathways to double-strand break repair". *Nature Reviews Molecular Cell Biology* 18.8 (2017), pp. 495–506.

[86] I. Brandsma and D. C. Gent. "Pathway choice in DNA double strand break repair: observations of a balancing act". *Genome Integrity* 3.1 (2012), p. 9.

[87] R. A. Shamanna, H. Lu, J. K. De Freitas, J. Tian, D. L. Croteau, and V. A. Bohr. "WRN regulates pathway choice between classical and alternative non-homologous end joining". *Nature Communications* 7.May (2016), pp. 1–12.

[88] D. S. Krasner, J. M. Daley, P. Sung, and H. Niu. "Interplay between Ku and Replication Protein A in the Restriction of Exo1-mediated DNA Break End Resection". *The Journal of biological chemistry* 290.30 (2015), pp. 18806–18816.

[89] G. Hua, T. H. Thin, R. Feldman, A. Haimovitz-Friedman, H. Clevers, Z. Fuks, and R. Kolesnick. "Crypt base columnar stem cells in small intestines of mice are radioresistant". *Gastroenterology* 143.5 (2012), pp. 1266–1276.

[90] A. Merlos-Suarez, F. M. Barriga, P. Jung, M. Iglesias, M. V. Cespedes, D. Rossell, M. Sevillano, X. Hernando-Momblona, V. da Silva-Diz, P. Munoz, H. Clevers, E. Sancho, R. Mangues, and E. Batlle. "The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse." *Cell stem cell* 8.5 (2011), pp. 511–524.

[91] C. Booth, G. Tudor, J. Tudor, B. P. Katz, and T. MacVittie. "The acute gastrointestinal syndrome in high-dose irradiated mice". *Health physics* 103.4 (2012), p. 383.

[92] H. Tian, B. Biehs, S. Warming, K. G. Leong, L. Rangell, O. D. Klein, and F. J. de Sauvage. "A reserve stem cell population in small intestine renders Lgr5-positive cells dispensable". *Nature* 478.7368 (2011), pp. 255–259.

[93] S. Asfaha, Y. Hayakawa, A. Muley, S. Stokes, T. A. Graham, R. E. Ericksen, C. B. Westphalen, J. Von Burstin, T. L. Mastracci, D. L. Worthley, C. Guha, M. Quante, A. K. Rustgi, and T. C. Wang. "Krt19+/Lgr5-Cells Are Radioresistant Cancer-Initiating Stem Cells in the Colon and Intestine". *Cell Stem Cell* 16.6 (2015), pp. 627–638.

[94] K. C. Roche, A. D. Gracz, X. F. Liu, V. Newton, H. Akiyama, and S. T. Magness. "SOX9 maintains reserve stem cells and preserves radioresistance in mouse small intestine". *Gastroenterology* 149.6 (2015), 1553–1563.e10.

[95] K. S. Yan, L. A. Chia, X. Li, A. Ootani, J. Su, J. Y. Lee, N. Su, Y. Luo, S. C. Heilshorn, M. R. Amieva, E. Sangiorgi, M. R. Capecchi, and C. J. Kuo. "The intestinal stem cell markers Bmi1 and Lgr5 identify two functionally distinct populations". *Proceedings of the National Academy of Sciences of the United States of America* 109.2 (2012), pp. 466–471.

[96] G. Hua, C. Wang, Y. Pan, Z. Zeng, S. G. Lee, M. L. Martin, A. Haimovitz-Friedman, Z. Fuks, P. B. Paty, and R. Kolesnick. "Distinct levels of

radioresistance in Lgr5+ colonic epithelial stem cells versus Lgr5+ small intestinal stem cells". *Cancer Research* 77.8 (2017), pp. 2124–2133.

[97] S. Umar. "Intestinal stem cells". *Current gastroenterology reports* 12.5 (2010), pp. 340–348.

[98] S. Nowsheen and E. S. Yang. "The intersection between DNA damage response and cell death pathways". *Experimental Oncology* 34.3 (2012), pp. 243–254.

[99] F. Murray-Zmijewski, E. A. Slee, and X. Lu. "A complex barcode underlies the heterogeneous response of p53 to stress." *Nature reviews. Molecular cell biology* 9.9 (2008), pp. 702–712.

[100] C.-L. Lee, J. M. Blum, and D. G. Kirsch. "Role of p53 in regulating tissue response to radiation by mechanisms independent of apoptosis". *Translational cancer research* 2.5 (2013), pp. 412–421.

[101] G. Lozano. "Mouse Models of p53 Functions". *Cold Spring Harbor Perspectives in Biology* 2.4 (2010).

[102] M. H. Barcellos-Hoff and F. A. Cucinotta. "New tricks for an old fox: Impact of TGF on the DNA damage response and genomic stability". *Science Signaling* 7.341 (2014).

[103] Y. Li, Y. Liu, Y. J. Chiang, F. Huang, Y. Li, X. Li, Y. Ning, W. Zhang, H. Deng, and Y.-G. Chen. "DNA Damage Activates TGF-b; Signaling via ATM-c-Cbl-Mediated Stabilization of the Type II Receptor TgfbRII". *Cell Reports* 28.3 (2019), pp. 735–745.

[104] Q. Liu, L. Ma, T. Jones, L. Palomero, M. A. Pujana, H. Martinez-Ruiz, P. K. Ha, J. Murnane, I. Cuartas, J. Seoane, M. Baumann, A. Linge, and M. H. Barcellos-Hoff. "Subjugation of TGFβ Signaling by Human Papilloma Virus in Head and Neck Squamous Cell Carcinoma Shifts DNA Repair from Homologous Recombination to Alternative End Joining". *Clinical Cancer Research* 24.23 (2018), pp. 6001–6014.

[105] S. Dambal, M. Shah, B. Mihelich, and L. Nonn. "The microRNA-183 cluster: the family that plays together stays together". *Nucleic acids research* 43.15 (2015), pp. 7173–7188.

[106] J. Yu, R. Lei, X. Zhuang, X. Li, G. Li, S. Lev, M. F. Segura, X. Zhang, and G. Hu. "MicroRNA-182 targets SMAD7 to potentiate TGFβ-induced epithelial-mesenchymal transition and metastasis of cancer cells". *Nature Communications* 7.1 (2016), p. 13884.

[107] H. Martinez-Ruiz, I. Illa-Bochaca, C. Omene, D. Hanniford, Q. Liu, E. Hernando, and M. H. Barcellos-Hoff. "A TGF -miR-182-BRCA1 axis controls the mammary differentiation hierarchy". *Science Signaling* 9.457 (2016).

[108] J. F. L. Chau, D. Jia, Z. Wang, Z. Liu, Y. Hu, X. Zhang, H. Jia, K. P. Lai, W. F. Leong, B. J. Au, Y. Mishina, Y. G. Chen, C. Biondi, E. Robertson, D. Xie, H. Liu, L. He, X. Wang, Q. Yu, and B. Li. "A crucial role for bone morphogenetic protein-Smad1 signalling in the DNA damage response". *Nature Communications* 3.May (2012), pp. 1–11.

[109] S. Vattulainen-Collanus, M. Southwood, X. D. Yang, S. Moore, P. Ghatpande, N. W. Morrell, G. Lagna, and A. Hata. "Bone morphogenetic protein signaling is required for RAD51-mediated maintenance of genome integrity in vascular endothelial cells". *Communications Biology* 1.1 (2018), p. 149.

[110] Y. Wang, J.-w. Huang, P. Calses, C. J. Kemp, and T. Taniguchi. "MiR-96 downregulates REV1 and RAD51 to promote cellular sensitivity to cisplatin and PARP inhibition". *Cancer research* 72.16 (2012), pp. 4037–4046.

[111] S. Kim, A. Hata, and H. Kang. "Down-regulation of mir-96 by bone morphogenetic protein signaling is critical for vascular smooth muscle cell phenotype modulation". *Journal of Cellular Biochemistry* 115.5 (2014), pp. 889–895.

[112] E. R. Fearon and B. Vogelstein. "A genetic model for colorectal tumorigenesis." *Cell* 61.5 (1990), pp. 759–767.

[113] L. N. Kwong and W. F. Dove. "APC and its modifiers in colon cancer". *Advances in experimental medicine and biology* 656 (2009), pp. 85–106.

[114] E. R. Fearon. "Molecular Genetics of Colorectal Cancer". *Annual Review of Pathology: Mechanisms of Disease* 6.1 (2011), pp. 479–507.

[115] S. D. Markowitz and M. M. Bertagnolli. "Molecular Basis of Colorectal Cancer". *New England Journal of Medicine* 361.25 (2009), pp. 2449–2460.

[116] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries". *CA: A Cancer Journal for Clinicians* 68.6 (2018), pp. 394–424.

[117] N. Barker, R. A. Ridgway, J. H. van Es, M. van de Wetering, H. Begthel, M. van den Born, E. Danenberg, A. R. Clarke, O. J. Sansom, and H. Clevers. "Crypt stem cells as the cells-of-origin of intestinal cancer". *Nature* 457.7229 (2009), pp. 608–611.

[118] L. Zhu, P. Gibson, D. S. Currle, Y. Tong, R. J. Richardson, I. T. Bayazitov, H. Poppleton, S. Zakharenko, D. W. Ellison, and R. J. Gilbertson. "Prominin 1 marks intestinal stem cells that are susceptible to neoplastic transformation". *Nature* 457.7229 (2009), pp. 603–607.

[119] Y. Nakanishi, H. Seno, A. Fukuoka, T. Ueo, Y. Yamaga, T. Maruno, N. Nakanishi, K. Kanda, H. Komekado, M. Kawada, A. Isomura, K. Kawada, Y. Sakai, M. Yanagita, R. Kageyama, Y. Kawaguchi, M. M. Taketo, S. Yonehara, and T. Chiba. "Dclk1 distinguishes between tumor and normal stem cells in the intestine". *Nature Genetics* 45.1 (2013), pp. 98–103.

[120] S. Schwitalla, A. A. Fingerle, P. Cammareri, T. Nebelsiek, S. I. Göktuna, P. K. Ziegler, O. Canli, J. Heijmans, D. J. Huels, G. Moreaux, R. A. Rupec, M. Gerhard, R. Schmid, N. Barker, H. Clevers, R. Lang, J. Neumann, T. Kirchner, M. M. Taketo, G. R. Van Den Brink, O. J. Sansom, M. C. Arkan, and F. R. Greten. "Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties". *Cell* 152.1-2 (2013), pp. 25–38.

[121] E. Torlakovic, E. Skovlund, D. C. Snover, G. Torlakovic, and J. M. Nesland. "Morphologic Reappraisal of Serrated Colorectal Polyps". *The American Journal of Surgical Pathology* 27.1 (2003).

[122] M. L. Bettington, N. I. Walker, C. Rosty, I. S. Brown, A. D. Clouston, D. M. McKeone, S.-A. A. Pearson, K. Klein, B. A. Leggett, and V. L. J. Whitehall. "A clinicopathological and molecular analysis of 200 traditional serrated adenomas". *Modern Pathology* 28.3 (2015), pp. 414–427. arXiv: `arXiv:1011.1669v3`.

[123] R. Erichsen, J. A. Baron, S. J. Hamilton-Dutoit, D. C. Snover, E. E. Torlakovic, L. Pedersen, T. Frøslev, M. Vyberg, S. R. Hamilton, and H. T. Sørensen. "Increased Risk of Colorectal Cancer Development Among Patients With Serrated Polyps". *Gastroenterology* 150.4 (2016), 895–902.e5.

[124] S. Sekine, S. Yamashita, T. Tanabe, T. Hashimoto, H. Yoshida, H. Taniguchi, M. Kojima, K. Shinmura, Y. Saito, N. Hiraoka, T. Ushijima, and A. Ochiai. "Frequent PTPRK-RSPO3 fusions and RNF43 mutations in colorectal traditional serrated adenoma". *Journal of Pathology* 239.2 (2016), pp. 133–138.

[125] L. B. Alexandrov and M. R. Stratton. "Mutational signatures: the patterns of somatic mutations hidden in cancer genomes". *Current opinion in genetics & development* 24.100 (2014), pp. 52–60.

[126] J. Guinney, R. Dienstmann, X. Wang, A. De Reyniès, A. Schlicker, C. Soneson, L. Marisa, P. Roepman, G. Nyamundanda, P. Angelino, B. M. Bot, J. S. Morris, I. M. Simon, S. Gerster, E. Fessler, F. De Sousa .E Melo, E. Missiaglia, H. Ramay, D. Barras, K. Homicsko, D. Maru, G. C. Manyam, B. Broom, V. Boige, B. Perez-Villamil, T. Laderas, R. Salazar, J. W. Gray, D. Hanahan, J. Tabernero, R. Bernards, S. H. Friend, P. Laurent-Puig, J. P. Medema, A. Sadanandam, L. Wessels, M. Delorenzi, S. Kopetz, L. Vermeulen, and S. Tejpar. "The consensus molecular subtypes of colorectal cancer". *Nature Medicine* 21.11 (2015), pp. 1350–1356.

[127] I. Martincorena, A. Roshan, M. Gerstung, P. Ellis, P. V. Loo, S. McLaren, D. C. Wedge, A. Fullam, L. B. Alexandrov, J. M. Tubio, L. Stebbings, A. Menzies, S. Widaa, M. R. Stratton, P. H. Jones, and P. J. Campbell. "High burden and pervasive positive selection of somatic mutations in normal human skin". *Science* 348.6237 (2015), pp. 880–886.

[128] S. Biswas, H. Davis, S. Irshad, T. Sandberg, D. Worthley, and S. Leedham. "Microenvironmental control of stem cell fate in intestinal homeostasis and disease". *Journal of Pathology* 237.2 (2015), pp. 135–145.

[129] K. J. Maloy, L. Salaun, R. Cahill, G. Dougan, N. J. Saunders, and F. Powrie. "CD4 + CD25 + T R Cells Suppress Innate Immune Pathology Through Cytokine-dependent Mechanisms". *The Journal of Experimental Medicine* 197.1 (2003), pp. 111–119.

[130] C. M. Nagamine, A. B. Rogers, J. G. Fox, and D. B. Schauer. "Helicobacter hepaticus promotes azoxymethane-initiated colon tumorigenesis in BALB/c-IL10-deficient mice". *International Journal of Cancer* 122.4 (2008), pp. 832–838.

[131] R. Rudling, A. B. Hassan, J. Kitau, N. Mandir, and R. A. Goodlad. "A simple device to rapidly prepare whole mounts of murine intestine". *Cell Proliferation* 39.5 (2006), pp. 415–420.

[132] J. B. Hiatt, C. C. Pritchard, S. J. Salipante, B. J. O'Roak, and J. Shendure. "Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation". *Genome Research* 23.5 (2013), pp. 843–854.

[133] T. Cancer and G. Atlas. "Comprehensive molecular characterization of human colon and rectal cancer". *Nature* 487.7407 (2012), pp. 330–337.

[134] P. Shen, W. Wang, S. Krishnakumar, C. Palm, A.-K. Chi, G. M. Enns, R. W. Davis, T. P. Speed, M. N. Mindrinos, and C. Scharfe. "High-quality DNA sequence capture of 524 disease candidate genes". *Proceedings of the National Academy of Sciences* 108.16 (2011), pp. 6549–6554.

[135] H. Li and R. Durbin. "Fast and accurate long-read alignment with Burrows–Wheeler transform". *Bioinformatics* 26.5 (2010), pp. 589–595.

[136] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. "The Sequence Alignment/Map format and SAMtools". *Bioinformatics* 25.16 (2009), pp. 2078–2079.

[137] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. "The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data". *Genome Research* 20.9 (2010), pp. 1297–1303.

[138] M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. "A framework for variation discovery and genotyping using next-generation DNA sequencing data". *Nature Genetics* 43.5 (2011), pp. 491–498.

[139] P. Cingolani, A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. "A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff". *Fly* 6.2 (2012), pp. 80–92.

[140] F. Blokzijl, R. Janssen, R. van Boxtel, and E. Cuppen. "MutationalPatterns: comprehensive genome-wide analysis of mutational processes". *Genome Medicine* 10.1 (2018), p. 33.

[141] R. Rosenthal, N. McGranahan, J. Herrero, B. S. Taylor, and C. Swanton. "deconstructSigs: delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution". *Genome Biology* 17.1 (2016), p. 31.

[142] L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, S. A. Aparicio, S. Behjati, A. V. Biankin, G. R. Bignell, N. Bolli, A. Borg, A. L. Børresen-Dale, S. Boyault, B. Burkhardt, A. P. Butler, C. Caldas, H. R. Davies, C. Desmedt, R. Eils, J. E. Eyfjörd, J. A. Foekens, M. Greaves, F. Hosoda, B. Hutter, T. Ilicic, S. Imbeaud, M. Imielinsk, N. Jäger, D. T. Jones, D. Jonas, S. Knappskog, M. Koo, S. R. Lakhani, C. López-Otín, S. Martin, N. C. Munshi, H. Nakamura, P. A. Northcott, M. Pajic, E. Papaemmanuil, A. Paradiso, J. V. Pearson, X. S. Puente, K. Raine, M. Ramakrishna, A. L. Richardson, J. Richter, P. Rosenstiel, M. Schlesner, T. N. Schumacher, P. N. Span, J. W. Teague, Y. Totoki, A. N. Tutt, R. Valdés-Mas, M. M. Van Buuren, L. Van 'T Veer, A. Vincent-Salomon, N. Waddell, L. R. Yates, J. Zucman-Rossi, P. Andrew Futreal, U. McDermott, P. Lichter, M. Meyerson, S. M. Grimmond, R. Siebert, E. Campo, T. Shibata, S. M. Pfister, P. J. Campbell, and M. R. Stratton. "Signatures of mutational processes in human cancer". *Nature* 500.7463 (2013), pp. 415–421.

[143] J. G. Tate, S. Bamford, H. C. Jubb, Z. Sondka, D. M. Beare, N. Bindal, H. Boutselakis, C. G. Cole, C. Creatore, E. Dawson, P. Fish, B. Harsha, C. Hathaway, S. C. Jupe, C. Y. Kok, K. Noble, L. Ponting, C. C. Ramshaw, C. E. Rye, H. E. Speedy, R. Stefancsik, S. L. Thompson, S. Wang, S. Ward, P. J. Campbell, and S. A. Forbes. "COSMIC: The Catalogue Of Somatic Mutations In Cancer". *Nucleic Acids Research* 47.D1 (2019), pp. D941–D947.

[144] K. J. Livak and T. D. Schmittgen. "Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2-ΔΔCT Method". *Methods* 25.4 (2001), pp. 402–408.

[145] C. S. Potten, M. Kellett, S. A. Roberts, D. A. Rew, and G. D. Wilson. "Measurement of in vivo proliferation in human colorectal mucosa using bromodeoxyuridine". *Gut* 33.1 (1992), pp. 71–78.

[146] H. Lee-Six, S. Olafsson, P. Ellis, R. J. Osborne, M. A. Sanders, L. Moore, N. Georgakopoulos, F. Torrente, A. Noorani, M. Goddard, P. Robinson, T. H. Coorens, L. O'Neill, C. Alder, J. Wang, R. C. Fitzgerald, M. Zilbauer, N. Coleman, K. Saeb-Parsy, I. Martincorena, P. J. Campbell, and M. R. Stratton. "The landscape of somatic mutation in normal colorectal epithelial cells". *Nature* 574.7779 (2019), pp. 532–537.

[147] A.-M. M. Baker, B. Cereser, S. Melton, A. G. Fletcher, M. Rodriguez-Justo, P. J. Tadrous, A. Humphries, G. Elia, S. A. C. McDonald, N. A. Wright, B. D. Simons, M. Jansen, and T. A. Graham. "Quantification of crypt and stem cell evolution in the normal and neoplastic human colon". *Cell reports* 8.4 (2014), pp. 940–947.

[148] S. Olafsson, R. E. McIntyre, T. Coorens, T. Butler, P. Robinson, H. Lee-Six, M. A. Sanders, K. Arestang, C. Dawson, M. Tripathi, K. Strongili, Y. Hooks, M. R. Stratton, M. Parkes, I. Martincorena, T. Raine, P. J. Campbell, and C. A. Anderson. "The mutational profile and clonal landscape of the inflammatory bowel disease affected colon". *bioRxiv* (2019), p. 832014.

[149] J. Bendl, M. Musil, J. Štourač, J. Zendulka, J. Damborský, and J. Brezovský. "PredictSNP2: A Unified Platform for Accurately Evaluating SNP Effects by Exploiting the Different Characteristics of Variants in Distinct Genomic Regions". *PLOS Computational Biology* 12.5 (2016), e1004962.

[150] Y. Choi and A. P. Chan. "PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels". *Bioinformatics* 31.16 (2015), pp. 2745–2747.

[151] M. Fujita, N. Matsubara, I. Matsuda, K. Maejima, A. Oosawa, T. Yamano, A. Fujimoto, M. Furuta, K. Nakano, A. Oku-Sasaki, H. Tanaka, Y. Shiraishi, R. N. Mateos, K. Nakai, S. Miyano, N. Tomita, S. Hirota, H. Ikeuchi, and H. Nakagawa. "Genomic landscape of colitis-associated cancer indicates the impact of chronic inflammation and its stratification by mutations in the Wnt signaling". *Oncotarget* 9.1 (2017), pp. 969–981.

[152] S. V. Kozlov, M. E. Graham, B. Jakob, F. Tobias, A. W. Kijas, M. Tanuji, P. Chen, P. J. Robinson, G. Taucher-Scholz, K. Suzuki, S. So, D. Chen, and M. F. Lavin. "Autophosphorylation and ATM Activation: ADDITIONAL SITES ADD TO THE COMPLEXITY ". *Journal of Biological Chemistry* 286.11 (2011), pp. 9107–9119.

[153] M. Gachechiladze, J. Škarda, A. Soltermann, and M. Joerger. "RAD51 as a potential surrogate marker for DNA repair capacity in solid malignancies". *International Journal of Cancer* 141.7 (2017), pp. 1286–1294.

[154] O. S. Gildemeister, J. M. Sage, and K. L. Knight. "Cellular redistribution of Rad51 in response to DNA damage: novel role for Rad51C". *The Journal of biological chemistry* 284.46 (2009), pp. 31945–31952.

[155] J. Greenhough, E. S. Papadakis, R. I. Cutress, P. A. Townsend, R. O. Oreffo, and R. S. Tare. "Regulation of osteoblast development by Bcl-2-associated athanogene-1 (BAG-1)". *Scientific Reports* 6 (2016), pp. 1–15.

[156] Y. Wang and T. Taniguchi. "MicroRNAs and DNA damage response: Implications for cancer therapy". *Cell Cycle* 12.1 (2013), pp. 32–42.

[157] A. J. Davis, S. So, and D. J. Chen. "Dynamics of the PI3K-like protein kinase members ATM and DNA-PKcs at DNA double strand breaks". *Cell cycle (Georgetown, Tex.)* 9.13 (2010), pp. 2529–2536.

[158] P. Ahmad, J. Sana, M. Slavik, P. Slampa, P. Smilek, and O. Slaby. "MicroRNAs Involvement in Radioresistance of Head and Neck Cancer". *Disease markers* 2017 (2017), p. 8245345.

[159] S.-M. Rapti, C. K. Kontos, I. N. Papadopoulos, and A. Scorilas. "High miR-96 levels in colorectal adenocarcinoma predict poor prognosis, particularly in patients without distant metastasis at the time of initial diagnosis". *Tumor Biology* 37.9 (2016), pp. 11815–11824.

[160] X. M. Xu, J. C. Qian, Z. L. Deng, Z. Cai, T. Tang, P. Wang, K. H. Zhang, and J.-P. Cai. "Expression of miR-21, miR-31, miR-96 and miR-135b is correlated with the clinical parameters of colorectal cancer". *Oncology letters* 4.2 (2012), pp. 339–345.

[161] A. L. Ress, V. Stiegelbauer, E. Winter, D. Schwarzenbacher, T. Kiesslich, S. Lax, S. Jahn, A. Deutsch, T. Bauernhofer, H. Ling, H. Samonigg, A. Gerger, G. Hoefler, and M. Pichler. "MiR-96-5p influences cellular growth and is associated with poor survival in colorectal cancer patients". *Molecular Carcinogenesis* 54.11 (2015), pp. 1442–1450.

[162] M. Vahabi, C. Pulito, A. Sacconi, S. Donzelli, M. D'Andrea, V. Manciocco, R. Pellini, P. Paci, G. Sanguineti, L. Strigari, G. Spriano, P. Muti, P. P. Pandolfi, S. Strano, S. Safarian, F. Ganci, and G. Blandino. "miR-96-5p targets PTEN expression affecting radio-chemosensitivity of HNSCC cells". *Journal of Experimental & Clinical Cancer Research* 38.1 (2019), p. 141.

[163] P. Guo, Y. Yu, Z. Tian, Y. Lin, Y. Qiu, W. Yao, and L. Zhang. "Upregulation of miR-96 promotes radioresistance in glioblastoma cells via targeting PDCD4." *International journal of oncology* 53.4 (2018), pp. 1591–1600.

[164] J. E. East, W. S. Atkin, A. C. Bateman, S. K. Clark, S. Dolwani, S. N. Ket, S. J. Leedham, P. S. Phull, M. D. Rutter, N. A. Shepherd, I. Tomlinson, and C. J. Rees. "British Society of Gastroenterology position statement on serrated polyps in the colon and rectum". *Gut* 66.7 (2017), 1181 LP –1196.

[165] H. A. Shihab, M. F. Rogers, J. Gough, M. Mort, D. N. Cooper, I. N. M. Day, T. R. Gaunt, and C. Campbell. "An integrative approach to predicting the

functional effects of non-coding and coding sequence variation". *Bioinformatics* 31.10 (2015), pp. 1536–1543.

[166] C. Morrison, E. Sonoda, N. Takao, A. Shinohara, K. Yamamoto, and S. Takeda. "The controlling role of ATM in homologous recombinational repair of DNA damage". *The EMBO journal* 19.3 (2000), pp. 463–471.

[167] K. Hatakeyama, K. Ohshima, T. Nagashima, S. Ohnami, S. Ohnami, M. Serizawa, Y. Shimoda, K. Maruyama, Y. Akiyama, K. Urakami, M. Kusuhara, T. Mochizuki, and K. Yamaguchi. "Molecular profiling and sequential somatic mutation shift in hypermutator tumours harbouring POLE mutations". *Scientific reports* 8.1 (2018), p. 8700.

[168] A. Maréchal and L. Zou. "DNA damage sensing by the ATM and ATR kinases". *Cold Spring Harbor perspectives in biology* 5.9 (2013), a012716.

[169] H. Lempiäinen and T. D. Halazonetis. "Emerging common themes in regulation of PIKKs and PI3Ks". *The EMBO journal* 28.20 (2009), pp. 3067–3073.

[170] Y. Zhang, A. G. Baranovskiy, T. H. Tahirov, and Y. I. Pavlov. "The C-terminal domain of the DNA polymerase catalytic subunit regulates the primase and polymerase activities of the human DNA polymerase α-primase complex". *The Journal of biological chemistry* 289.32 (2014), pp. 22021–22034.

[171] Q. Pan, X. Lou, J. Zhang, Y. Zhu, F. Li, Q. Shan, X. Chen, Y. Xie, S. Su, H. Wei, L. Lin, L. Wu, and S. Liu. "Genomic variants in mouse model induced by azoxymethane and dextran sodium sulfate improperly mimic human colorectal cancer". *Scientific reports* 7.1 (2017), p. 25.

[172] A. J. McCarthy, S. Serra, and R. Chetty. "Traditional serrated adenoma: an overview of pathology and emphasis on molecular pathogenesis". *BMJ open gastroenterology* 6.1 (2019), e000317–e000317.

[173] J. A. Lee, H. E. Park, S.-Y. Yoo, S. Jeong, N.-Y. Cho, G. H. Kang, and J. H. Kim. "CpG Island Methylation in Sessile Serrated Adenoma/Polyp of the Colorectum: Implications for Differential Diagnosis of Molecularly High-Risk Lesions among Non-dysplastic Sessile Serrated Adenomas/Polyps". *Journal of pathology and translational medicine* 53.4 (2019), pp. 225–235.

[174] P. Mombaerts, J. Iacomini, R. S. Johnson, K. Herrup, S. Tonegawa, and V. E. Papaioannou. "RAG-1-deficient mice have no mature B and T lymphocytes." *Cell* 68.5 (1992), pp. 869–877.

[175] O. Boulard, S. Kirchberger, D. J. Royston, K. J. Maloy, and F. M. Powrie. "Identification of a genetic locus controlling bacteria-driven colitis and associated cancer through effects on innate inflammation." *The Journal of experimental medicine* 209.7 (2012), pp. 1309–1324.

[176] J. A. Eaden, K. R. Abrams, and J. F. Mayberry. "The risk of colorectal cancer in ulcerative colitis: a meta-analysis." *Gut* 48.4 (2001), pp. 526–535.

[177] S. Grivennikov, E. Karin, J. Terzic, D. Mucida, G.-Y. Yu, S. Vallabhapurapu, J. Scheller, S. Rose-John, H. Cheroutre, L. Eckmann, and M. Karin. "IL-6 and Stat3 are required for survival of intestinal epithelial cells and development of colitis-associated cancer." *Cancer cell* 15.2 (2009), pp. 103–113.

[178] J.-M. Zha, H.-S. Li, Q. Lin, W.-T. Kuo, Z.-H. Jiang, P.-Y. Tsai, N. Ding, J. Wu, S.-F. Xu, Y.-T. Wang, J. Pan, X.-M. Zhou, K. Chen, M. Tao, M. A. Odenwald, A. Tamura, S. Tsukita, J. R. Turner, and W.-Q. He. "Interleukin 22 Expands Transit-Amplifying Cells While Depleting Lgr5(+) Stem Cells via Inhibition of Wnt and Notch Signaling". *Cellular and molecular gastroenterology and hepatology* 7.2 (2019), pp. 255–274.

[179] C. C. Maley, A. Aktipis, T. A. Graham, A. Sottoriva, A. M. Boddy, M. Janiszewska, A. S. Silva, M. Gerlinger, Y. Yuan, K. J. Pienta, K. S. Anderson, R. Gatenby, C. Swanton, D. Posada, C.-I. Wu, J. D. Schiffman, E. S. Hwang, K. Polyak, A. R. A. Anderson, J. S. Brown, M. Greaves, and D. Shibata. "Classifying the evolutionary and ecological features of neoplasms". *Nature Reviews Cancer* 17.10 (2017), pp. 605–619.

[180] N. Huyghe, P. Baldin, and M. Van den Eynde. "Immunotherapy with immune checkpoint inhibitors in colorectal cancer: what is the future beyond deficient mismatch-repair tumours?" *Gastroenterology Report* 8.1 (2019), pp. 11–24.

[181] T. Hay, J. Matthews, L. Pietzka, A. Lau, A. Cranston, R. Boulter, A. Nygren, A. Douglas-Jones, G. Smith, N. Martin, M. O'Connor, and A. Clarke. "PARP-1 inhibitor monotherapy and combination therapy in a preclinical mouse model of Brca2 mutant breast cancer". *Breast Cancer Research : BCR* 10.Suppl 2 (2008), P95–P95.