

Correlative and functional analysis of genomic instability and immune response in colorectal and endometrial cancer

Submitted for the degree of Doctor of Philosophy

Mark Glaire

Green Templeton College
University of Oxford
Trinity 2021



Abstract

Cancer is a common disease, which results in considerable amounts of morbidity and mortality worldwide. It is now recognised that the tumour immune response is essential in determining cancer biology. However, anti-tumour immunity varies considerably between cancers, and even within the same cancer type. In part, it is determined by the tumour mutational landscape. But much remains to be understood about how the interaction with the cancer genome shapes the tumour immune microenvironment, and influences clinical outcomes.

In this research, I demonstrated that tumour CD8⁺ cell infiltrate is significantly associated with reduced recurrence risk and improved overall survival in stage II/III colorectal cancer (CRC). Importantly, this relationship was independent of key confounding factors, including genomic instabilities that have been shown to influence both anti-tumour immunity and clinical outcomes. Furthermore, this relationship was not consistent across all stage II/III CRCs, but rather varied according to tumour and nodal status. Mismatch repair deficiency (MMRd) is a form of genomic instability that contributes to the development of both CRC and endometrial cancer (EC). I showed that MMRd is associated with improved outcomes in CRC, but not EC. This might be a reflection of differences in the immune landscape between MMRd CRC and EC: MMRd CRCs had much greater levels of infiltrate T lymphocytes. Furthermore, MMRd ECs showed an enrichment for loss of function *JAK1* mutations, which was associated with abrogated IFN γ signalling activation. Finally, this research demonstrated a murine model to allow functional investigation of MMR loss and its impact on the tumour microenvironment. Conditional, inducible *Mlh1* loss resulted in a specific phenotype, with the development of intestinal polyps. In contrast to human disease, these polyps showed CD8⁺ T cell exclusion, which may be a secondary to Beta-2 Microglobulin loss.

In summary, this work constitutes a correlative and functional study of the associated between the immune response and clinical outcome, and how this is influenced by tumour genomic instability.

Acknowledgements

First, and foremost, thank you to my supervisor David. Your unfaltering patience, continual support and pertinent guidance will always be remembered. It cannot be expressed how truly grateful I am.

Thank you to Professor Emad El-Omar for inspiring a pursuit of academia; his mentorship work has had a transformational effect on many careers. Thank you also to Professor Ian Tomlinson for providing the first opportunity to work at the University of Oxford, and for the support and encouragement to undertake a DPhil.

Thank you to Wellcome Trust for providing me with the funding to carry out this DPhil. It is a truly outstanding organisation and I hope that its success continues long into the future.

Thank you to the Church (and before that Tomlinson) group as a whole: a wonderful bunch to work, and socialise, with. With special mention to Matty, for guidance on all lab-based issues; to Noel, for showing me how mouse work should be done, and to Charlotte, for all of her help (and putting up with my awful jokes). Thank you to our collaborators at the Leiden University Medical Centre, with special thank you Noel for showing me Leiden during my time there. Thank you to Neil, for help with all things endometrial, for being a good research partner and an even better friend.

Thanks to the staff and students of Green Templeton College for providing a welcoming environment in which to enjoy the non-DPhil aspects of life at the University of Oxford. Particular thanks to Allison, for many kind words and sage advice.

Thank you to Mum and Dad, who have been sources of eternal support and encouragement. Cheers to Tom and Josh for being kind enough to listen (or pretend to listen) to endless discussions of cancer and the immune response.

56	Table of Contents	
57	<i>Abstract</i>	<i>i</i>
58	<i>Acknowledgements</i>	<i>ii</i>
59	<i>Table of Figures</i>	<i>vii</i>
60	<i>Table of Tables</i>	<i>xv</i>
61	<i>Glossary of abbreviations</i>	<i>xix</i>
62	<i>Publications arising from my work</i>	<i>xxi</i>
63	1. Introduction	1
64	1.1. Colorectal cancer	3
65	1.1.1. Epidemiology.....	3
66	1.1.2. Aetiology	3
67	1.1.3. Pathogenesis	4
68	1.1.4. Clinical features	7
69	1.1.5. Management	8
70	1.2. Endometrial cancer	10
71	1.2.1. Epidemiology.....	10
72	1.2.2. Aetiology	11
73	1.2.3. Pathogenesis	11
74	1.2.4. Clinical management	12
75	1.3. Genomic instability in colorectal and endometrial cancer	15
76	1.3.1. DNA Polymerase ϵ (<i>POLE</i>) mutations.....	16
77	1.3.2. The DNA Mismatch Repair System	18
78	1.4. The immune microenvironment in cancer	18
79	1.5. Determinants of the anti-tumour immune response	25
80	1.5.1. Tumour mutational landscape and antigens	25
81	1.5.2. Immune escape	26
82	1.5.3. Host genetics.....	27
83	1.5.4. The gut microbiome.....	28
84	1.6. Clinical implications of the tumour immune microenvironment	30
85	1.7. Summary and outstanding questions	31
86	2. Methods	33
87	2.1. TCGA data analysis	33
88	2.1.1. RNAseq data	33
89	2.1.2. TCGA tumour mutational analysis.....	33
90	2.2. Cell culture	34
91	2.2.1. Cell culture conditions and media	34
92	2.3. General histology and tissue staining	34

93	2.3.1.	Tissue storage	34
94	2.3.2.	Tissue processing and sectioning.....	34
95	2.3.3.	Haematoxylin and eosin staining.....	34
96	2.3.4.	Immunohistochemistry	35
97	2.4.	SDS-PAGE electrophoresis and immunoblotting	36
98	2.4.1.	Cellular protein lysate collection	36
99	2.4.2.	Protein lysate quantification	36
100	2.4.3.	Western Blot	37
101	2.4.4.	Antibodies used Western Blot experiments.....	37
102	2.5.	Mice	38
103	2.5.1.	Mutations and background strains	38
104	2.5.2.	Animal husbandry.....	38
105	2.5.3.	DNA Extraction	38
106	2.5.4.	Genotyping	38
107	2.5.5.	Mouse Intestinal adenoma quantification	43
108	2.6.	Statistical analysis	43
109	2.6.1.	Graph preparation	43
110	3.	<i>The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome</i>	
111		<i>in stage II/III colorectal cancer.....</i>	44
112	3.1.	Introduction	44
113	3.1.1.	Previous Work	45
114	3.1.2.	Limitations of previous work.....	48
115	3.1.3.	Rationale for study and aims for chapter	49
116	3.2.	Methods.....	50
117	3.2.1.	Trial cohorts used in this study	50
118	3.2.2.	Inclusion of cases for biomarker study	50
119	3.2.3.	Determination of tumour T cell infiltrate by immunohistochemistry	52
120	3.2.4.	Tumour molecular analysis.....	52
121	3.2.5.	Validation series	53
122	3.2.6.	Statistical analysis.....	53
123	3.2.7.	Ethical approval	54
124	3.3.	Results	55
125	3.3.1.	QUASAR2 and VICTOR patient details	55
126	3.3.2.	Immune cell infiltrate	58
127	3.3.3.	Missing Data Imputation	66
128	3.3.4.	CD8 ⁺ cell infiltrate and clinical outcome	66
129	3.3.5.	Prognostic association of CD8 ⁺ cell infiltrate according to tumour and nodal risk strata.....	71
130	3.3.6.	Prognostic implication of <i>BRAF</i> mutation and genomic status on clinical outcome in the	
131		combined QUASAR2 and VICTOR population	76
132	3.3.7.	External validation of results	80
133	3.3.8.	Work planned but not completed due to COVID-19 pandemic	84
134	3.4.	Discussion.....	87
135	3.4.1.	CD8 ⁺ cell infiltrate and clinical outcome in early-stage CRC.....	87
136	3.4.2.	Strengths	88
137	3.4.3.	Limitations	89
138	3.4.4.	Future work.....	90
139	4.	<i>The consequences of DNA mismatch repair deficiency in colorectal cancer and</i>	
140		<i>endometrial cancer.....</i>	91

141	4.1. Introduction	91
142	4.1.1. The DNA Mismatch Repair System	91
143	4.1.2. Mismatch Repair Deficiency in Endometrial and Colorectal cancer	94
144	4.1.3. Mismatch repair deficiency and the tumour immune microenvironment	96
145	4.1.4. MMR status as a biomarker in cancer treatment	97
146	4.1.5. Rationale for study and aims for this chapter	99
147	4.2. Methods	100
148	4.2.1. Identification of studies for meta-analysis.....	100
149	4.2.2. Multispectral immunofluorescence quantification of T cell infiltrate	100
150	4.2.3. Image acquisition and analysis	102
151	4.2.4. TCGA analysis.....	102
152	4.2.5. Cell lines	103
153	4.2.6. Interferon-gamma stimulation	103
154	4.2.7. Western blotting.....	104
155	4.2.8. Gene Set Enrichment Analysis	104
156	4.2.9. Statistical analyses	104
157	4.3. Results	105
158	4.3.1. Mismatch repair deficiency and clinical outcome in colorectal and endometrial cancer	105
159	4.3.2. T cell infiltrate in CRC and EC	107
160	4.3.3. Validation in TCGA series	114
161	4.3.4. Immune escape mutations.....	121
162	4.3.5. <i>JAK1</i> LOF mutations are associated with reduced IFN- γ pathway activation in MSI ECs....	124
163	4.3.6. <i>JAK1</i> loss of function mutations cause protein loss and abolishes IFN- γ response <i>in vitro</i>	127
164	4.4. Discussion	129
165	4.4.1. The consequences of MMR loss in colorectal and endometrial cancer.....	129
166	4.4.2. Strengths	130
167	4.4.3. Limitations	131
168	4.4.4. Future Work.....	132
169	5. Functional analysis of mismatch repair loss in intestinal cancer	133
170	5.1. Introduction	133
171	5.1.1. The mouse as a tool for cancer research	133
172	5.1.2. Mouse models of intestinal cancer	135
173	5.1.3. Modelling tumour-immune interaction.....	138
174	5.1.4. Mouse models of mismatch repair deficiency and cancer	139
175	5.2. Rationale for study and aims of chapter	143
176	5.3. Materials and methods	144
177	5.3.1. Murine housing and breeding	144
178	5.3.2. Genotyping PCR.....	145
179	5.3.3. Animal monitoring, sacrifice, necropsy and tissue preparation	145
180	5.3.4. Experimental mice	145
181	5.3.5. Quantification of intestinal polyps.....	146
182	5.3.6. Tissue storage	146
183	5.3.7. Histological analysis	146
184	5.3.8. Immune cell quantification	146
185	5.4. Results	148
186	5.4.1. Gastrointestinal tumorigenesis in <i>Vil</i> -CreERT2, <i>Apc</i> ^{+/Δ} <i>Mlh1</i> ^{Δ/Δ} mice	148
187	5.4.2. The immune microenvironment	157
188	5.4.3. CD19 ⁺ B lymphocyte infiltration.....	165
189	5.4.4. Planned analyses not completed due to the COVID19 pandemic	168

190	5.5. Discussion.....	170
191	5.5.1. Limitations.....	172
192	5.5.2. Future work.....	172
193	6. Conclusions.....	174
194	6.1. Tumour infiltrating CD8⁺ T lymphocytes are independently associated with improved	
195	recurrence free survival and overall survival in stage II/III colorectal cancer patients	174
196	6.2. The prognostic association of CD8⁺ T lymphocyte varies according to tumour and	
197	nodal status	175
198	6.3. The consequences of mismatch repair deficiency vary between endometrial and	
199	colorectal cancer	175
200	6.4. A murine model of <i>Mlh1</i> loss is associated with intestinal tumorigenesis	176
201	6.5. CD8⁺ T cell exclusion and <i>B2m</i> loss are common features of mouse intestinal	
202	adenomas.....	177
203	6.6. Future Perspectives	177
204	7. Appendices	179
205	7.1. Gene set enrichment analysis results.....	179
206	7.1.1. Top 30 upregulated gene sets for MSI CRC vs MSS CRC	179
207	7.1.2. Top 30 downregulated gene sets for MSI CRC vs MSS CRC	180
208	7.1.3. Top 30 upregulated gene sets for MSI EC vs MSS EC	181
209	7.1.4. Top 30 downregulated gene sets for MSI EC vs MSS EC	182
210	8. References.....	183
211		
212		

Table of Figures

Figure 1-1. Cancer is a leading cause of premature death (death of those below age 70 years) in 112 countries worldwide. Data from the World Health Organisation and figure reproduced from GLOBOCAN 2020 report. [5]	vii
Figure 1-2: Normal epithelium to polyp to colorectal cancer sequence. Colorectal cancers result from stepwise neoplastic transformation of normal epithelial cells. There are two well defined pathways: the classic adenoma-carcinoma sequence, and the serrated neoplasia pathway. Whilst each of these are characterised by distinct genetic alterations, they share dysregulation of similar signalling pathways. APC: adenomatous polyposis coli. SMAD4: SMAD family member 4. CTNNB1: catenin- β 1. PIK3CA: phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit- α . PTEN: phosphatase and tensin homologue. TGF β : transforming growth factor β . TGFB2: TGF β receptor 2. Reproduced from [3]	5
Figure 1-3. The immune microenvironment in cancer. Elements of the immune response can either be pro or anti-tumour. Neutrophils are typically considered to have pro-tumorigenic effects, and this may be in part from production of CXCL5. In contrast, elements of the adaptive immune response have clear anti-tumour effects. In particular, CD8 ⁺ cytotoxic T-lymphocytes (CTLs) and NK cells are able to directly destroy tumour cells. The former do so after activation by antigen presentation through the MHC class I pathway, whilst the latter are able to act independently. Tumour cells may avoid CTL destruction by upregulation of immune checkpoints such as PD-L1. T _H 1 helper cells exert anti-tumour effects by driving an IFN- γ immune based response; they are activated primarily by antigen presentation via MHC class II interaction with antigen presenting cells such as dendritic cells. B-cells act as professional antigen presenting cells, or induce humoral macrophage phagocytosis by antibody production. CTL: cytotoxic T-lymphocytes. NK cells: natural killer cells. TCR: T-cell receptor. PD-L1: programmed cell death ligand 1. PD1: programmed cell death protein 1. Figure adapted from Bruni et al [4]. Further details of the immune cells represented are available in Tables 1.5. and 1.6.	20

239	Figure 2-1. Gel electrophoresis for the products of Mlh1 floxed PCR: Mlh1 WT (wildtype) Mlh1	
240	Het (Heterozygous) and Mlh1 Hom (Hom)	40
241	Figure 2-2. Gel electrophoresis for the products of Apc flox PCR: WT (wildtype) and Het	
242	(heterozygous)	41
243	Figure 2-3. Gel electrophoresis for the products of Cre PCR: Cre + (Heterozygous Vil-CreERT2	
244	and WT (wildtype)	43
245	Figure 3-1. CONSORT diagram showing the inclusion and exclusion of CRC cases from	
246	QUASAR2 and VICTOR trials. RT: radiotherapy. TMA: tissue microarray (reproduced from Glaire	
247	et al, 2019).....	56
248	Figure 3-2: T-cell infiltrate quantification. Representative image showing IHC based quantification	
249	of CD3 ⁺ (A) and CD8 ⁺ (B) in TMA sections using the ImmunoPath automated software. For analysis,	
250	each histopathological image is divided into subsections, and the software is trained to recognise	
251	marker positive and negative cells. Total number of positive and negative cells are given per unit area	
252		60
253	Figure 3-3 CD8⁺ cell infiltration in the VICTOR and QUASAR2 combined cohort. Histogram	
254	with overlaid kernel density plot showing unadjusted CD8 ⁺ cell fraction in the VICTOR and	
255	QUASAR2 combined cohort (A) and the Log(2) transformed CD8 ⁺ cell fraction with representative	
256	images (B)	61
257	Figure .3-4. The association between CD8⁺ cell infiltrate and clinical outcome in the VICTOR	
258	and QUASAR2 trials. (A). Multivariable adjusted hazard ratio for risk of recurrence and CD8 ⁺ cell	
259	density in QUASAR2 and VICTOR (B). Multivariable adjusted hazard ratio for risk of death and	
260	CD8 ⁺ cell density in QUASAR2 and VICTOR (C). Kaplan Meier curves showing risk of recurrence	
261	with patient cohort divided at a median cut point. (D). Kaplan Meier curves showing overall survival	
262	with patient cohort divided at a median cut point. Reproduced from Glaire et al 2019 [1].....	68
263	Figure 3-5. Kaplan Meier curve showing the probability of recurrence according to primary	
264	tumour and nodal stage in QUASAR2 and VICTOR cohorts. Reproduced from Glaire et al [1] .	72

265	Figure 3-6. Multivariable adjusted hazard ratio showing association of CD8+ cell infiltrate and	
266	recurrence risk divided into low risk (pT1-3), intermediate risk (pT1-3 & pN1-2, pT4 & pN0)	
267	and high risk (pT4 & pN1-2). Reproduced from Glaire et al 2019 [1]	73
268	Figure 3-7. 3 year recurrence free probability according to risk strata: low risk (pT1-3),	
269	intermediate risk (pT1-3 & pN1-2, pT4 & pN0) and high risk (pT4 & pN1-2). Reproduced from	
270	Glaire et al 2019 [1]	74
271	Figure 3-8. Kaplan Meier curve showing probability of recurrence according CD8+ infiltrate	
272	high or low and by tumour risk strata. Reproduced from Glaire et al 2019 [1].....	76
273	Figure 3-9.. Kaplan Meier curve showing the probability of recurrence according to primary	
274	tumour and nodal stage in the external validation cohort. Reproduced from Glaire et al 2019 [1]82	
275	Figure 3-10 Multivariable adjusted hazard ratio showing association of CD8A expression and	
276	recurrence risk divided into low risk (pT1-3), intermediate risk (pT1-3 & pN1-2, pT4 & pN0)	
277	and high risk (pT4 & pN1-2) in the external validation cohort. Reproduced from Glaire et al 2019	
278	[1]	83
279	Figure 3-11. Representative imagex showing 7-plex co-immunofluorescence staining. These are	
280	example images of the analysis of QUASAR2 and SCOT trial samples that have been stained using	
281	the multispectral, immunofluorescence panel detailed in Table 3.12	85
282	Figure 4-1. Diagrammatic representation of the MMR system in eukaryotic cells. The MMR	
283	system is responsbile to correcting base-base mismatches, and insert-deletion loops that are	
284	introduced during genomic replication. Loss of the MMR system results in a mutator phenotype and	
285	genomic instability known as microsatellite instability (MSI). MMR: Mismatch Repair Figure	
286	reproduced from [6]	93
287	Figure 4-2. T cell infiltrate quantification by multispectral immunofluorescence staining in full	
288	faced FFPE tumour sections. This quantification method has been previously published and validated	
289	[2]. FFPE: formalin fixed, paraffin embedded.....	101
290	Figure 4-3. Forest plot showing the meta-analysis results for the association between MMRd and	
291	risk of disease recurrence in CRC and EC. MMRd was significantly associated with a reduced risk	
292	of disease recurrence in CRC (HR 0.62 95%CI 0.55-0.69) but not in EC (HR 1.10 95%CI 0.82-1.39).	

293	MMRd: mismatch repair deficient. MMRp: mismatch repair proficient. c. 95%CI: 95% confidence	
294	intervals	106
295	Figure 4-4. Forest plot showing the meta-analysis results for the association between MMRd and	
296	overall survival in CRC and EC. MMRd was significantly associated with a reduced risk of disease	
297	recurrence in CRC (HR 0.69 95%CI 0.59-0.79) but not in EC (HR 1.33 95%CI 0.98-1.68). MMRd:	
298	mismatch repair deficient. MMRp: mismatch repair proficient. HR: hazards ratio. 95%CI: 95%	
299	confidence intervals	107
300	Figure .4-5. Multispectral immunofluorescent quantification of immune cells in FFPE	
301	CRC and EC tissue A. Shows the imaging and quantification methodology used to analyse	
302	multispectral IF stained FFPE tumour slides. B. Example images of CRC MMRp, MMRd	
303	Sporadic and MMR germline tumours. C. Example images of CRC MMRp, MMRd Sporadic	
304	and MMR germline tumours. C. Example images of EC MMRp, MMRd Sporadic and MMR	
305	germline tumours. MMRp: Mismatch repair proficient. MMRd: mismatch repair deficient. LS:	
306	Lynch Syndrome. Pan-CK: pan cytokeratin. DAPI: 4',6-diamidino-2-phenylindole	110
307	Figure 4-6. T-cell infiltrate in CRC and EC MMRp and MMRd tumours quantified by	
308	multispectral IF staining A) Intraepithelial CD3 ⁺ and CD8 ⁺ cell infiltrate. B) IntrastromalCD3 ⁺ and	
309	CD8 ⁺ cell infiltrate. Individual data points coloured according to whether MMRd was sporadic or	
310	Lynch syndrome. P values between MMR groups for each cancer type determined by non-parametric	
311	Mann-Whitney testing and P for interaction by aligned rank ANOVA, MMRp – mismatch repair	
312	proficient; MMRd – mismatch repair deficient. CRC: colorectal cancer. EC: endometrial cancer ...	111
313	Figure 4-7. Intrastromal and intraepithelial FoxP3⁺ cell infiltrate in CRC and EC MMRp and	
314	MMRd tumours; quantified by multispectral IF staining. P values between MMR groups for each	
315	cancer type determined by non-parametric Mann-Whitney testing and P for interaction by Aligned	
316	Rank ANOVA. MMRp – mismatch repair proficient; MMRd – mismatch repair deficient. CRC:	
317	colorectal cancer. EC: endometrial cancer	113
318	Figure 4-8. Heatmap showing relative tumour proportion of tumour-infiltrating immune cells,	
319	IFNγ response score and expression of immune checkpoints. The cancer type, MMR status and	
320	immune subgroup (C1-C6). MSI: microsatellite unstable. MSS: microsatellite stable. CRC: colorectal	

321	cancer EC: endometrial cancer. NK cells: Natural killer cells. The right column graphically represents	
322	the Log10 transformed P value for comparisons of CRC MSS Vs MSI (red), EC MSS. Vs MSI (blue)	
323	(both determined by Mann-Whitney U Test) and P _{INTERACTION} value (green) (aligned rank ANOVA	
324	test).....	115
325	Figure 4-9. Immune cell infiltrate in TCGA CRC and EC cancers determined by CIBERSORT	
326	processing of bulk RNAseq data. P values between MMR groups for each cancer type determined by	
327	non-parametric Mann-Whitney testing and P for interaction by Aligned Rank ANOVA. CRC:	
328	colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable.	
329	NK cells: natural killer cells.....	116
330	Figure 4-10. Immune checkpoint expression in TCGA CRC and EC RNAseq data. P values	
331	between MMR groups for each cancer type determined by non-parametric Mann-Whitney testing and	
332	P for interaction by Aligned Rank ANOVA. LAG3: Lymphocyte Activation Gene 3. VSIR: V-set	
333	Immunoregulatory Receptor. TIGIT: T-cell Immunoreceptor with Immunoglobulin and ITIM domain.	
334	HAVCR2: Hepatitis A Virus Cellular Receptor 2	117
335	Figure 4-11. Gene set enrichment analysis results comparing CRC MSI cancers with CRC MSS	
336	cancers, and EC MSI cancers with EC MSS cancers. Notably, in CRC MSI cancers a number of the	
337	upregulated pathways are related to IFN- γ signalling.. CRC: Colorectal cancer. EC: endometrial	
338	cancer. MSS: microsatellite stable. MSI: microsatellite unstable.....	119
339	Figure 4-12. The total non-silent mutation rate, indel neoantigens and SNV neoantigens for	
340	TCGA CRC and EC MSS and MSI cancer. P value for comparisons within cancers generated by	
341	non-parametric Mann-Whitney testing and P for interaction by Aligned Rank ANOVA. CRC:	
342	colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable.	
343	SNV: single nucleotide variants.....	120
344	Figure 4-13. MLH1 methylation and expression A. MLH1 promoter methylation in TCGA CRC	
345	and EC MSS and MSI cancers B. MLH1 expression in TCGA CRC and EC MSS and MSI cancers.	
346	CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite	
347	unstable	121

348	Figure 4-14. The frequency and type of immune evasion associated mutations in TCGA MSI	
349	CRC and EC. P values for comparison of mutation frequency between cancer types obtained using	
350	Fisher's exact test. CRC: Colorectal cancer. EC: endometrial cancer.....	123
351	Figure 4-15. IFN-γ and JAK1 loss of function (LOF) mutations in TCGA CRC and EC. A IFN- γ	
352	pathway response in TCGA CRC and EC MSS and MSI cancers B. Linear regression examining the	
353	association of MMR status, mutational load and immune escape mutations on IFN- γ signalling C.	
354	JAK1 expression in TCGA CRC and EC MSI cancers according to JAK1 mutational status. D IFN- γ	
355	response in TCGA CRC and EC MSI cancers according to JAK1 mutational status. P values for	
356	within cancer comparison generated by Mann-Whitney test and P for interaction by aligned rank	
357	ANOVA. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI:	
358	microsatellite unstable. fs: frameshift mutation	125
359	Figure 4-16. JAK1 mutations and IFN-γ pathway activation in EC cancer cell lines in vitro. A.	
360	JAK1 expression in EC cell lines under resting conditions. B. IFN- γ stimulation of JAK1 WT and	
361	mutant cell lines. JAK1 mutation was associated with reduced activation of the JAK-STAT pathway	
362	in response to IFN- γ	128
363	Figure 5-1. Outlines the breeding plan used to obtain experimental mice for this work. A.	
364	breeding to achieve Vil-CreERT2 ^{+/delta} Apc ^{+/fl} Mlh1 ^{fl/fl} . B. breeding to achieve Vil-	
365	CreERT2 ^{+/delta} Apc ^{+/fl} Mlh1 ^{+/fl} . C. breeding to achieve Vil-CreERT2 ^{+/delta} Apc ^{+/fl} Mlh1 ^{+/+}	144
366	Figure 5-2. H&E images of representative GI polyps seen in Apc del, Apc del/Mlh1 del and Apc	
367	del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x magnification.	149
368	Figure 5-3. Gastrointestinal (GI) tract polyps. in the Apc del, Apc del/Mlh1 del and Apc del/Mlh1	
369	null and Mlh1 null A. total polyp counts for 90, 120 and all time points for the four genotypes	
370	included in this work. B GI Polyp formation in the proximal, medial and distal intestine, and colon for	
371	each of the four genotypes. C. GI polyps in the four genotypes by sex. NS: not significant	151
372	Figure 5-4: Example of phenotype of Apc del/Mlh1 null mice at 120 days. The severity of the	
373	phenotype required culling of the mice for humane reasons. A. Macroscopically, the mice had	
374	distended and necrotic caeca and appendices B. Evidence of multiple polyps throughout the	
375	gastrointestinal tract.	152

376	Figure 5-5. H&E staining of an invasive caecal carcinoma in a Mlh1 null mouse (Apc wildtype).	
377	This mouse had to be culled early due to severe ill health, and necropsy revealed a large caecal	
378	tumour with intestinal obstruction.	153
379	Figure 5-6. Examples of IHC from CD8⁺ cells in normal intestinal mucosa of Black-6 WT and	
380	Vil-CreERT 2, mice shown at 2x, 10x and 20x magnification. CD8⁺ cell infiltrate was	
381	widespread across normal mucosal tissue.....	158
382	Figure 5-7. Quantification of IHC CD8⁺ positive cells using the QuPath analysis software	
383	platform. Adenomas are manually identified in FFPE gut roll sections. ROIs for quantification are	
384	annotated as appropriate (adenoma – red; normal – yellow). Quantification is then conducted: marker	
385	positive and negative cells are identified. Initially, the quantification is refined by adjusting thresholds	
386	before quantification of all samples. ROI: region of interest.....	159
387	Figure 5-8. CD8⁺ quantified by IHC in A) normal intestinal epithelium and B) intestinal polyps in	
388	Apc del, Apc del/Mlh1 del Apc del/Mlh1 null and Mlh1 null mice	159
389	Figure 5-9. Examples of IHC for CD8⁺ cells in normal intestinal mucosa and intestinal adenomas	
390	in Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x	
391	magnification. Visual inspection revealed marked CD8 ⁺ cell exclusion in polyps in each of the fur	
392	genotypes.....	160
393	Figure 5-10. Examples of IHC from CD4⁺ cells in normal intestinal mucosa of Black-6 WT and	
394	Vil-CreERT 2 mice. Shown at 2x, 10x and 20x magnification. CD4 ⁺ cell infiltrate was widespread	
395	across normal mucosal tissue	162
396	Figure5-11 Examples of IHC for CD4⁺ cells in normal intestinal mucosa and intestinal polyps in	
397	Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x	
398	magnification.....	163
399	Figure 5-12. CD4⁺ quantified by IHC in A) normal intestinal epithelium and B) intestinal polyps in	
400	Apc del, Apc del/Mlh1 del Apc del/Mlh1 null and Mlh1 null mice	164
401	Figure 5-13. Examples of IHC from C19⁺ cells in normal intestinal mucosa of WT and Vil-CreERT	
402	² , mice shown at 2x, 10x and 20x magnification. CD19 ⁺ cell infiltrate was sparse in normal intestinal	
403	epithelium, except for Peyer's patch areas (red arrow).....	165

404	Figure 5-14 Examples of IHC for CD19⁺ cells in normal intestinal mucosa and intestinal polyps	
405	in Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x	
406	magnification. Note the red arrow denotes the presence of a Peyer's patch, which acts as an internal	
407	positive control for CD19 IHC staining.	166
408	Figure 5-15. Examples of IHC for B2m cells in normal intestinal mucosa and GI polyps in Apc	
409	del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x	
410	magnification. B2m loss was universal to all polyps in each of the four genotypes examined.	168
411		
412		

413	Table of Tables	
414	Table 1-1. The consensus molecular subtype classification of CRC as proposed by Guinney et al	
415	[48]. SCNA: somatic copy number alteration. MSI: microsatellite instability. EMT: epithelial to	
416	mesenchymal transition.....	7
417	Table 1-2. The TNM staging classification for CRC	8
418	Table 1-3. Factors associated with a higher risk of disease recurrence post-surgery in stage II	
419	CRC patients [49]. CEA: carcinoembryonic antigen	9
420	Table 1-4. The TNM and FIGO staging systems for endometrial carcinoma	13
421	Table 1-5. Summary of the key innate immune cells , their function and clinical relevance in cancer.	
422	TGF- β : transforming growth factor beta. IL: interleukins. IFN- γ : interferon gamma APC: antigen	
423	presenting cells. TNF- α : tumour necrosis factor alpha. TLR: Toll-like receptor. OS: overall survival.	
424	PFS: progression free survival CRC: colorectal cancer. LUAD: lung adenocarcinoma. HCC:	
425	hepatocellular carcinoma. SKCM: skin cutaneous melanoma. HNSC: head and neck squamous cell	
426	carcinoma.GC: gastric cancer. BRCA: breast invasive carcinoma. ccRCC: clear-cell renal cell	
427	carcinoma	23
428	Table 1-6 Summary of the key adaptive immune cells , their function and clinical relevance in	
429	cancer. TGF- β : transforming growth factor beta. IL: interleukins. IFN- γ : interferon gamma APC:	
430	antigen presenting cells. TNF α : tumour necrosis factor alpha. TLR: Toll-like receptor.CRC:	
431	colorectal cancer. LUAD: lung adenocarcinoma. HCC: hepatocellular carcinoma. SKCM: skin	
432	cutaneous melanoma. HNSC: head and neck squamous cell carcinoma.GC: gastric cancer. BRCA:	
433	breast invasive carcinoma. NSCLC: non-small cell lung cancer	24
434	Table 2-1 Details of the antibodies used for T-cell quantification human CRC tissue	35
435	Table 2-2: Details of the antibodies used for immune cell quantification in murine FFPE tissues	
436		36
437	Table 2-3. Details of the primary antibodies used in Western Blot experiments	37
438	Table 2-4. Details of the secondary detection antibodies used in Western Blot experiments	37
439	Table 2-5 Mlh1 floxed allele PCR reagents	39

440	Table 2-6 APC floxed allele PCR reagents	40
441	Table 2-7 Mlh1 floxed allele PCR reagents	42
442	Table 3-1. Summary of the T cell populations and markers most commonly used in CRC	
443	prognostic studies.	46
444	Table 3-2. Summary of selected previous studies examining T cell infiltrate and clinical outcome	
445	in CRC. OS: Overall Survival. DFS: Disease Free Survival. RFS: Recurrence Free Survival. IM:	
446	Invasive Margin. CT: Centre of Tumour. IHC: Immunohistochemistry. MV: multivariable	47
447	Table 3-3. Comparison of the baseline characteristics of the QUASAR2 and VICTOR trial cases	
448	used in this biomarker study. pT: pathological tumour (T) stage; MMR: DNA mismatch repair;	
449	MMRp: mismatch repair proficient; MMRd: mismatch repair deficient; POLE-mutant: pathogenic	
450	POLE exonuclease domain mutation; KRAS-mutant: KRAS mutation at codons 12, 13 or 61. BRAF-	
451	mutant: BRAF hotspot V600E mutation. *determined by unpaired Student's t-test. †determined by	
452	Fisher exact test (in cases which marker status was determined). Reproduced from Glaire et al 2019	
453	[1]	57
454	Table 3-4. Comparison of baseline characteristics of VICTOR and QUASAR2 cases included in	
455	biomarker study vs. cases not included. *determined by unpaired Student's t-test. †determined by	
456	Fisher exact test (in cases which marker status was determined). Reproduced from Glaire et al 2019	
457	[1]	58
458	Table 3-5 Preliminary analyses of CD3 ⁺ and CD8 ⁺ cell fraction in the QUASAR2 trial samples.	
459	AIC: Akaike's Information criterion. P value calculated by Cox proportional hazards model	63
460	Table 3-6. Relationship between clinicopathological/molecular characteristics and tumour CD8 ⁺	
461	cell density in the combined QUASAR2 and VICTOR trial population. IQR: interquartile range.	
462	pT : primary tumour stage. POLE-mutant: pathogenic POLE mutation. MMR: mismatch repair	
463	MMR-p: mismatch repair proficient. MMR-d: mismatch repair deficient. KRAS-mutant: KRAS	
464	mutation at codons 12, 13 or 6. BRAF-mutant: BRAF hotspot V600E mutation * P value determined	
465	by non-parametric Mann-Whitney test. [∇] Determined by log-rank test.....	64
466	Table 3-7. Multiple linear regression analysis of the relationship between	
467	clinicopathological/molecular characteristics and CD8 ⁺ cell density in the combined QUASAR2	

468	and VICTOR trial samples. POLE-mutant: pathogenic POLE mutation. MMRd: mismatch repair	
469	deficiency. MMRp: mismatch repair proficient. CIN: chromosomal instability. CRC: colorectal cancer.	
470	KRAS-mutant: KRAS mutation at codons 12, 13 or 6. BRAF-mutant: BRAF hotspot V600E	
471	mutation.....	65
472	Table 3-8. Univariable and multivariable analyses of time to colorectal cancer recurrence and	
473	overall survival in pooled VICTOR and QUASAR2 trial population. pT : primary tumour stage.	
474	POLE-mutant: pathogenic POLE mutation. MMR: DNA mismatch repair MMRp: mismatch repair	
475	proficient. MMRd: mismatch repair deficient. KRAS-mutant: KRAS mutation at codons 12, 13 or 61.	
476	BRAF-mutant: BRAF hotspot V600E mutation. Reproduced from Glaire et al [1].....	70
477	Table 3-9. Estimated probabilities of colorectal cancer recurrence and overall survival	
478	according to tumour risk strata and CD8+ cell density dichotomized at sample median in the	
479	pooled QUASAR2 and VICTOR trial population Glaire et al 2019 [1].....	74
480	Table 3-10. Estimated probabilities of colorectal cancer recurrence and overall survival	
481	according to tumour risk strata and CD8+ cell density dichotomized at sample median in the	
482	pooled QUASAR2 and VICTOR trial population	75
483	Table 3-11. Details of the seven publicly available independent series included in the external	
484	validation cohort	80
485	Table 3-12. Details markers being using for multispectral, immunofluorescence	
486	immunoprofiling project of ~5000 cancers from QUASAR2 and SCOT trial	86
487	Table 4-1. Genetic alterations of Lynch Syndrome and the associated phenotype. The exact	
488	phenotype exhibited by LS sufferers depends on the underlying genetic abnormality. Reproduced	
489	from [284]	95
490	Table 4-2. Tumour infiltrating lymphocyte density by marker and compartment according to	
491	cancer type and MMR status. Values shown are sample medians with interquartile rangers (IQR) in	
492	parentheses. P values calculated by non-parametric Mann-Whitney U Test. Test for interaction were	
493	performed by aligned rank ANOVA. MMRp: Mismatch repair proficient. MMRd: Mismatch repair	
494	deficient.....	112

495	Table 4-3. Multiple linear regression analysis of the effect of MMR status, mutational load and	
496	immune escape mutation on the IFN- γ response in TCGA CRC and EC cancers. LOF – Loss of	
497	Function. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite unstable. MSI	
498	microsatellite stable.....	126
499	Table 5-1. Summary of the key murine models for the study of intestinal carcinogenesis. APC:	
500	adenomatous polyposis coli. AOM: azoxymethane. Cdx2: caudal type homeobox 2.....	137
501	Table 5-2. Summary of the key murine models for the study of the immune response and cancer	
502	development. Ifng: Interferon gamma. Pfp: perforin. TRAIL: TNF- α related apoptosis inducing	
503	ligand. B2m: beta-2 microglobulin. Stat1: signal transducer and activator of transcription 1	138
504	Table.5-3. Summary of the key murine models for the study of the immune deficiency and	
505	cancer genetic modification. Ifng: Interferon gamma. Pfp: perforin. Stat1: signal transducer and	
506	activator of transcription 1	139
507	Table 5-4. Summary of selected studies that have investigated mismatch repair loss in mice...	142
508	Table 5-5. Gastrointestinal polyps in experimental mice. Median number of polyps in the Apc del,	
509	Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null. IQR: Interquartile Range.....	148
510	Table 5-6. Site of gastrointestinal polyp formation. Median number of polyps in the Apc del, Apc	
511	del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null according to gastrointestinal section.....	155
512	Table 5-7. Multiple linear regression analysis of the relationship between total polyp number	
513	and mouse sex, timepoint and genotype	156
514	Table 5-8 CD8⁺ cell infiltrate in Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null	
515	normal intestine and polyps. Median values of CD8 ⁺ cells/mm ² presented with IQR in parentheses	159
516	Table 5-9 CD4⁺ cell infiltrate in Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null	
517	normal intestine and polyps. Median values of CD4 ⁺ cells/mm ² presented with IQR in parentheses	164
518		

519 **Glossary of abbreviations**

520

AOM	Azoxymethane
APC	Adenomatous Polyposis Coli
B ₂ M	Beta-2 Microglobulin
BER	Base Excision Repair
CIN	Chromosomal Instability
CRC	Colorectal cancer
CTL	Cytotoxic T-lymphocytes
DFS	Disease Free Survival
EC	Endometrial Cancer
EMT	Epithelial-to-mesenchyme transition
FAP	Familial adenomatous polyposis
FFPE	Formalin Fixed Paraffin Embedded
FIGO	International Federation of Gynaecological and Obstetrics
H&E	Haematoxylin and Eosin
HR	Hazard Ratio
ICI	Immune Checkpoint Inhibitors
IFN- γ	Interferon gamma
IHC	Immunohistochemistry
JAK1	Janus Kinase 1
LOH	Loss of Heterozygosity
LS	Lynch Syndrome
MAP	MUTYH associated polyposis
MDSC	Myeloid Derived Suppressor Cell
MHC	Major Histocompatibility Complex
MMR	Mismatch Repair Deficiency

MSI	Microsatellite instability
OS	Overall Survival
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PFS	Progression Free Survival
POLE	DNA Polymerase ϵ
SNV	Single Nucleotide Variants
TAA	Tumour Associated Antigens
TCGA	The Cancer Genome Atlas
TILs	Tumour Infiltrating Lymphocytes
TLS	Tertiary Lymphoid Structure
TMA	Tissue microarray
TMB	Tumour Mutational Burden
TNF- α	Tumour Necrosis Factor alpha
UICC	Union for Interventional Cancer Control
WHO	World Health Organisation

Publications arising from my work

Glaire MA, Ryan NAJ, Ijsselsteijn ME, Kedzierska K, Obolenski S, Ali RCrosbie EJ, Bosse T, de Miranda N, Church DN.

Discordant prognosis of mismatch repair deficiency in colorectal and endometrial cancer reflects variation in antitumour immune response and immune escape.

Under review.

Glaire MA, Domingo E, Sveen A, Bruun J, Nesbakken A, Nicholson G, Novelli M, Lawson K, Oukrif D, Kildal W, Danielsen HE, Kerr R, Kerr D, Tomlinson I, Lothe RA, Church DN. *Tumour-infiltrating CD8⁺ lymphocytes and colorectal cancer recurrence by tumour and nodal stage.* Br J Cancer. 2019. PMID: 31388185

Glaire MA, Brown M, Church DN, Tomlinson I. *Cancer predisposition syndromes: lessons for truly precision medicine.* J Pathol. 2017. PMID: 27859368.

de Jonge MM, Auguste A, van Wijk LM, Schouten PC, Meijers M, Ter Haar NT, Smit VTHBM, Nout RA, **Glaire MA**, Church DN, Vrieling H, Job B, Boursin Y, de Kroon CD, Rouleau E, Leary A, Vreeswijk MPG, Bosse T. *Frequent Homologous Recombination Deficiency in High-grade Endometrial Carcinomas.* Clin Cancer Res. 2019. PMID: 30413523

Workel HH, Lubbers JM, Arnold R, Prins TM, van der Vlies P, de Lange K, Bosse T, van Gool IC, Eggink FA, Wouters MCA, Komdeur FL, van der Slikke EC, Creutzberg CL, Kol A, Plat A, **Glaire MA**, Church DN, Nijman HW, de Bruyn M. A *Transcriptionally Distinct CXCL13⁺ CD103⁺ CD8⁺ T-cell Population Is Associated with B-cell Recruitment and Neoantigen Load in Human Cancer.* Cancer Immunol Res. 2019. PMID: 30872264.

Ryan NAJ, **Glaire MA**, Blake D, Cabrera-Dandy M, Evans DG, Crosbie EJ. *The proportion of endometrial cancers associated with Lynch syndrome: a systematic review of the literature and meta-analysis.* Genet Med. 2019 PMID: 31086306

557 Ryan NAJ, Blake D, Cabrera-Dandy M, **Glaire MA**, Evans DG, Crosbie EJ.
 558 *The prevalence of Lynch syndrome in women with endometrial cancer: a systematic*
 559 *review protocol.*
 560 Syst Rev. PMID: 30115102
 561
 562 Ryan NAJ, Walker TDJ, Bolton J, Ter Haar N, Van Wezel T, **Glaire MA**, Church
 563 DN, Evans DG, Bosse T, Crosbie EJ.
 564 *Histological and Somatic Mutational Profiles of Mismatch Repair Deficient Endometrial Tumours of*
 565 *Different Aetiologies.*
 566 Cancers (Basel). PMID: 34572765
 567
 568 Jiang D, Hveem TS, **Glaire M**, Church DN, Kerr DJ, Yang L, Danielsen HE.
 569 *Automated assessment of CD8⁺ T-lymphocytes and stroma fractions*
 570 *complement conventional staging of colorectal cancer.*
 571 EBioMedicine. 2021 PMID: 34479131
 572
 573 Domingo E, Freeman-Mills L, Rayner E, **Glaire MA**, Briggs S, Vermeulen L,
 574 Fessler E, Medema JP, Boot A, Morreau H, van Wezel T, Liefers GJ, Lothe RA,
 575 Danielsen SA, Sveen A, Nesbakken A, Zlobec I, Lugli A, Koelzer VH, Berger MD,
 576 Castellví-Bel S, Muñoz J; Epicolon consortium, de Bruyn M, Nijman HW, Novelli M,
 577 Lawson K, Oukrif D, Frangou E, Dutton P, Tejpar S, Delorenzi M, Kerr R, Kerr D,
 578 Tomlinson I, Church DN.
 579 *Somatic POLE proofreading domain mutation, immune response, and prognosis in colorectal cancer:*
 580 *a retrospective, pooled biomarker study.*
 581 Lancet Gastroenterol Hepatol. 2016 PMID: 28404093.
 582
 583 Temko D, Van Gool IC, Rayner E, **Glaire MA**, Makino S, Brown M, Chegwiddden L,
 584 Palles C, Depreeuw J, Beggs A, Stathopoulou C, Mason J, Baker AM, Williams M,
 585 Cerundolo V, Rei M, Taylor JC, Schuh A, Ahmed A, Amant F, Lambrechts D, Smit VT,
 586 Bosse T, Graham TA, Church DN, Tomlinson I.
 587 *Somatic POLE exonuclease domain mutations are early events in sporadic endometrial and*
 588 *colorectal carcinogenesis, determining driver mutational landscape, clonal neoantigen burden and*
 589 *immune response.*
 590 J Pathol. 2018 PMID: 29604063

1. Introduction

Cancer is a common disease that results in considerable morbidity and mortality worldwide [5], with an estimated 20 million new diagnoses, and 10 million cancer deaths annually. The World Health Organisation (WHO) ranks cancer as the first or second leading cause of death for those below 70 years old in 112 of 183 countries (Figure 1.1) [5]. Moreover, global cancer incidence and mortality are increasing, reflecting an aging population and increasingly prevalent risk factors [7].

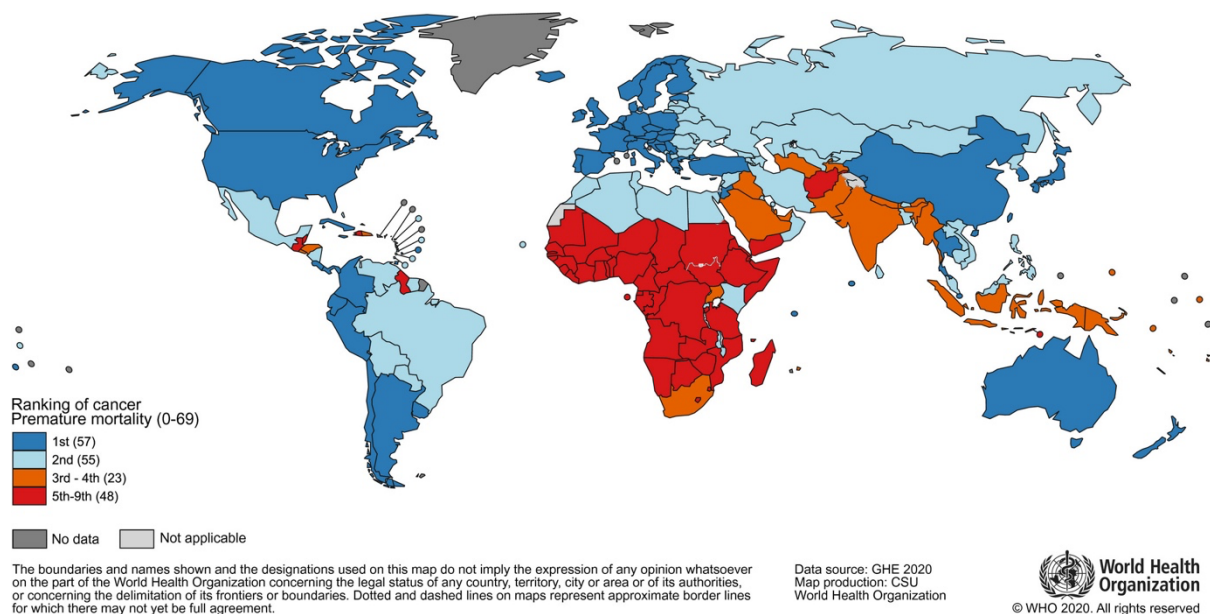


Figure 1-1. Cancer is a leading cause of premature death (death of those below age 70 years) in 112 countries worldwide. Data from the World Health Organisation and figure reproduced from GLOBOCAN 2020 report. [5]

Cancer is a disease caused by genetic and epigenetic aberrations that result in the loss of tumour suppressor gene function (genes that normally function to restrain cell growth), and gain of oncogene function (genes that serve to promote cellular growth) [8]. Whilst there is considerable biological variation between different types of neoplastic disease, underlying genetic aberrations result in fundamental characteristics that are shared by all cancer types; often termed the hallmarks of cancer [9, 10]. The original, highly influential, publication which proposed these, defined six biological hallmarks that enable tumour growth and metastatic spread [10]. These included:

- Sustaining proliferative signalling

- Evading growth suppressors
- Activating invasion and metastasis
- Enabling replicative immortality
- Inducing angiogenesis
- Resisting cell death

Considerable advances in the field of oncology research led to the inclusion of two more hallmarks in the latest iteration of this work, published in 2011, less than a decade after the first [9]:

- Deregulating cellular energetics
- Avoiding immune destruction

The inclusion of immune evasion as a hallmark of cancer was the culmination of longstanding interest in the intricate relationship between cancer and the immune response, first postulated by pathologist Rudolf Virchow over 100 years ago [11]. In the past three decades, this connection has become the focus of renewed attention, and substantial advances in understanding [12]. This has resulted in the development of anti-cancer immunotherapies, in particular immune checkpoint inhibitors (ICIs), which have provided dramatic responses in previously treatment-refractory cancers [13]. Immune checkpoint inhibitors work by potentiating a pre-existing, but ineffective, immune response [14]. However, their efficacy is far from universal across cancer types, or even within the same type of cancer [13]. In part, cancer immunity is determined by the underlying tumour's mutational landscape: higher mutational loads have been associated with a stronger immune response [15, 16]. But this alone is not sufficient to explain the variation in the quantity or quality of the antitumour immune response. Efforts are needed to fully elucidate how the tumour immune microenvironment is shaped by the interaction with the cancer genomic landscape and other factors. This will not only improve our understanding of the fundamentals of cancer biology, but may also help fully realise the therapeutic potential of the antitumour immune response.

1.1. Colorectal cancer

1.1.1. Epidemiology

Colorectal cancer (CRC) describes cancers arising from the epithelium of the colon or rectum [17]. It is the third commonest cancer worldwide in males and females, and as such is the cause of much morbidity and mortality [18, 19]. In the UK, it accounts for 11% of all new cancer diagnoses, with over 40,000 new cases, and 17,000 deaths, each year. The mean CRC incidence across the UK is 63.7 per 100,000 people [20]. It is commoner in males than females, with 56% of new cases being diagnosed in men. Fortunately, in the last decade CRC incidence has decreased by approximately 5%, and mortality rates by 10% [20]. This decrease in mortality may partly be explained by the introduction of the NHS bowel cancer screening programme for those aged 60 and over in the UK; the original trial data suggested CRC screening could reduce mortality rates by up to 15% [21]. The incidence of CRC varies by age with over 40% of cases diagnosed in people aged 75 and over. There is a sharp increase in incidence at ages 50-54, and the peak incidence is reached at 85-98 years old [20].

1.1.2. Aetiology

Both genetic and environmental factors play important roles in CRC development. 10-20% of CRC patients have a positive family history of the disease [22]; however, only 5-7% can be directly attributed to a Mendelian familial cancer syndrome [23]. Such Mendelian syndromes are often classified as either polyposis (characterised by the development of multiple pre-malignant adenomatous polyps – in some cases thousands – in addition to increased CRC risk), or non-polyposis (in which the increased CRC risk occurs in the absence of polyposis, although CRCs still typically develop from pre-malignant adenomatous lesions) [24]. The most common familial syndrome contributing to CRC development, and the predominant non-polyposis condition, is Lynch syndrome (previously referred to as hereditary non-polyposis colorectal cancer – HNPCC) [25], which will be discussed further below and in Chapter 4. Familial adenomatous polyposis (FAP) is the next most common hereditary condition associated with CRC [23]. It is an autosomal dominant condition that is caused by inheritance of a mutated copy of the *APC* gene [26]. The disease is characterised by very

large numbers of colonic polyps that begin to form in the second and third decade of life. The associated risk of developing CRC is almost 100%, and sufferers are often offered a prophylactic total colectomy or a proctocolectomy [27]. *MUTYH* associated polyposis (MAP), another important familial polyposis disease, is an autosomal recessive condition caused by defective base excision repair (BER) pathway, which leads to accumulation of DNA damage caused by reactive oxygen species, including DNA methylation, deamination and hydroxylation [28]. Sufferers develop multiple adenomas, although fewer than seen in FAP and typically no more than 100, and have a lifetime risk of CRC development of 60-70% [29, 30]. Other familial polyposis syndromes include *POLE/POLD1* polyposis, Peutz-Jeghers syndrome, Juvenile polyposis syndrome and *GREM1* mixed polyposis syndrome [23].

The risk of developing CRC is also moderately influenced by modifiable risk factors. Lifestyle choices such as eating red meat [31], prolonged heavy smoking [32, 33], and alcohol consumption all increase risk [34]. Moderate alcohol consumption (2-3 units/day) is associated with a 20% increased lifetime risk, whilst more substantial consumption correlates with a 50% increase [3]. Increased bodyweight has also been associated with significantly increased CRC risk in epidemiological studies; it is estimated that every one unit increase in body mass index (BMI) results in 2-3% CRC risk increase [35]. Type II diabetes mellitus is a risk factor, independent of the association with obesity and a sedentary lifestyle [36]. Conversely, there are a number of factors that can decrease the risk of CRC, which include: dairy product consumption, high amounts of dietary fibre, fruit and vegetable intake, vitamin D and calcium supplementation [3]. Low dose daily aspirin has also been shown to decrease CRC risk, but this may only apply to individuals of a certain genotype [37].

1.1.3.Pathogenesis

The factors detailed above promote the acquisition of genetic and epigenetic aberrations in normal colonic epithelial cells, which in combination may perturb the cell's normal regulatory processes. A key, and early event in colorectal carcinogenesis is the development of genomic instability; evidence of which is found in the majority of early precancerous lesions [38, 39]. Once genomic instability has

arisen, there is an acceleration in the accumulation of the genetic and epigenetic alterations in oncogenes and tumour suppressor genes that drive the malignant transformation of colonic cells [40]. A commonly held model is that loss of tumour suppressor gene function and gain of oncogene function leads to the development of cancer stem cells [41]. Studies from mice show that *Lgr5*⁺ stem cells, which reside between the base and the 4⁺ position in the colonic crypt, are one population capable of such transformation [42, 43], though there may be others.

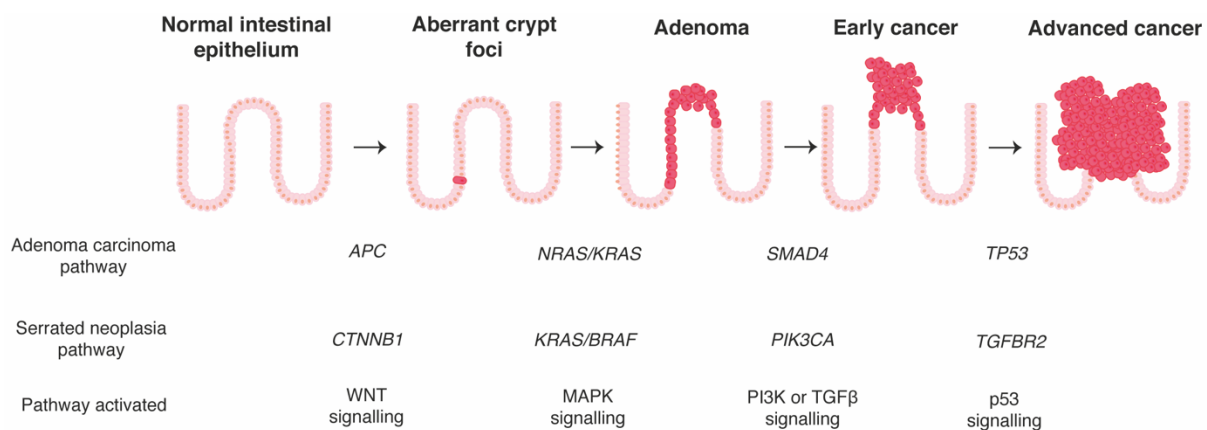


Figure 1-2: Normal epithelium to polyp to colorectal cancer sequence. Colorectal cancers result from stepwise neoplastic transformation of normal epithelial cells. There are two well defined pathways: the classic adenoma-carcinoma sequence, and the serrated neoplasia pathway. Whilst each of these are characterised by distinct genetic alterations, they share dysregulation of similar signalling pathways. *APC*: adenomatous polyposis coli. *SMAD4*: SMAD family member 4. *CTNNB1*: catenin-β1. *PIK3CA*: phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-α. *PTEN*: phosphatase and tensin homologue. *TGFβ*: transforming growth factor β. *TGFBR2*: TGFβ receptor 2. Reproduced from [3]

There are two well described pathways for the development of CRC. The “classic” adenoma-carcinoma sequence, and the serrated-neoplasia pathway (Figure 1.2) [3]. Each of these is characterised by distinct genetic and molecular features. The classic adenoma-carcinoma sequence accounts for 70-90% of CRCs and was first described by Fearon and Vogelstein (Figure 1.2) [44]. It begins with the formation of a small adenoma (<1cm) that arises from an aberrant crypt. This adenoma develops into a large adenoma (>1cm) and finally into a colorectal cancer. Each stage in this sequence is accompanied by histological and epigenetic/genetic changes. The entire process is estimated to take around 10-15 years [45]. The classic adenoma-carcinoma sequence is characterised by a form of genomic instability known as chromosomal instability (CIN), and the initiating event is most commonly *APC* loss [46]. The polyps formed during this process are of tubular or tubular-

villous histology. Alternatively, approximately 5-10% of CRCs arise from polyps of sessile serrated histology [45]. In this case the sequence of progression is hyperplastic polyp to sessile polyp to adenocarcinoma. The exact molecular and genetic characteristics of CRCs arising from the sessile pathway depends on the “sidedness” of the cancer [47]. Those that originate on the right side of the colon (caecum, ascending and transverse colon) typically have a CpG island methylator phenotype (CIMP) and exhibit microsatellite instability (MSI), a type of genomic instability that will be discussed in depth below. In contrast, left sided tumours more frequently harbour *KRAS* mutations and an attenuated form of CIMP [47].

Further classification of CRCs has also been made on the basis of gene expression data. Although multiple classifiers have been proposed, perhaps the most important initiative pooled data from multiple series to propose the consensus molecular subtypes (CMS) of colorectal cancer [48]. This comprises four main groups: CMS1 – 4 (summarised in Table 1.1). Each of these subtypes are characterised by distinct molecular and genetic features.

Table 1-1. The consensus molecular subtype classification of CRC as proposed by Guinney et al [48].
 SCNA: somatic copy number alteration. MSI: microsatellite instability. EMT: epithelial to mesenchymal transition

Consensus molecular subtype	Proportion of CRCs (%)	Key Features	Side	Sex association	Prognosis
CMS1 (MSI)	14	MSI hypermutated with CIMP. Frequent <i>BRAF</i> mutations	Right	Female	Better in early stage; worse after relapse
CMS2 (canonical)	37	Increased SCNA; frequent <i>APC</i> and <i>TP53</i> mutations with WNT and MYC activation	Left	Male	-
CMS3 (metabolic)	13	Mixed MSI status. Both CIMP and SCNA low. High rate of <i>KRAS</i> mutations	Right = Left	Male = Female	-
CMS4 (mesenchymal)	23	Increased TGF- β pathway activation with EMT. Poorly responsive to chemotherapy	Left (distal)	Male	Worse relapse free survival and overall survival

1.1.4. Clinical features

1.1.4.1. Diagnosis and staging

In the first instance, it is recommended that the diagnosis of a primary CRC lesion be made by full colonoscopy [49]. This has the advantage of allowing for tumour marking and obtaining biopsies for tissue diagnosis; additionally, it also allows for removal of synchronous early cancerous or precancerous lesions. To assess distal disease spread, patients undergo computed tomography scans of the thorax, abdomen and pelvis with iv contrast; in addition to baseline haematology, biochemistry and CEA measurements [49].

Staging of CRCs are made in accordance with the tumour, node, metastasis (TNM) classification developed by the Union for International Cancer Control (UICC) [50] (Table 1.2). TNM staging determines the primary course of treatment, and in addition is useful for prognostication.

734 Table 1-2. *The TNM staging classification for CRC*

AJCC Stage	T (Primary Tumour)	N (Regional Lymph Nodes)	M (Distant Metastasis)	Description
0	Tis	N0	M0	Earliest stage – carcinoma <i>in situ</i> . Tumour has not grown beyond the mucosa
I	T1, T2	N0	M0	The tumour has grown through the submucosa (T1), or into the muscularis propria (T2). There is no lymph node (N0) or distant site spread (M0)
IIA	T3	N0	M0	The tumour has either grown into the subserosa or pericolic/rectal tissues (T3); or has invaded into the visceral peritoneum (T4a); or has grown into nearby tissue or organs (T4b). No lymph node (N0) or distal spread (M0)
IIB	T4a	N0		
IIC	T4b	N0		
IIIA	T1 or T2	N1/N1c	M0	Tumour stage T1 – T4 with lymph node involvement. Metastasis in 1 regional lymph node (N1a), in 2-3 regional lymph nodes (N1b), or in the non-peritonealised pericolic/perirectal soft tissue without regional lymph node involvement (N1c). Alternatively, metastasis in 4-6 regional lymph nodes (N2a), or more than 7 (N2b)
	Or			
	T1	N2a		
IIIB	T3 or T4a	N1/N1c		
	Or			
	T2 or T3	N2a		
	Or			
IIIC	T1 or T2	N2b		
	T4a	N2a		
	Or			
	T3 or T4a	N2b		
	Or			
	T4b	N1 or N2		
IVA	Any T	Any N	M1a	Any tumour and any nodal stage, with distal disease spread: metastasis confined to one organ (M1a), more than one organ (M1b) or metastasis to peritoneum with/without other organ involvement (M1c)
IVB	Any T	Any N	M1b	
IVC	Any T	Any N	M1c	

735

736 **1.1.5. Management**

737 The choice of management strategies is dependent on the disease stage; however, in essentially all
738 cases surgical resection is required if cure is to be achieved. For some early stage CRCs (pTis and
739 pT1 pedunculated polyp carcinomas), it is possible to resect the tumour through minimally invasive
740 endoscopic techniques with proper follow up [49]. Locally infiltrative cancers, however, cannot be
741 resected endoscopically and as such require laparoscopic or open surgery. The exact operation, and
742 extent of colonic resection is dictated by the tumour's anatomical location. In addition to removal of

the bowel segment, excision of the local lymphatic drainage is also required [49]. Although at present there is some debate of the benefit of extensive lymphadenopathy versus more limited dissection, it is generally accepted that ≥ 12 lymph nodes should be retrieved to enable accurate prognostication [45].

Following surgery, adjuvant (post-operative) chemotherapy is often given with the intention of sterilising possible micrometastatic disease and thus increasing probability of cure [49]. The decision to administer such therapy is based on balancing the risk of recurrence with the expected chemotherapy benefits and the potential for harmful side effects. At present, adjuvant chemotherapy is recommended for all stage III CRCs, and a subset of stage II cancers with factors indicative of an increased recurrence risk (Table 1.3). Chemotherapy regimens generally involve either 5-fluorouracil (typically given as capecitabine – an oral 5-FU treatment) alone or in combination with oxaliplatin [49].

Table 1-3. Factors associated with a higher risk of disease recurrence post-surgery in stage II CRC patients [49]. CEA: carcinoembryonic antigen

Major prognostic factors
<12 lymph nodes sampled
pT4 stage including perforation
Minor Prognostic factors
High tumour grade
Vascular invasion
Perineural invasion
Tumour presenting with obstruction
High peri-operative CEA levels

For stage IV disease there is some debate over the benefit of offering resection of the primary tumour, compared with no resection and administration of systemic chemotherapy [51]. Localised resections of metastases in the lung and liver can be undertaken, with a view to long term disease control or even

possibly a cure. Whilst overall the prognosis for stage IV disease remains poor, the recent introduction of ICI therapies for the relatively small group of patients with mismatch repair deficient (MMRd)/MSI disease has provided some hope [52]. Pembrolizumab (anti PD-L1 humanised monoclonal antibody treatment) has been approved by NICE for previously untreated, metastatic MMRd/MSI CRCs [53].

The prognosis for CRC patients is very much dependent on disease stage at diagnosis. Whilst early-stage disease is associated with excellent outcomes, the outlook remains poor for patients with advanced CRC. Stage I disease is associated with a 93.1% 5-year survival rate, stage II 84.1%, stage III 64.9% and stage IV 10.3% [45, 49]. There is an urgent to develop better treatment modalities for metastatic CRC, and for those patients with disease recurrence despite optimal treatment.

1.2. Endometrial cancer

1.2.1. Epidemiology

Endometrial cancer (EC), or more formally, uterine corpus endometrial cancer, refers to malignancies arising from endometrial lining of the uterus. It is the fourth most common cancer in women worldwide, and the commonest gynaecological cancer in high-income countries [54]. In the UK, there are 9,400 new cases diagnosed each year, accounting for approximately 5% of all cancer diagnoses in women. More than a quarter of these diagnoses are in women aged 75 and over [55]. Unlike most other malignancies, there has been an increase in the incidence of EC over the last four decades. The age standardised incidence rate was 18.4 in 1993, but this rose to 29.1 in 2017 [55]. It has been suggested that this may be in part due to a reduction in hysterectomies for benign reasons, or increasing obesity levels [55]. There are approximately 2,500 deaths from EC every year; this has risen by a quarter in the last decade in the UK [55].

1.2.2. Aetiology

Similarly to CRC, risk factors for EC can be considered as environmental and genetic. Approximately 3% of EC are associated with LS. Female LS sufferers have a 40-60% lifetime risk of EC, and the median age of diagnosis is 48 years [56, 57]. Environmental risk factors also play a considerable role in the development of EC [58]. The environmental risk factors are those that lead to the uterus receiving unopposed oestrogen exposure, which can either be from endogenous or exogenous sources [59]. In resource rich countries, obesity and metabolic syndrome associated conditions, such as diabetes and polycystic ovarian syndrome, are becoming increasingly important risk factors. For example in the USA, 57% of all endometrial cancers are attributable to obesity [60]. The lifetime risk of EC for women with a BMI in the normal range is 3%. For every 5-unit increase in BMI this increases by more than 50% [61]. Alternatively, exogenous oestrogen sources can play a role in EC development. These include oestrogen only hormone replacement therapy, and tamoxifen treatment, which has an anti-oestrogenic effect in the breast but a pro-oestrogenic effect in the endometrium. The risk of EC is doubled after 5 years of tamoxifen treatment [62, 63]. In contrast, the oral contraceptive pill protects against the development of EC, with a risk reduction of 30-40% with long term use [64].

1.2.3. Pathogenesis

ECs arise from the epithelial lining of the uterus, and historically has been divided into two types on the basis of risk factors and tumour histopathology. The first is endometrioid type cancers (Type 1) which are oestrogen dependent and account for 75-80% of ECs. The second is non-endometrioid carcinomas (Type 2), which are oestrogen independent and are further subdivided into serous cell (10%) and clear-cell carcinomas (4%), along with other rarer histotypes. Endometrioid endometrial cancers (EEC) are recognised to arise from precursor lesions in both pre- and post-menopausal women [65]. The carcinogenesis process is believed to begin with uninterrupted endometrial proliferation, caused by excessive and unopposed oestrogen exposure. This leads to endometrial hyperplasia, and from there localised lesions of premalignant atypia arise, known as endometrial intraepithelial neoplasia (EIN) [66]. These may then progress into endometrial adenocarcinomas, which are characterised by myometrial invasion. The transformation from normal epithelium to

cancer is driven by the acquisition of *PTEN*, *PIK3CA* and *KRAS* mutations [58]. In contrast, endometrial serous carcinomas are similar to high-grade serous ovarian and fallopian carcinomas, and are characterised by *p53* mutations and CIN. Clear-cell carcinomas are relatively rare, and have a mixture of Type 1 and Type 2 features [65].

This traditional pathological grouping of EC was refined by work from the TCGA consortium, which sought to perform detailed molecular analysis of endometrial cancers using whole exome sequencing, methylation profiling and gene expression analysis. The results from this landmark study suggested four molecular EC subsets [67]:

1. Ultramutated ECs: characterised by *POLE* mutations, and a very high mutational load (232×10^{-6} mutations per Mb)
2. Hypermutated ECs: cancers that have microsatellite instability due to defective DNA mismatch repair, most commonly caused by *MLH1* promoter methylation. This subset has mutation burden of 18×10^{-6} mutations per Mb (approximately tenfold higher than MSS cancers)
3. Copy number low ECs: typically endometrioid tumours with a lower mutation burden (2.9×10^{-6} mutations per Mb), proficient mismatch repair and low levels of copy number alterations
4. Copy number high ECs: typically of serous histology (though some are endometrioid) with low mutation burden (2.39×10^{-6} mutations per Mb). The majority had *TP53* mutations, and a high frequency had *FBXW7* (22%) and *PPP2R1A* (22%) mutations

1.2.4. Clinical management

A large percentage of ECs are diagnosed at an early stage due to the presentation of abnormal uterine bleeding in the majority of cases (90%). A histological diagnosis of EC was traditionally made from samples obtained from dilation and curettage (D&C). However, the accuracy of this method has been called into question [68] and there is an emphasis on the use of less invasive techniques such as endometrial biopsy, vaginal ultrasound scanning and hysteroscopy [69].

Staging is conducted using the TNM classification and the International Federation of Gynaecology and Obstetrics system (FIGO) [50, 70, 71] (Table 1.4). It provides important information for prognostication, and aids decisions about the use of adjuvant therapies. The mainstay of management for endometrial cancer is surgery. Typically, for stage I disease patients undergo total hysterectomy with bilateral salpingo-oophorectomy; for stage II disease this often accompanied by systematic pelvic lymphadenectomy with or without para-aortic lymph node removal. In the cases of stage III and IV disease, maximal surgical debulking is typically performed for tumours that are resectable [68].

Table 1-4. The TNM and FIGO staging systems for endometrial carcinoma

TNM Classification	FIGO system	Description
Tx		Primary tumour cannot be assessed
T0		No evidence of primary tumour
T1	I	Tumour confined to the corpus uteri
T1a	IA	Tumour limited to the endometrium, or invades less than half of myometrium
T1b	IB	Tumour invades half or more of myometrium
T2	II	Tumour invades cervical stroma; no extension beyond the uterus
T3	III	
T3a	IIIA	Tumour invades the serosa of the corpus uteri or adnexa
T3b	IIIB	Vaginal or parametrial involvement (direct extension or metastasis)
T4	IVA	Tumour invading the bladder mucosa, bowel mucosa or both
N1, N2	IIIC	Metastasis to pelvic or para-aortic lymph nodes
N1	IIIC1	Metastasis to pelvic lymph nodes
N2	IIIC2	Metastasis to para-aortic lymph nodes, with or without pelvic lymph node involvement
M0		No distant metastasis
M1	IVB	Distant metastasis

Following surgical resection of EC, post-operative radiotherapy and sometimes chemotherapy may be given with the aim of curing localised disease and distant metastases respectively, with their use determined by patient and tumour factors. Grade 1-2, stage I tumours with less than 50% myometrial involvement do not require adjuvant radiotherapy [68]. The remaining stage I cases are further stratified on the basis of age, tumour grade, histological features, the degree of myometrial invasion and lymph node involvement. In the past, patients with early stage disease defined to high risk of recurrence – those with grade 3 tumours and invasion beyond 50% of the myometrium – were offered pelvic radiation therapy [68]. However, the results from two large prospective studies showed that there was no benefit from this approach [72]. The presence of spread to lymph nodes in the pelvic or paraaortic regions deems patients at high risk of local and distant recurrence. At present there isn't a consensus about the best way to treat these patients. The PORTEC-3 trial showed that patients with stage III benefited from chemoradiotherapy in the form of four cycles of carboplatin and paclitaxel chemotherapy: it was associated with improved overall survival and 5-year recurrence free survival [73]. However, the COG -249 trial did not show a benefit of adding chemotherapy to radiotherapy in the treatment [74]. For patients with advanced or recurrent disease, treatment options remain relatively limited. Chemotherapy is used for advanced disease, and hormonal therapy for those not suitable for chemotherapy [68]. In the USA, the FDA has approved the use the immune checkpoint inhibitor pembrolizumab for the treatment of all metastatic MSI/MMRd cancers, including endometrial.

Overall, the 5-year survival rate for EC is good, largely reflecting the fact that most cases are diagnosed at an early stage. Approximately three quarters of women diagnosed to EC will survive 5 years or more. In the UK, the 5-year survival for stage 1 disease is 92.2%, for stage II 74.1%, stage III 47.8% and for stage IV is 15.1% [55].

1.3. Genomic instability in colorectal and endometrial cancer

As discussed previously, human cancers arise due to the acquisition of DNA alterations that collectively lead to unchecked cell proliferation, and loss of normal homeostatic mechanisms [9]. The mutagenic agents that cause DNA damage can be broadly classified as endogenous and exogenous. In melanoma and lung cancer, for instance, the major contributing mutagenic forces are exogenous (UV radiation and tobacco smoking respectively). In contrast, DNA alterations in CRC and EC are thought to arise mainly from endogenous agents (although there is increasing awareness of the mutagenic role of the microbiome in CRC) [75]. To ensure the integrity of the genome is maintained, cells employ a number of mechanisms to prevent and repair DNA damage should it arise. Failure of one, or more of these mechanisms, leads to genomic instability – a characteristic of almost all human cancers [40]. The exact form of genomic instability depends on which mechanism has failed, but can be broadly classified into three groups [76]:

- I. Chromosomal instability (CIN), which is the increase in the rate of changes in chromosome number and structure compared to normal cells. This can be caused by failure of chromosome segregation during mitosis, or failure of mitotic checkpoints
- II. Mutations at the DNA base level, which includes point mutations (single nucleotide variants – SNVs), small insertions and deletions (indels) and DNA microsatellite losses or expansions. These are often caused by inaccurate DNA synthesis, failure of Base Excision Repair (BER)/Nucleotide Excision Repair (NER), or MMRd
- III. Gross chromosomal rearrangements, copy number variants (CNV) and loss of heterozygosity (LOH). These result from the combination of single stranded DNA gaps (ssDNA) or double strand breaks (DSBs), and dysfunctional homologous recombination, which is responsible for repairing these breaks.

Understanding of the mutational processes that contribute to cancer development has greatly increased in the last decade. Previously research has focused on the discovery of driver mutations, which are key somatic mutations central to carcinogenesis. However, it has been recognised that cancer genomes contain not only driver mutations, but also numerous passenger mutations. These

may not have a causative role in cancer development, but they can provide an “historical log” of the mutational forces that have acted upon cancer cell genomes. Efforts to decipher this have been greatly aided by the advent of next generation sequencing technologies, which have vastly increased the speed and scale of sequencing efforts [77]. In particular, work by Alexandrov and colleagues, using data from TCGA and other sources, have produced a compendium of the mutational signatures detected in cancer [78]. SNVs, or point mutations, are a common type of DNA alteration and involve the replacement of one base for another in one of six possible ways (C·G→A·T, C·G→G·C, C·G→T·A, T·A→A·T, T·A→C·G and T·A→G·C). While it has long been known that some point mutations occur more commonly with specific types of DNA damage, more recently it also became clear that the sequence flanking the mutation is important and affects the rate of mutation. Considering the bases 5’ and 3’ to the mutation, it is possible to define 96 different trinucleotide mutation patterns (6 possible base pair substitutions multiplied by 16 possible adjacent nucleotides) [77]. Alexandrov and colleagues found that the exact number and frequency of the 96 variants could be correlated with particular mutagenic agents and processes. For instance, in adenocarcinomas, squamous cell carcinomas and small cell carcinomas of the lung there was increased levels of C>A substitutions with a transcription strand bias, which was defined as Signature 4 (Catalogue of Somatic Mutations in Cancer (COSMIC) SBS4). This corresponds to DNA damage caused by tobacco smoking. In melanoma samples, there was a high frequency of Signature 7, which is C>T mutations at trinucleotides CpCpN and TpCpN, which are the result of UV radiation. Analysis of the mutational signatures in CRC and EC show that apart from Signature 1B (age associated) these two cancers share specific forms of DNA damage associated with errors introduced during genome replication: Signatures 6, 14 and 15, the result of MMR deficiency [78], and Signature 10a and b, caused by *POLE* mutations [79].

1.3.1. DNA Polymerase ϵ (*POLE*) mutations

Accurate and efficient DNA replication is essential to ensuring genomic stability; an error rate of 1×10^{-9} per base or one mutation per genome doubling is testament to the effectiveness of the processes

involved [80]. This achieved by a combination of the proofreading activities of DNA polymerases, and post-replication surveillance and repair conducted by the MMR system (mentioned below, and discussed fully in chapter 4). In humans, there are three replicative DNA polymerases: α , δ and ϵ . Synthesis of the leading and lagging DNA strands is initiated by the Pol α – RNA primase complex; the DNA synthesis for the leading strand is carried out by Pol δ , and the lagging strand by Pol ϵ [81]. Both of these DNA polymerases are composed of four subunits, the largest of which contains the catalytic and exonuclease domains, and is coded for by *POLD1* and *POLE* respectively [82]. Incorporation of an incorrect base results in pausing of the polymerase activity, and activation of the exonuclease function. This results in excision of the aberrant base, and replacement with the correct base to allow DNA synthesis to resume [83]. Loss of the proofreading function of Pol δ or Pol ϵ results in genomic instability. In yeast models using *S. cerevisiae*, disruption of the exonuclease domain in either of these two polymerases resulted in 100 fold increase in the mutation rate [84, 85]. Furthermore, mice with a similar loss of polymerase proofreading function also show an increase in mutation rate and develop tumours, including gastrointestinal malignancies[83, 86].

Mutations in DNA polymerase proofreading domains have similar consequences in humans. Analyses of CRC exome data by TCGA found 16% of the tumours to carry substantially elevated mutational load (>10 mut/Mb) [87]. The majority of these exhibited microsatellite instability, in keeping with loss of DNA mismatch repair, typically due to *MLH1* silencing by hypermethylation, However, unexpectedly one quarter of hypermutated tumours were found to be microsatellite stable. These tumours were found to carry somatic *POLE* mutations within the exonuclease domain – most commonly a proline to arginine/histidine substitution at codon 286 (*POLE*^{P286R/H}), or substitutions at codons 411(*POLE*^{V411L}) and 459 (*POLE*^{S459F}) [87]. Interestingly, the mutation burden of these cancers was even greater than that of the hypermutated MSI phenotype, and as such they were labelled “ultramutated”. Subsequent studies demonstrated that somatic *POLE* proofreading mutations also occur in endometrial cancer with even greater prevalence [67]. Overall, *POLE* proofreading domain mutations occur in 1-2% of CRCs and 7-12% of ECs, and are also observed, albeit less frequently, in cancers of the brain, pancreas, ovary, breast and stomach [88].

In addition to somatic gene mutations, loss of polymerase proofreading function can occur as a result of germline mutations. Work by Palles and colleagues showed that heterozygous germline mutations in *POLD1*(*POLD1*^{S478N}) and *POLE*(*POLE*^{L424V}) were present in affected individuals in families with hereditary CRC unexplained by other predisposition variants [89]. The *POLD1* variant was also associated with the development of EC [89]. These mutations were confirmed to have a deleterious effect on DNA proofreading function in yeast models [84].

The consequences of *POLE* mutations for tumour biology, and clinical outcomes are important. *POLE* mutant CRCs and ECs are characterised by a vigorous immune reaction, with high densities of tumour infiltrating T-lymphocytes [90, 91]. This is thought to be the consequence of their very high mutational burden leading to an increased repertoire of tumour neoantigens [92]. In both EC and CRC, *POLE* mutations are associated with more favourable prognosis [90, 91].

1.3.2. The DNA Mismatch Repair System

The DNA mismatch repair system is also essential for maintaining genomic stability. This is the focus of Chapter 4, and will be discussed in detail there; hence a brief summary only is provided here. A highly evolutionary conserved system, it is responsible for correcting insertion deletion loops (IDLs) and base-base pair mismatches that are introduced during genomic replication [93]. Loss of mismatch repair machinery (MMRd) results in a genomic instability referred to as MSI, and is an important contributor to the development of both CRC and ECs. MMRd characterises 15-20% of CRCs and 20-30% of EC [94]. Pertinently, MMRd in CRC is associated with a vigorous anti-tumour immune response, improved prognosis and more favourable response to ICI [95]. However, the clinical consequences are less well characterised in EC.

1.4. The immune microenvironment in cancer

Virchow's original observations on the antitumor immune response focused on the frequent presence of leukocytes in human tumour tissue [11]. He hypothesised that this was the result of cancers

commonly arising at sites of chronic inflammation [96]. In the years since Virchow's work, our understanding of cancer and the immune system has been greatly enhanced and refined. It is widely accepted that the immune system can promote cancer development through chronic, persistent and unresolved inflammation [97]. It has been suggested that chronic inflammation increases cancer risk in all organs and tissues [98]. Specific key examples include gastric cancer associated with *Helicobacter pylori* infection, hepatocellular cancer occurring secondary to Hepatitis B or C infection and CRC risk from inflammatory bowel disease (IBD) [12]. However, it is also now clear that the immune system is essential in preventing cancer development, as evidenced by human studies as well as functional analyses of mouse models [99]. Whether the immune system exerts a pro- or anti-tumorigenic effects appears to depend not only on the specific immune cell subtype(s) involved but also the microenvironmental context (Figure 1.3). Chronic inflammation is a common consequence of a seemingly diverse group of factors recognised to increase cancer risk, including smoking, obesity and viral infections [12]. In the earliest stages of malignancy, chronic inflammation creates an environment that propagates malignant transformation, and may even contribute to genomic instability developing [100, 101]. Broadly speaking, the establishment of an inflammatory response is predominantly driven by the innate immune system, key elements of which are summarised in Table 1.5 and Figure 1.3. Neutrophils, which are the most abundant immune cell type and the first attracted to sites of inflammation, are recognised to favour tumour growth [4]. Meta-analyses have found that neutrophils are associated with worse outcomes in practically all cancer types. Templeton and colleagues showed that a neutrophil to lymphocyte ratio of more than four correlated with worse disease-free survival (DFS), progression-free survival (PFS) and overall survival (OS) in numerous different solid malignancies [102]. It was suggested this adverse prognostic association could reflect the pro-inflammatory nature of neutrophils that prevents an effective anti-tumour immune response. In particular, the cytokine CXCL5, produced by neutrophils, is associated with worse survival in several cancers [103-106]. The production of other proinflammatory cytokines, such TNF- α and IL-6, by neutrophils and other innate immune cells promotes malignant cell transformation as well as encouraging further immune cell infiltration [107].

1013

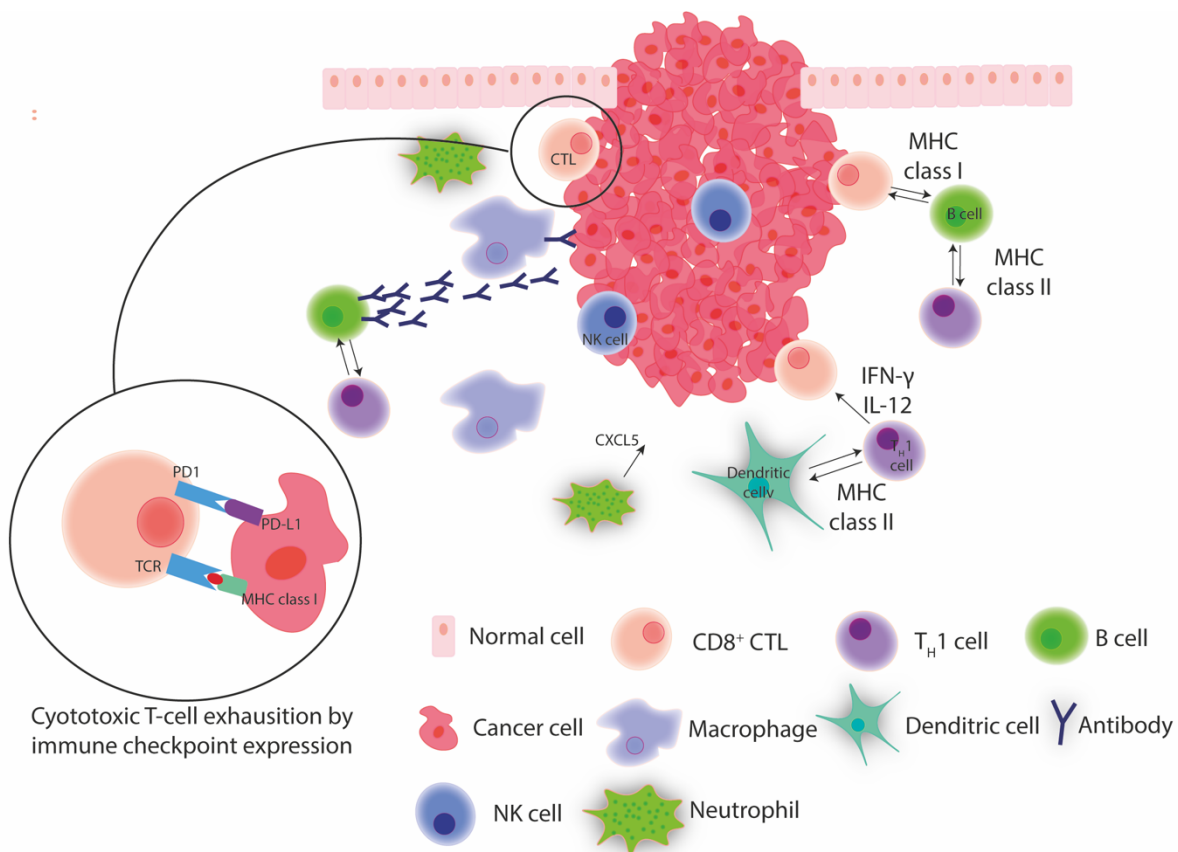


Figure 1-3. The immune microenvironment in cancer. Elements of the immune response can either be pro or anti-tumour. Neutrophils are typically considered to have pro-tumorigenic effects, and this may be in part from production of CXCL5. In contrast, elements of the adaptive immune response have clear anti-tumour effects. In particular, CD8⁺ cytotoxic T-lymphocytes (CTLs) and NK cells are able to directly destroy tumour cells. The former do so after activation by antigen presentation through the MHC class I pathway, whilst the latter are able to act independently. Tumour cells may avoid CTL destruction by upregulation of immune checkpoints such as PD-L1. T_H1 helper cells exert anti-tumour effects by driving an IFN- γ immune based response; they are activated primarily by antigen presentation via MHC class II interaction with antigen presenting cells such as dendritic cells. B-cells act as professional antigen presenting cells, or induce humoral macrophage phagocytosis by antibody production. CTL: cytotoxic T-lymphocytes. NK cells: natural killer cells. TCR: T-cell receptor. PD-L1: programmed cell death ligand 1. PD1: programmed cell death protein 1. Figure adapted from Bruni et al [4]. Further details of the immune cells represented are available in Tables 1.5. and 1.6.

1014 In contrast to the pro-tumorigenic effects of the immune system, there is a substantial and ever-
 1015 expanding body of evidence that has demonstrated an anti-tumour role for the immune system [4].
 1016 Adaptive immune cells that play a key role in the tumour microenvironment are summarised in Table
 1017 1.6 and Figure 1.3. A key feature of the adaptive immune response is recognition of specific antigens,
 1018 or neoantigens [107]. The somatic alterations that accompany carcinogenesis leads to an altered
 1019 repertoire of neoantigens that are presented via MHC pathways [92]. This identifies malignant cells as
 1020 non-self. The recruitment of adaptive immune cells is driven by antigen presenting cells, and
 1021 cytokines such as IFN- γ and IL-12 [107]. The interaction between the immune system and cells

undergoing neoplastic transformation is a dynamic process, and is referred to as immunoediting [108].

There are three possible outcomes of this interaction (frequently referred to as “The Three E’s of Immunoediting”) [108]:

- Elimination: nascent cancer cells are destroyed by an effective anti-tumour immune response
- Equilibrium: when cancer cells survive elimination but face continual selective pressure from an effective adaptive immune response
- Escape: selection pressure from the immune response in the equilibrium phase results in the development of clones that can grow in spite of intact anti-tumour immunity

T-cells, and in particular CD8⁺ cytotoxic T lymphocytes (CTLs), are an essential part of effective anti-cancer immunity. CD8⁺ cells are predominantly responsible for defence against intracellular pathogens such as bacteria and viruses, but are also capable of directly destroying cells that have undergone malignant transformation. They are activated via antigens presented by the MHC class I on tumour cells themselves, or via professional antigen presenting cells. Their anti-tumorigenic effect appears to be almost universal across all cancer types studied with numerous studies showing their positive association with improved clinical outcomes, including in CRC and EC [109]. Indeed, a recent review by Bruni and colleagues, revealed that CD8⁺ T cells had the strongest association with patient survival of all immune cell subtypes analysed [4]. The potential of CD8⁺ cells as a prognostic biomarker has been especially well studied in CRC [110] and will be discussed in greater detail in Chapter 3. Additionally, CD8⁺ cells are also a predictive marker, with their presence within the immune microenvironment being associated with ICI responsiveness [111]. Other notable adaptive immune cell subtypes that are typically associated with anti-tumour effects including T_H1 CD4⁺ cells and B-cells [109]. T_H1 CD4⁺ are a subset of T helper cells that are associated with an IFN- γ dominated anti-tumour response. IFN- γ can directly exert anti-tumorigenic effects [112], in addition to promoting activation and differentiation of other T-cells, including CD8⁺ T-cells [113, 114]. In the tumour microenvironment, B-cells contribute to anti-tumour immunity by both acting as antigen presenting cells, and hence participating in T-cell activation, and also production of tumour specific antibodies, which can induce macrophage driven phagocytosis [115].

1049

1050 The role of other immune cells in cancer biology is less clear. Macrophages are conventionally
1051 subclassified into M1 (pro-inflammatory) and M2 (anti-inflammatory) subtypes [116]. To date there
1052 has not been consistent demonstration that either macrophage subset are pro- or anti-tumorigenic with
1053 some studies reporting favourable associations [117-119] and others the opposite [120-123]. In part,
1054 this may reflect the fact that the M1/M2 classification lacks nuance and fails to adequately represent
1055 the diversity in macrophage population [4]. Future work that would enable more precise immune cell
1056 characterisation will help these efforts.

1057

Introduction

1058 **Table 1-5. Summary of the key innate immune cells, their function and clinical relevance in cancer.** *TGF- β* : transforming growth factor beta. *IL*: interleukins. *IFN- γ* :
 1059 *interferon gamma* *APC*: antigen presenting cells. *TNF- α* : tumour necrosis factor alpha. *TLR*: Toll-like receptor. *OS*: overall survival. *PFS*: progression free survival *CRC*:
 1060 colorectal cancer. *LUAD*: lung adenocarcinoma. *HCC*: hepatocellular carcinoma. *SKCM*: skin cutaneous melanoma. *HNSC*: head and neck squamous cell carcinoma. *GC*:
 1061 gastric cancer. *BRCA*: breast invasive carcinoma. *ccRCC*: clear-cell renal cell carcinoma

Immune cell type	Description	Clinical relevance in cancer	Refs
Neutrophils	Most abundant immune cell. Recruited to acute inflammation sites by chemokines (e.g CXCL8). Phagocytosis of pathogens and release of antimicrobial factors	Associated with worse outcomes across almost all cancer types – independent of stage and subtype	[102]
Macrophages	Tissue resident, differentiated monocytes; conventionally classified as M1 (pro-inflammatory response – secretion of IL-12 and ROS; promotion of T _H 1-type response) and M2 (anti-inflammatory effects by secretion of IL-10 and TGB- β)	Possible positive correlation with survival in CRC. Contrastingly TAMs (usually pro-tumorigenic M2 phenotype) associated with worse OS in BRCA, GC, ovarian, bladder and thyroid.	[117-119, 121-123]
Dendritic cells (DCs)	Professional APC; essential for T-cell activation and maturation	Dependent on particular subset. Increased levels of CD208 ⁺ associated with worse disease outcome in CRC, whereas mature/activated DCs associated with improved outcomes. Increased levels of cDC1 associated with improved survival in SKCM, BRCA, HNSC and LUAD	[124-126]
Natural Killer (NK) cells	Destruction of viral infected cells, and early malignant cells. Capable of killing tumour cells without priming or prior stimulation. Secrete IFN γ and TNF α , which acts on other immune cells like macrophages and dendritic cells	Associated with improved outcomes in CRC, HCC and ccRCC	[127-131]
Myeloid-derived suppressor cells (MDSC)	Immature myeloid cells with widespread immunosuppressive activity: inhibit T-cell activation and prevent NK cell mediated killing. Two main groups: polymorphonuclear MDSCs (PMN-MDSCs) and monocytic MDSCs (M-MDSCs)	Associated with worse PFS and OS in multiple solid tumours, including: NSCLC, melanoma, thyroid, bladder, HCC, GC and CRC	[132-136]
Mast cells	Long term tissue resident cells; important role in immune response to parasitic infection and allergic reactions. Release cytokines upon activation (IgE or TLR) that affect functioning of neighbouring immune cells and promote systemic inflammation	Associated with improved prognosis in prostate, BRCA and CRC; poorer prognosis in bladder cancer and GC	[137-140]

Introduction

1062 **Table 1-6 Summary of the key adaptive immune cells, their function and clinical relevance in cancer.** *TGF- β* : transforming growth factor beta. *IL*: interleukins. *IFN- γ* :
 1063 *interferon gamma* *APC*: antigen presenting cells. *TNF α* : tumour necrosis factor alpha. *TLR*: Toll-like receptor. *CRC*: colorectal cancer. *LUAD*: lung adenocarcinoma. *HCC*:
 1064 *hepatocellular carcinoma*. *SKCM*: skin cutaneous melanoma. *HNSC*: head and neck squamous cell carcinoma. *GC*: gastric cancer. *BRCA*: breast invasive carcinoma. *NSCLC*:
 1065 *non-small cell lung cancer*

Immune cell type	Description	Clinical relevance in cancer	Refs
Cytotoxic CD8 ⁺ T-cells	Defence against intracellular pathogens (viruses and bacteria) and tumour surveillance. Destruction of malignant cells via cytokine production, release of cytotoxic granules and induction of apoptosis by Fas/FasL interactions	Strongly associated with improved OS and DFS in numerous cancers at different anatomical sites, including CRC and EC	[141-149]
T helper (T _{H1})	Defence against intracellular pathogens by production of IFN γ and TNF α , which leads to macrophage and CD8 ⁺ cell activation	Strongly associated with improved outcomes in almost all cancer types, including CRC, BRCA, RCC and ovarian	[150-154]
T helper (T _{H2})	Defence against extracellular pathogens; secretion of IL-4, IL5, IL-10 and IL-13. Stimulate B-cell antibody production	Presence in TME of most aggressive cancers; associated with worse outcomes in pancreatic cancer	[155]
T helper (T _{H17})	Activated by APCs secreting IL-23. Responsible for defence against extracellular bacteria and fungi by secretion of IL-17	Associated with poorer prognosis in CRC, NSCLC and HCC; improved OS in oesophageal, GC, ovarian and cervical cancers.	[150, 156, 157]
T helper (T _{FH})	Specialised subset of CD4 ⁺ cells: facilitate germinal centre formation, and drive B-cell proliferation and maturation	Associated with improved outcomes in CRC and BRCA	[158, 159]
T _{reg} cells	Responsible for regulating and suppressing other immune cells: maintenance of peripheral tolerance thus preventing autoimmunity and chronic disease	Associated with improved outcomes in CRC, bladder, and head and neck cancers; poor prognosis in cervical and renal.	[160-165]
B-cells	Integral to adaptive humoral immunity. Antibody production (plasma cells), and antigen presentation	Increased B cells associated with improved outcomes in CRC, HCC, melanoma, ovarian, NSCL and BRCA.	[120, 158, 166-170]

1066

1.5. Determinants of the anti-tumour immune response

The composition of the tumour immune microenvironment is ultimately determined by the interplay between the tumour characteristics, the immune system and other factors, such as host genetics or the microbiome.

1.5.1. Tumour mutational landscape and antigens

The basis of an anti-tumour response lies in identification of malignant cells as non-self, a process driven by recognition of tumour associated antigens (TAA), which may derive from expression of mutated peptides or aberrant expression of peptides which are normally non-expressed [92]. The first group, referred to as neoantigens, are novel peptides arising from DNA mutations. The widely accepted paradigm is that tumours with higher mutational burdens are more likely to possess peptides capable of binding to HLA molecules and stimulating the adaptive immune response. In a seminal study, Rooney and colleagues conducted a pan-cancer study looking at the effects of tumour characteristics on anti-tumour immunity [171]. The authors proposed the CYT metric, derived as the geometric mean of *perforin 1 (PRF1)* and *granzyme-B (GZMB)* expression levels, as a composite measure of tumour cytolytic activity, itself a reflection of the action of CD8⁺ and NK cells. Across cancers, including CRC, increasing tumour mutational count was significantly correlated with increased cytolytic activity. The important contribution of a tumour's overall mutational burden is further reinforced by the active immune microenvironment, and responsiveness to ICI, common to melanoma and lung cancer; two cancers with characterised by large numbers of DNA alterations caused by external mutagens [172]. In CRC and EC, DNA alterations from external mutagens play a limited role in the carcinogenesis process [17]. Rather genomic instabilities, resulting from disruption of processes involved in DNA maintenance and repair, are the principal cause of higher mutation counts. CRC tumours with MMRd or *POLE* mutations have very high mutation counts and typically exhibit increased tumour T-cell infiltrate [87, 90]. In EC, *POLE* proofreading domain mutations, and to a lesser extent MMRd, are associated with an increased immune response [67]. This will be discussed in depth in Chapter 4.

In addition to neoantigens, there is evidence that aberrantly expressed self-antigens in cancer may be capable of eliciting an immune response. Specific examples of this are T-cell based immunity to the cancer testis antigen (CTA) NY-ESO1 in myeloma and sarcoma [173-176], and MAGE-A in melanoma [177-179]. Across TCGA cancers, the role of CTAs was found to be limited [171], but inducing CTA specific immune responses may represent a promising therapeutic avenue.

Alternatively, it is recognised that a number of malignancies are caused by oncogenic viruses, which are capable of eliciting highly specific CTL based immunity. Across all TCGA cancers, Rooney and colleagues found that limited evidence viral infection, except for particular cancer types. Notably, these were human papilloma virus (HPV) infection in cervical cancer, Epstein Barr virus in gastric cancer and hepatitis B in liver cancers. Consistent with our current understanding of the aetiology of CRC and EC, there is a limited role of oncogenic viruses in determining immunogenicity in either of these malignancies [171].

1.5.2. Immune escape

As mentioned above, the anti-tumour immune response exerts a continual selection pressure on cancer cells. Reaching the escape stage of the immunoediting necessitates cancers acquiring mechanisms of evading immune destruction [180]. The adaptive immune response to cancer is driven by recognition of cancer specific neoantigen [92]. In their analysis of TCGA cancers, Rooney and colleagues analysed the actual number of neoantigens observed compared with the expected number of neoantigens derived from a model based on the non-silent mutation rate [171]. Interestingly, their results showed significant depletion of neoantigens seen in CRC, consistent with immune elimination of the most immunogenic cells. Furthermore, their results revealed that tumours with a strong adaptive immune response, as defined by high CYT levels, showed an enrichment for mutations in antigen presentation and extrinsically induced apoptosis pathways. Across all cancers, including CRC, the most enriched immune escape alteration was *CASP8* loss of function mutations. A mechanistic explanation for this enrichment is the associated capacity to evade TRAIL-induced apoptosis; in mice it has been shown that *Casp8* knockout confers resistance to CTL activity [181]. The second most

frequently mutated gene was *B2M*, an integral part of the MHC class I antigen presentation pathway that is essential for CTL immunity [182]. These mutations were more frequent in tumours with high CYT. Similarly, mutations in *HLA-A* and *HLA-B*, also components of the MHC class I pathway, were enriched in tumours with high levels of cytotoxic activity. This pattern of immune escape mutations was confirmed in a subsequent CRC specific study, which showed MSI CRCs had an enrichment for *B2M* (11% of tumours), *HLA-A* (39%) of and *HLA-B* (33%) [183]. Taken together these results suggest that adaptive immunity selects for cancer cells harbouring genetic mutations that abrogate antigen presentation and/or immune induced apoptosis.

There also exist non-genetic mechanisms of immune evasion, and perhaps the best characterised are immune checkpoints. Immune checkpoints are features of normal physiology, which contribute to maintaining peripheral and central immune tolerance. The discovery of immune checkpoints began with the mechanistic description of T-cell driven cytotoxicity, which was made possible by advancements in molecular biology in the late 1980s/early 1990s [12]. Almost simultaneously, CTLA-4 was discovered in mice, a member of the immunoglobulin superfamily, it was recognised to bind with high affinity to CD28 [184]. However, its actual function remained controversial until the development of a CTLA-4 null mice conclusively showed its crucial role in regulating T-cell based immunity. PD-1 was discovered in the early 1990s by a group of Japanese based researchers who were screening for genes involved in apoptosis [185]. Later work showed that PD-1 exerted immunomodulatory effects through binding with the PD-L1 present on APCs [186].

From our current knowledge of cancer immunity, and the process of immunoediting, it is clear that acquisition of immune evasion mechanisms is an essential step in carcinogenesis. However, aside from those described above, our understanding of immune evasion in cancer remains limited.

1.5.3. Host genetics

Variations in the tumour immune landscape between patients can be explained not only by differences in cancer's genetic makeup, but also by the inherited host germline line variants. A large, trans-ethnic

meta-analysis demonstrated that levels of circulating leukocytes are influenced by single nucleotide polymorphisms (SNPs) [187]. There is also evidence of heritable factors influencing the adaptive immune response. In two large studies of adult female twins, it was found that the functioning of adaptive immune cells, in particular of B-cells, CD4⁺ and CD8⁺ T-cells, was determined to a greater degree by genetic factors rather than environmental [188, 189].

Correspondingly there is also evidence for the role of host genetics in defining the immune landscape in cancer. For example, in melanoma immunomodulatory expression quantitative trait loci (eQTL) that are significantly associated with better overall survival were identified [190]. Expansion of these findings was provided in a comprehensive analysis of 30 non-haematological cancers by Sayaman and colleagues [191]. Their results showed that approximately 20% of tumour immune trait variation could be explained by inherited variants. Furthermore, it was suggested that the abundance of CTLs, NK and T_{FH} cells, and the degree of IFN- γ signalling were the traits most strongly influenced by germline factors. The relevance of CTLs and IFN- γ signalling will be highlighted in Chapters 3 and 4 respectively. Identification of germline variants that participate in shaping the immune microenvironment will not only further out understanding of cancer biology but may also be directly translatable to clinical practice. This is evident in the reported associations between polymorphisms in immune related genes such as *HLA* [192], *CTLA4* [193], *IL2* and *IL21* [194] and response to ICI treatment.

1.5.4. The gut microbiome

Aside from host and cancer genetics, perhaps the most pertinent influencer of the tumour immune microenvironment is the gut microbiome, which refers to the genetic composition of all the microbes that exist within the human gastrointestinal tract. It has been implicated in a number of human diseases, including cancer. The term microbiome is used in contrast to microbiota, which describes the microbes themselves. Whilst consisting of approximately 100 trillion micro-organisms, the gut microbiota is primarily comprised of two major phyla, *Bacteroidetes* and *Firmicutes* [195]. Its

specific relevance to cancer relates to two main areas: cancer development, particularly CRC, and influencing response to immune checkpoint therapies.

Alterations in the composition of the gut microbiome have been associated with various stages of CRC development, from early polyp formation to metastatic disease. Significantly increasing abundance of *Fusobacterium nucleatum* has been found with increasing disease stage, from carcinoma in situ to late-stage disease [196]. Moreover, higher levels of *Atopobium parvulum* and *Acintomyces odontolyticus* were found in patients with multiple polyps compared with a solitary lesion [196]. Evidence supporting a causative role for *F. nucleatum* has been derived from experimental models: it has been shown to increase CRC growth *in vitro* [197], and accelerate CRC xenograft growth in mice [198]. It is postulated that dysbiosis may contribute to CRC development through a variety of different mechanisms, including production of bacterial associated carcinogenic metabolites, epithelial-to-mesenchymal transition (EMT) or by immune dysregulation with/without inducing chronic inflammation [75].

In addition to influencing cancer development, the microbiome also appears to be integral in determining response to cancer immunotherapies, namely ICIs. Early work in mice models suggested that antibiotic treatment abrogated anti-tumour immunity [199]. Further to this, it was shown that the mouse microbiome was altered by anti-CTLA-4 and PD-L1 treatment: specifically causing increased levels of *Bacteroidale* and *Burkholderiales*, and decreased abundance of *Clostridiales*. Administration of broad-spectrum antibiotic prior to ICI treatment resulted in decreased efficacy of the treatment. But this effect could be reversed by pre-feeding the mice oral *Bacteroides fragilis*, and *Bacteroides thetaiotaomicron* or *Burkholderia cepacia* [200, 201]. These results are relatable to human disease. It was found that responders to anti-PD1 therapy had significantly higher numbers *Clostridiales*, *Ruminococcaceae*, and *Faecalibacterium* compared with non-responders [75, 202]. The microbiome may represent an exciting future therapeutic target, for both cancer prevention and treatment.

1.6. Clinical implications of the tumour immune microenvironment

Detailed analysis of the cellular elements that comprise the tumour immune landscape have raised the prospect of identifying immune-based prognostic biomarkers. The first demonstration of the prognostic utility of tumour infiltrating lymphocytes was demonstrated in gastric cancer almost 100 years ago [12], but the last two decades have seen a rapid expansion in this area. Immune based biomarkers may serve to compensate for deficiencies in current prognostication strategies, such as TNM staging, to improve patient stratification and aid in identifying those most at risk of disease recurrence and in need of further monitoring and/or treatment. The potential of immune based biomarkers has perhaps been best studied in CRC, and this will be discussed in greater depth in Chapter 3. In particular, work has focused on the prognostic association of infiltrating T-lymphocytes, including CD3⁺ and CD8⁺ cells [110]. Studies published to date have agreed on the positive prognostic association of CD8⁺ and CD3⁺ cells; however, the clinical translation of this work has been limited by the failure to use clinical trial study populations, and, for the majority, to adjust for key confounding variables, such as MMRd status. The most advanced effort to develop an immune based biomarker is the “Immunoscore”, which is prognostic tool based on IHC quantification of CD3⁺ and CD8⁺ cells in the tumour centre and invasive margin [148]. Whilst its latest iteration did include adjustment for key variables (not-published at the time of Chapter 3 completion), its clinical use remains limited by forced dichotomisation for analysis.

Harnessing the immune system to treat cancer is not a novel concept. However, the field of immuno-oncology has seen a great renaissance in recent years with the discovery of immune checkpoints, as described above. The therapeutic implications of these checkpoint discoveries were demonstrated in early mice models. Treatment with anti-PD-L1 antibody resulted in reduced tumour formation and metastases, as well as augmented T-cell mediated tumour destruction [203]. These promising pre-clinical results were soon translated into trials of human therapies, which began with treatment of metastatic melanoma by CTLA-4 blockade in 2011 [204] and followed shortly after by anti-PD1 treatment in non-small cell lung cancer (NSCLC) [205]. The results of these treatments have been

remarkable for a minority of patients: durable results lasting long after treatment cessation have been observed in approximately 20% of melanoma patients [206]. Importantly, ICI responsive seems to be dependent on the pre-treatment immune landscape. Increased responsiveness is associated with a “hot” immune microenvironment characterised by increased levels of CTLs and counter balanced by increased immune checkpoint expression [15, 111]. As described above, increased tumour immunogenicity is associated with higher levels of somatic DNA mutational burden. Intuitively, this would suggest that ICIs would be ineffective in unselected CRC and EC patients, as the majority of these cancers have a relatively low mutational burden. Clinical trials proved this to be broadly true, although responses have been observed in metastatic non-hypermutated ECs [207]. However, later studies concentrated the treatment of CRC patients (and some EC patients) with MMRd cancers, which produced far more promising results (discussed in greater depth in Chapter 4) [208]. Unfortunately, it remains the case that only a minority of patients overall respond. As such, immunotherapies that will be effective irrespective of pre-treatment tumour immunity are highly desirable. Examples of potentials include cancer vaccines, and CAR T-cell therapy. Whilst there have been some exciting results, particularly for CAR T-cell therapy [209], their use is not yet widespread.

1.7. Summary and outstanding questions

Cancer is a disease of a disordered genome, which enables uncontrolled cell proliferation, tissue invasion and metastatic spread [10]. Whilst it has long been suggested that the immune response has important implications for cancer development and progression, research advances in the last four decades have now demonstrated that it is an essential part of cancer biology. To this extent, immune evasion is now a recognised central hallmark of cancer [9]. Furthermore, we better understand that prognosis and response to treatment is in part determined by the composition of the tumour immune microenvironment, and that the tumour immune microenvironment is a function of the interaction between the immune system and the tumour genome. Furthering our understanding of this essential relationship is a vital research endeavour.

There are many key questions that remain to be addressed; I will attempt to answer a select few in the research conducted during my DPhil and presented in this thesis:

- Previous work has highlighted that quantification of tumour infiltrating immune cells in CRC holds promise as a prognostic biomarker. However, it has not fully shown whether this prognostic association is independent of underlying genomic and molecular. Furthermore, previous studies have not shown whether immune cell quantification as a continuous variable provides relevant prognostic information.
- The consequences of MMRd have been relatively well characterised in CRC, making it an important prognostic and predictive marker for this disease, but far less so in EC. It is vital to understand whether this form of genomic has the same consequences in cancers of different tissues, particularly in regards to shaping the anti-cancer immune response. This will have important implications for our understanding of the genetic-immune interaction, and also in dictating the clinical relevance of genomic instability
- Much of our understanding of the relationship between immune response and tumour genetics is derived from observational, correlative work on human samples. At present, there are no studies reporting murine models that would enable work to improve our functional understanding of this relationship. The development of mice models that will allow delineation of specific aspects of this immune-genetic interaction is critical. This will help to investigate key issues, such as mechanisms of immune evasion.

2. Methods

2.1. TCGA data analysis

2.1.1. RNAseq data

2.1.1.1. FASTA file download and preparation

Prior to RNAseq analyses, it was necessary to download and pre-process raw FASTA files. The reason for doing so was to avoid batch effect that may have been introduced by other downstream pipelines [210]. FASTA files from TCGA were downloaded using the ‘TCGA Biolinks’ package (version 3.13, <https://bioconductor.org/packages/release/bioc/html/TCGAbiolinks.html>).

2.1.1.2. RNAseq quantification

To produce transcript- level quantification for RNAseq, the package ‘Salmon’ (<https://combine-lab.github.io/salmon/>). This is a computationally efficient package that employs the use of quasi-mapping to achieve accurate and rapid RNAseq quantification [211]. Results are produced as Transcripts Per Million (TPM).

2.1.1.3. Differential gene expression analysis

Following transcript level quantification, RNAseq analyses were then performed using the DESeq2 package (<https://bioconductor.org/packages/release/bioc/html/DESeq2.html>) (Love). DESeq2 tests for differential gene expression using a negative binomial distribution.

2.1.2. TCGA tumour mutational analysis

2.1.2.1. MC3 MAF files

MC3 (Multi-Centre Mutation Calling in Multiple Cancer) MAF files were downloaded from (<https://gdc.cancer.gov/about-data/publications/mc3-2017>) [212]

2.1.2.2. *Mutational frequency analysis*

Analysis of mutational frequency and production of “Oncoprint” plot (Figure 4.14) was done in R using the ‘maftools’ package (<https://bioconductor.org/packages/release/bioc/html/maftools.html>).

2.2. Cell culture

2.2.1. Cell culture conditions and media

Details of cell culture experiments, including cell lines and media used are provided in Section 4.2.5.

2.3. General histology and tissue staining

All histological worked described in Chapter 4 was conducted at the Pathology Department, Leiden University Medical Centre (LUMC), Leiden, Netherlands. Histological procedures involving murine tissue were carried out at the Wellcome Centre Human Genetics, University of Oxford.

2.3.1. Tissue storage

All FFPE tissue was stored at room temperature in appropriate conditions. Cut FFPE slides were stored at 4°C if not processed after cutting. All human tissues were stored in compliance with the Human Tissue Act 2004, and under Oxfordshire Research Ethics Committee B (Approval No 05\Q1605\66).

2.3.2. Tissue processing and sectioning

Formalin-fixed tissues and organs were processed by automated system and embedded. Tissue blocks were then sectioned at 4µm thickness using a Leica microtome

2.3.3. Haematoxylin and eosin staining

Haematoxylin and Eosin (H&E) staining of cut FFPE slides was conducted manually. FFPE slides underwent rehydration in xylene (2x5 mins) and then graded alcohol (100%, 90% and 70% ethanol) series to deionised water (dH₂O). Slides were then stained with haematoxylin stain for 30 seconds,

before immediately being rinsed in running water until clear. The slides were then dipped in Acid alcohol (1% HCl & 70% EtOH) briefly before emersion in dH₂O. The slides were then stained with eosin, and rehydrated through dH₂O, graded ethanol series to xylene. The slides were then mounted with DPX before being viewed by light microscopy

2.3.4. Immunohistochemistry

FFPE slides underwent rehydration in xylene (2x5 mins) and then graded alcohol (100%, 90% and 70% ethanol) series to dH₂O. Endogenous peroxidase activity was quenched with 3% hydrogen peroxide for 15 mins at room temperature before washing twice in 1x Phosphate buffered saline with 0.1% Tween (0.1% PBST). Antigen retrieval was achieved using a Citrate Buffer solution 1M (pH 6.0) in a pressure cooker at 125°C for 5 minutes. The slides were then cooled to room temperature in ice water for 20 minutes, before washing 3X in PBST and blocking with 5% serum (same source as secondary antibody) at room temperature for one hour. Slides were then incubated overnight at 4°C with primary antibody at the optimised concentration. Following washes, slides were incubated for 2 hours with species-specific secondary horseradish peroxidase conjugated antibody at 1:200 concentration. Signal amplification was performed by the Vectastain kit (Vector Laboratories) and labelling visualised with 3,3-diaminobenzidine (DAB), (Vector laboratories) as per manufacturer's instructions before counterstaining with haematoxylin (Vector laboratories).

2.3.4.1. Antibodies used for IHC based quantification of T-cells in CRC

The antibodies used for T-cell quantification in CRC TMAs (Chapter 3) are detailed in Table 2.1

Table 2-1 Details of the antibodies used for T-cell quantification human CRC tissue

Target	Host Species	Target species	Manufacturer (clone)	Optimised conc.
CD8	Rabbit	Human	Leica Biosystems (4B11)	Premade formulation
CD3	Rabbit	Human	Abcam (EPR19514)	1:100

2.3.4.2. Antibodies used for IHC immunoprofiling of mouse FFPE samples

The antibodies used for immunoprofiling of mouse samples (Chapter 5) are detailed in Table 2.2

Table 2-2: *Details of the antibodies used for immune cell quantification in murine FFPE tissues*

Target	Host Species	Target species	Manufacturer (clone)	Optimised conc.
CD8a	Rabbit	Mouse	Cell Signalling Technology (D4W2Z)	1:800
CD4	Rabbit	Mouse	Abcam (EPR19514)	1:500
Beta-2 Microglobulin	Rabbit	Mouse	Abcam (EPR21752-214)	1:2000
CD19	Rabbit	Mouse	Cell Signalling Technology (D4V4B)	1:400

2.3.4.3. Image analysis

Stained FFPE mouse tissues slides were viewed by brightfield microscopy and images obtained by Leica Aperio slide scanner, imager and software (Leica, United Kingdom). Positive marker cell analysis was conducted using QuPath and details are provided in the relevant Chapters.

2.4. SDS-PAGE electrophoresis and immunoblotting

2.4.1. Cellular protein lysate collection

The cells for Western Blot analysis were first washed with ice cold PBS and then lysed in RIPA buffer with protease and phosphatase inhibitors (Pierce Halt Cocktail, ThermoFisher). The lysates were then kept on ice if they were going to be used immediately, or stored at -20°C long term.

2.4.2. Protein lysate quantification

Protein lysates concentration was determined using the CB-XTM Protein Assay (GBiosciences). For each quantification, a standard curve was calculated from a six-protein standard of proteins concentrations: 0, 0.2, 0.4, 0.6, 0.8 and 1mg/ml. Protein lysate samples were diluted 1:10 before mixing 50µl with 1ml of CB-X assay. Following incubation at room temperature for 1 hour, the samples' absorbance was read at 595nm wavelength.

2.4.3. Western Blot

30µg protein from of each sample was separated using NuPAGE 3 to 8%, Tris-Acetate gels (Invitrogen) and blotted onto PVDF membrane (Millipore) by wet transfer. Membranes were blocked with 5% milk in 0.1% TBST for one hour, incubated with primary antibody overnight at 4°C and 1 hour at room temperature; washed with TBST x 3, incubated with loading control antibody to GAPDH for 1 hour at room temperature, washed x 3 and incubated with anti-mouse IgG/anti-rabbit IgG as appropriated 1:15000 in 5% milk at room temperature for 1 hour. Membranes were then imaged using the LI-COR imaging system.

2.4.4. Antibodies used Western Blot experiments

Details of the primary (Table 2.3) and secondary antibodies (Table 2.4) used in Western Blotting.

Table 2-3. Details of the primary antibodies used in Western Blot experiments

Target	Host Species	Target Species	Manufacturer (clone)	Optimised conc.
JAK1	Mouse	Human	BD Biosciences (AB_397627)	1:500
STAT1	Mouse	Human	Thermo Fisher Scientific (AHO0832)	1:500
Phospho-STAT1	Mouse	Human	Thermo Fisher Scientific (33-3400)	1:500
HLA Class I Heavy Chain	Mouse	Human	Nordic MUBio (MUB2037P)	1:1000
GAPDH	Rabbit	Human	Cell Signalling Technology (5174)	1:1000

Table 2-4. Details of the secondary detection antibodies used in Western Blot experiments

Target	Host Species	Target Species	Manufacturer (clone)	Optimised conc.
Mouse IgG	Rabbit	Mouse	LI-COR Bioscience (LIC-926-68070)	1:1500
Rabbit IgG	Mouse	Rabbit	LI-COR Biosciences (LIC-925-68071)	1:1500

2.5. Mice

2.5.1. Mutations and background strains

Details of the mice used in this work are detailed in Chapter 5.3.3

2.5.2. Animal husbandry

Further details are provided in Chapter 5.3.1. All mice were housed in the animal unit of the Wellcome Centre for Human Genetics, Roosevelt Drive, University of Oxford. The mouse work was conducted under Project licence no. PDF0B94C3 and Personal licence: 12D1D1096

2.5.3. DNA Extraction

DNA was extracted from mouse ear tissue by the hot sodium hydroxide and Tris (HotSHOT) protocol (Truett et al). HotSHOT protocol DNA extraction was preformed by adding 50µl of Alkaline lysis reagent (25 mM NaOH, 0.2 mM EDTA, pH 12.0) to mouse tissue sample (obtained by ear clipping), and the mixture incubated at 95°C for 1 hour. 50µl of Neutralising agent (40 mM Tris-HCl) was then added, before the samples were centrifuged and stored at -4°C before further processing

2.5.4. Genotyping

2.5.4.1. Polymerase Chain Reaction (PCR)

Genotyping was preformed using the primers and reagents listed below. All reactions were preformed using a MJ research DNA Tetrad 2 Thermal Cycler.

2.5.4.2. Gel electrophoresis

PCR products were separated and visualised by gel electrophoresis: 2.0% agarose in 0.5M TBE (Tris-borate- EDTA) buffer was heated in the microwave for 2 minutes and cooled under running water. 2µg/ml of ethidium bromide was added and the solution allowed to set. 2µl of PCR marker (Bio-Rad) was loaded and 8µl of PCR sample under TBE buffer. The gel was run at 120V for 20 minutes, then visualised by transillumination by UV light with a Syngene Gene Genius.

2.5.4.3. *Mlh1* floxed allele PCR conditions

Primers:

Primer name	Sequence (5'-3')
5armWTF:	GGACCTAGTAGATGGTATCCAGGTT
CritWTR	ATGCTGCTACCACAGGGA

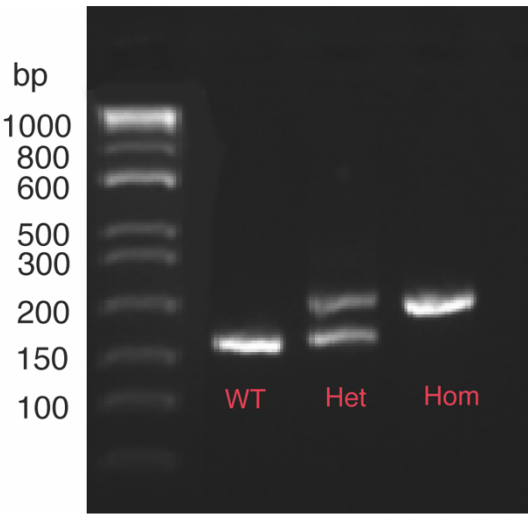
Table 2-5 *Mlh1* floxed allele PCR reagents

Reagent	Working Conc.	Volume (μ l)
dH ₂ O		10.5
Buffer (10x)	1	2.5
MgCl ₂ (50mM)	3mM	0.75
dNTPs	0.2mM	2
WT F Primer	1 μ M	1
WT R Primer	1 μ M	1
Q	1 X	5
Taq	0.2U/ μ l	0.25
Total volume (μ l)		23
DNA sample		2

PCR conditions

Step	Temp (°C)	Time (minutes)
1	95	5
2	94	1
3	55	1
4	72	1
5		Repeat 2-4 for 40 cycles
6	72	5
7	4	Hold indefinitely

Figure 2-1. **Gel electrophoresis for the products of *Mlh1* floxed PCR:** *Mlh1* WT (wildtype) *Mlh1* Het (Heterozygous) and *Mlh1* Hom (Hom)



2.5.4.4. *Apc* floxed allele PCR conditions

Primers:

Primer name	Sequence (5'-3')
<i>Apcf1</i> F	GCCATCCCTTCACGTTAG
<i>Apcf1</i> R	TTCCACTTTGGCATAAGGC

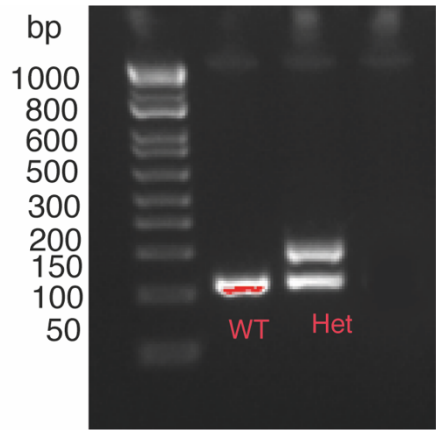
Table 2-6 **APC floxed allele PCR reagents**

Reagent	Working Conc.	Volume (μ l)
dH ₂ O		10.5
Buffer (10x)	1	2.5
MgCl ₂ (50mM)	3mM	0.75
dNTPs	0.2mM	2
WT F Primer	1 μ M	1
WT R Primer	1 μ M	1
Q	1 X	5
Taq	0.2U/ μ l	0.25
Total volume (μ l)		23
DNA sample		2

1434 **PCR conditions**

Step	Temp (°C)	Time (minutes)
1	95	5
2	94	1
3	55	1
4	72	1
5		Repeat 2-4 for 40 cycles
6	72	5
7	4	Hold indefinitely

1441 *Figure 2-2. Gel electrophoresis for the products of Apc flox PCR: WT (wildtype) and Het (heterozygous)*



2.5.4.5. *Vil-CreERT2 floxed allele PCR conditions*

Primers:

Primer name	Sequence (5'-3')
Cre F	CGGTCGATGCAACGAGTGATGAGG
Cre R	CCAGAGACGGAAATCCATCGCTCG

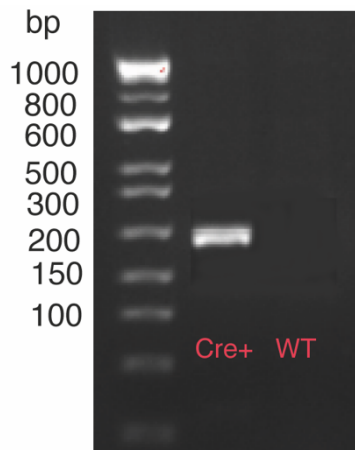
Table 2-7 Mlh1 floxed allele PCR reagents

Reagent	Working Conc.	Volume (μ l)
dH ₂ O		11.5
Buffer (10x)	1	2.5
MgCl ₂ (50mM)	3mM	0.75
dNTPs	0.2mM	2
WT F Primer	0.5 μ M	0.5
WT R Primer	0.5 μ M	0.5
Q	1 X	5
Taq	0.2U/ μ l	0.25
Total volume (μ l)		23
DNA sample		2

PCR conditions

Step	Temp (°C)	Time (minutes)
1	95	5
2	94	1
3	55	1
4	72	1
5		Repeat 2-4 for 40 cycles
6	72	5
7	4	Hold indefinitely

Figure 2-3. **Gel electrophoresis for the products of Cre PCR:** Cre + (Heterozygous Vil-CreERT2 and WT (wildtype)



2.5.5. Mouse Intestinal adenoma quantification

After mice intestinal tissue had been formalin fixed for 24 hours and then contained in 70% for a further 24 hours, intestinal adenomas were counted. The intestinal tissue was briefly placed in methylene blue stain for 30 seconds and then rinsed in PBS. The polyps were then counted with light microscopy. Polyps were recorded according to their size and location

2.6. Statistical analysis

All statistical analysis was conducted using R, Version 4.0.1 (<https://cran.r-project.org>). Details of specific R packages used and statistical tests applied are detailed in the relevant chapters.

2.6.1. Graph preparation

All graphs shown in this thesis and submitted with manuscripts were prepared initially in R with package “ggplot2” (<https://cran.r-project.org/web/packages/ggplot2>) and edited in Adobe Illustrator.

3. The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

3.1. Introduction

CRC is a significant cause of morbidity and mortality worldwide [19]. It causes approximately 250,000 deaths each year in Europe alone [54]. CRCs are staged according the AJCC TNM staging system (explained in the Introduction; Table 1.2), which guides prognostication and treatment strategies [49]. This is a composite score derived from pathological assessment of the primary tumour, evidence of lymph node involvement and the presence/absence of distant metastases [50]. Prognosis correlates with disease stage, and fortunately, the majority of patients are diagnosed at an early stage. Patients with stage II (primary tumour stage (pT) 3 or 4 without spread to regional lymph nodes or distant metastases) or stage III (any pT stage with spread to regional lymph nodes but no metastatic disease) are generally managed by surgery with curative intent in the first instance. This is on the whole successful, with post resection 5-year survival rates for stage II and stage III disease of 68-83% and 45-65% respectively [50]. As mentioned previously (Introduction 1.1.5), to reduce recurrence risk postoperatively, all stage III patients and a subset of stage II defined with high-risk characteristics (Table 1.3) are offered adjuvant chemotherapy (a fluoropyrimidine $\pm$ oxaliplatin) provided they are fit for this treatment [49]. It is estimated that fluoropyrimidine monotherapy reduces the absolute risk of death by 3-5% in high risk stage II and, and 10-15% in stage III, with oxaliplatin based agents contributing a further 4-5% reduction [49].

The use of pathological staging described above does not fully delineate the substantial variation in the clinical outcomes of stage II and III CRCs, and is alone insufficient to accurately stratify patients according to their risk of recurrence. Consequently, many patients with stage II/III CRC who have been cured by surgery alone receive needless cytotoxic chemotherapy from which they do not benefit, and which puts them at risk of harmful side effects/toxicities and death. Additionally, patients who are destined to recur despite standard adjuvant chemotherapy (and who may be candidates from trials of

novel therapies) are not reliably identified. Furthermore, apart from MMRd there is a paucity of markers to identify patients who might benefit from novel immunotherapies such as immune checkpoint inhibitors (ICIs), adoptive T cell transfer or targeted cancer vaccines. At present, ICIs are the best and most widely used immunotherapy agent. In the UK, they are currently only licensed for the treatment of metastatic CRC [53]. However, there is evidence from CTLA-4 and PD-L1 blockade in melanoma that these therapies may also be beneficial in the non-metastatic setting; one-year of ICI treatment was shown to reduce post-resection recurrence risk for stage III melanoma patients [213, 214].

Importantly, ICI treatment appears to work best in tumours with a pre-existing anti-tumour immune response [13] and as such, it is logical that MMRd status would identify those patients who are likely to respond. However, in reality only approximately 50% of patients with MMRd metastatic colorectal cancer actually respond to PD-1 inhibition [52]. Thus, there is a clear, and urgent, need for further biomarkers to improve prognostication and prediction of therapeutic benefit.

3.1.1.Previous Work

As noted in the Introduction, a substantial, and increasing, body of evidence demonstrates that the immune system plays a critical role in tumour development and progression [97]. The prognostic implication of tumour infiltrating lymphocytes in cancer was first noted 100 years ago, in gastric carcinoma [12]. Over the last two decades, there have been considerable efforts to understand the relationship between particular populations of tumour infiltrating lymphocytes and clinical outcome in CRC patients. The cell populations examined in these studies are outlined in Table 3.1. Early work highlighted the possible prognostic association between CD4⁺ and CD8⁺ T-cells and clinical outcome in CRC [143, 215]. This was expanded upon by Galon and colleagues, who first demonstrated that tumours with high levels of CD45RO⁺ cells (effector memory T-cells) had lower risk of metastatic disease, and improved overall survival [146]. This was followed by a study looking at T-cell infiltration in 415 patients. The levels of CD3⁺, CD8⁺ and CD45RO⁺ cells were quantified at both the tumour centre (CT) and invasive margin (IM). Cancers were classified as either “Hi” or “Lo”

depending on whether the density of the T-cell infiltrate in either the CT or IM was greater than the median. Patients with any of the three T-cell marker densities above the median in the CT, IM or both were shown to have a greater overall survival [153]. This work provided the basis for the “Immunoscore”, an immunohistochemistry (IHC) based prognostication tool. In the latest iteration, total T-cell and cytotoxic T-cell populations were quantified by staining for CD3 and CD8 in the tumour centre and invasive margin. Using a mean from the four measurements, two markers in two tumour compartments, tumours were scored as “Immunoscore Low” (immune cell density in the 0-25th centile), “Immunoscore Intermediate” (25th -70th centile) or “Immunoscore High” (70th -100th centile). Studies of the Immunoscore are only a few of many that have examined the association between tumour T-cell infiltrate and clinical outcome in CRC. The key results of these and from selected additional publications are summarised in Table 3.2.

Table 3-1. Summary of the T cell populations and markers most commonly used in CRC prognostic studies.

Marker	Lymphocyte population	Description	Refs.
CD3	All T-cells	CD3 ⁺ common marker to all T-cell populations	[146, 153]
CD8	Cytotoxic T-lymphocytes (CTLs)	CTLs protect against intracellular pathogens (viruses and bacteria), as well destroy malignant cells. Predominantly activated via antigen MHC I antigen presentation	[148, 153]
CD4	T-helper cells	Essential for initiating and shaping almost all forms of adaptive immunity; primarily activated via MHC II pathway	[216]
FoxP3	T regulatory cells	Suppress and regulate other immune cells, which maintains peripheral tolerance and prevents autoimmunity/chronic inflammation	[163, 164, 217]
CD45RO	Memory T-cells	Provide antigen specific T-cell memory. Responsible for initiating immune response against previously encountered antigens	[146, 147]

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

Table 3-2. Summary of selected previous studies examining T cell infiltrate and clinical outcome in CRC.
OS: Overall Survival. DFS: Disease Free Survival. RFS: Recurrence Free Survival. IM: Invasive Margin. CT: Centre of Tumour. IHC: Immunohistochemistry. MV: multivariable

Marker quantification	Study Population	Clinical Trial	MMR adjusted for	Analysis method	Key results	Ref
Effector memory T-cells (CD3 ⁺ CD45RO ⁺) by IHC	959 patients; stage I – IV CRC.	No	No	Median dichotomisation	CD3 ⁺ CD45RO ⁺ cells associated with absence of early metastatic disease and prolonged survival	[146]
CD3 ⁺ , CD8 ⁺ and CD45RO ⁺ at IM and CT by IHC	415 patients; stage I – IV	No	No	Median dichotomisation	Increased densities of all markers associated with Improved DFS	[153]
Tumour cytotoxic T-cells (CTL) quantified by gene expression analysis	1203 patients; stage I-IV	No	No	“Overexpressed” tumours compared with non-overexpressed	CTLs not significantly associated with survival	[218]
CD3 ⁺ and CD8 ⁺ T-cells at IM by IHC	573 patients; stage II CRC.	No	No	Dichotomisation at lower tertile	Worse RFS and OS in tumours with low CD3 ⁺ /CD8 ⁺ infiltrate	[219]
CD8 ⁺ , CD3 ⁺ and FoxP3 ⁺ by IHC	555 patients; stage I – IV CRC	No	No	Optimal cut point dichotomisation	CD8 ⁺ positively associated with OS in MV analyses only in right sided tumours. FoxP3 ⁺ independently associated with OS in rectal tumours only	[220]
CD8 ⁺ by IHC and IF	1162 patients	No	No	Median dichotomisation	CD8 ⁺ low tumours associated with worse OS	[221]
CD3 ⁺ and CD8 ⁺ by IHC	215 patients; stage I – IV	No	No	Semi-quantitative marker scoring; grouped into low and high infiltrated tumours	CD8 ⁺ and CD3 ⁺ significantly associated with OS.	[222]
CD8 ⁺ and CD4RO ⁺ by IHC	694 patients; stage I-IV	No	Yes	“Immunoscore”: value of 0-4 based on the densities. “I4” corresponds to the CD8 and CD45RO densities in the highest quartiles in both CT & IM	Immunoscore significantly associated with improved DSS and OS; better predictor of survival and recurrence than MSI status	[223]
CD3 ⁺ and CD8 ⁺ by IHC at IM and CT	799 patients; stage I – IV	No	Yes	“Immunoscore”: value of 0-4 based. “I4” corresponds to the CD8 and CD3 densities in the highest quartiles CT & IM	Higher Immunoscore significantly associated with improved DFS, DSS and OS	[224]
CD3 ⁺ and CD8 ⁺ by IHC at IM and CT	2861 patients; stage I – IV	No	Yes	“Immunoscore”: value of 0-4 based. “I4” corresponds to the CD8 and CD3 densities in the highest quartiles CT & IM	Higher Immunoscore significantly associated improved DFS	[148]
CD3 ⁺ and CD8 ⁺ by IHC at IM and CT	600 patients; stage III	Yes	Yes	“Immunoscore”: value of 0-4 based. “I4” corresponds to the CD8 and CD3 densities in the highest quartiles CT & IM	Higher Immunoscore significantly associated with improved DFS	[144]
CD3 ⁺ and CD8 ⁺ by IHC at IM and CT	1,322 patients; stage III	Yes	Yes	“Immunoscore”: value of 0-4 based. “I4” corresponds to the CD8 and CD3 densities in the highest quartiles CT & IM	Higher Immunoscore significantly associated with improved DFS	[225]

3.1.2. Limitations of previous work

Whilst previous studies have demonstrated a potential role for tumour infiltrating lymphocytes as a biomarker in CRC, there are limitations to these bodies of work. Firstly, they have mostly been conducted using non-experimental cohorts, which pose a much greater risk for bias than clinical trial series [226]. Secondly, whilst a small number of previous studies have suggested that the prognostic association of tumour infiltrating lymphocytes (TILs) persists after adjustment for MMRd [148], they have not adjusted for other molecular variables that are associated with both clinical outcome and antitumour immune response. For instance, CIN is associated with a reduced T cell infiltrate and poorer prognosis in cancer [227, 228], whilst *POLE* proofreading domain mutations confer a more vigorous immune response, particularly T-cell response, and are associated with better clinical outcomes [90]. Thirdly, it has not been addressed whether the prognostic association is uniform across all CRCs, or whether it varies according to tumour stage, or other key molecular factors, such as *BRAF* and *KRAS* mutation status. Mutations in both of these genes have been associated with worse clinical outcome in non-MMRd CRCs [229, 230]. Moreover, KRAS has been shown to have immunomodulatory effects [231]. Mice models of pancreatic cancer have shown that *Kras* knockout has immunosuppressive effects [232], and in human CRC, *KRAS* mutations are associated with reduced immune cell infiltration and checkpoint expression [233, 234]. Finally, previous work has largely approached outcome analyses by forced dichotomization of TIL levels. Whilst superficially attractive to clinicians, this approach discards important information [235], which may lead to an underestimation of the variation in outcome between groups. Furthermore, dichotomising tumours according to any given cut point reduces the sensitivity of a prognostic marker. For instance, dichotomisation at a median cut point would mean that two tumours with immune cell densities at the 49th and 51st centile would be classified as “low” and “high” respectively. Despite the densities differing by just two centiles, they would be predicted to have different clinical outcomes, and as such may receive radically different treatment regimes. In the most recent iteration of Immunoscore, this risk has been somewhat mitigated by producing an aggregate score from the quantification of multiple markers in different tumour compartments [110, 236]. Additionally, dichotomisation risks missing non-linear associations between variables of interest and clinical outcomes [235, 237].

3.1.3. Rationale for study and aims for chapter

There is a pressing need for biomarkers to improve the management of patients with early-stage CRC. Immune cell-based biomarkers hold considerable promise in this regard; however, due to limitations of previous research highlighted above this has yet to be fully realised. This study aimed to address some of these issues. Using a large clinical-trial patient cohort, I wanted to study the prognostic association of CD8⁺ infiltrating lymphocytes in early stage (stage II and III) CRC. The detailed clinicopathological and molecular annotation enabled multivariable analyses with adjustment for key confounding variables. In addition, it was possible to assess whether the prognostic association is uniform across all early-stage CRC. Finally, CD8⁺ cell infiltrate density was analysed as a continuous variable to avoid some of the pitfalls incurred by forced dichotomisation.

The primary and secondary aims for this study were to assess the prognostic association between CD8⁺ cell infiltrate, measured as a continuous variable, with time to recurrence (TTR) and overall survival (OS) in the pooled VICTOR and QUASAR2 clinical trials. Exploratory analyses were to include examining the association according to clinical, pathological and molecular features.

3.2. Methods

3.2.1. Trial cohorts used in this study

This study was a retrospective analysis of two large clinical trials: the QUASAR2 and VICTOR trials [238, 239]. The QUASAR2 trial examined the addition of bevacizumab to oral capecitabine as adjuvant therapy for stage II/III patients who had had curative-intent surgery [238]. Bevacizumab is humanised, monoclonal antibody that targets vascular endothelial growth factor, and had been shown to be effective in treating metastatic disease [240]. In total, 1,952 patients were recruited for this trial and 1,941 included in the final intention-to-treat analysis [238]. In the VICTOR trial, stage II/III CRC patients who had had curative-intent surgery with or without adjuvant chemotherapy were randomised to receive either the addition of rofecoxib, a selective cyclooxygenase (COX) 2 inhibitor, or placebo to standard treatment. An initial 2,434 patients were recruited. However, the trial was terminated early because of the global withdrawal of rofecoxib, at which point 2,327 patients had been enrolled (1,167 and 1,160 in the treatment and control arms respectively) [239]. The exclusion for QUASAR2 and VICTOR trials are shown in Table 3.3.

3.2.2. Inclusion of cases for biomarker study

Tissue microarrays (TMAs) are constructed of punch samples taken from tumour blocks. Both trials consented patients for optional donation of FFPE tumour blocks for research. The TMAs were constructed from FFPE blocks collected from 764 participants in the VICTOR trial and 1,144 in the QUASAR2 trial. The included cases were all those with tissue available in TMA format and clinical outcome data.

The association between tumour-infiltrating CD8+ lymphocytes and clinical outcome in stage II/III colorectal cancer

QUASAR2
History of cancer
Inflammatory bowel disease
Active poetical ulcer requiring treatment in previous 2 years
Lack of physical integrity of upper GI tract
Malabsorption syndrome
Inability to take oral medication
Moderate or severe renal impairment (creatinine clearance <30mL/min)
Absolute neutrophil count < 1.5 x 10 ⁹ cells per L
Platelet count < 100 x 10 ⁹ cells per L
Total bilirubin > 1.5 times upper limit of normal
Alanine aminotransferase, aspartate aminotransferase or alkaline phosphatase > 2.5 times upper limit of normal
Proteinuria worse than 500mg per 24 hours
Previous chemotherapy, immunotherapy or infradiaphragmatic radiotherapy
Use of any investigational drug, agent or procedure within 4 weeks of randomisation
Chronic use of full dose anticoagulants, high dose aspirin, antiplatelet drugs or known bleeding diathesis
Concomitant treatment with sorivudine or its chemically related analogues
History of uncontrolled seizures, CNS disorders or psychiatric disorders that precluded giving informed consent
Clinical important cardiac disease
Known coagulopathy
Known allergy to Chinese hamster ovary cell proteins
Pregnancy, lactation or no use of contraception in premenopausal women
VICTOR
Peptic ulceration or GI bleeding in previous year
Receiving long term nonsteroidal anti-inflammatory drug therapy (except for low-dose aspirin)
Younger than 18 years old
History of cancer
Inflammatory bowel disease
Severe congestive heart failure
Pregnancy, lactation or no use of contraception in premenopausal women

Table 3-3. ***Multiple linear regression analysis of the relationship between clinicopathological/molecular characteristics and CD8+ cell density in the combined QUASAR2 and VICTOR trial samples. CNS – Central Nervous System***

3.2.3. Determination of tumour T cell infiltrate by immunohistochemistry

CD8⁺ and CD3⁺ cellular infiltrate was determined in the cancer tissue samples by immunohistochemistry (IHC). FFPE slides underwent deparaffinisation, antigen retrieval and blockade of intrinsic peroxide activity, before overnight incubation at 4°C with CD8 antibody (Leica Biosystems PA0183, mouse clone 4B11) and CD3 (Leica NCL-L-CD3-565, mouse clone LN10, diluted 1:100). For the VICTOR trial, CD8⁺ cell quantification was conducted in triplicate cores, whilst for QUASAR2 CD3⁺ and CD8⁺ cells were quantified in duplicate cores. The IHC was conducted and slides scanned the Clinical Diagnostic Histopathology Laboratory, UCL Hospital, London.

Quantification of marker-positive and negative cell numbers was done by computerized image analyses using ImmunoPath 1.3.9.0 (Room4, Crowborough, UK). This method of quantifying marker positive cells has been previously published and validated [241]. The ImmunoPath software platform allows for selection and annotation of scanned CRC slides. Each TMA is broken into smaller squares of interest for analysis. ImmunoPath is trained to recognise marker positive and negative cells, and produce a final count for both. In previous work, the accuracy of ImmunoPath was found to be comparable to manual counting techniques [241]. Automated quantification of CD8⁺ and CD3⁺ was carried out by the group of Havard Danielsen at Oslo University, Norway. In addition to basic quality control (QC) conducted by the Danielsen group, I performed detailed QC of all the stained TMA cores. This included checks to see if they contained at least 1000 cells in total, and visual inspection to ensure no loss of tumour material following immunostaining.

3.2.4. Tumour molecular analysis

Molecular annotation of the VICTOR and QUASAR2 trials had been previously conducted and reported prior to commencement of this work [90, 238, 242]. For tumour molecular analysis, tumour DNA was extracted from FFPE samples. MSI status was determined using the standard five marker Bethesda panel (BAT25, BAT26, D2S123, D5S346, and D17S250). Cases were defined to be MSI if $\geq$

2/5 (40%) markers were positive. CIN status was determined by through preparation of FFPE sections and quantification of nuclear DNA content by Ploidy Work Station Grabber. *POLE* proofreading domain mutations were defined as recurrent variants known to cause an ultramutated phenotype and an enrichment of associated mutational signatures. In the VICTOR trial, *POLE* mutational status was determined by sequencing of the entire exonuclease domain whereas in QUASAR2, targeted sequencing of hotspot exons 9,13 and 14 was performed. For *KRAS* mutations, direct DNA sequencing was performed for exon 12,13 and 61. For *BRAF* mutations, codon 600 was sequenced.

3.2.5.Validation series

To externally validate results, independent patient series were sought. Potential series for inclusion were either widely known (such as TCGA) or found by searching the NCBI Gene Expression Omnibus (GEO). All series selected for inclusion were required to have documentation of primary tumour and nodal stage, *CD8A* expression and clinical outcome.

3.2.6.Statistical analysis

The REMARK (Reporting Recommendation for Tumour Marker Prognostic Studies) guidelines [243] provide a gold standard for studies such as these; all analyses and reporting were conducted in accordance with these. Comparison of the baseline demographic, clinicopathological and tumour molecular factors between trials were assessed by unpaired student's t-test, non-parametric Mann-Whitney test, or Fisher's exact test according to whether the data were continuous or categorical, normally distributed or not. The primary and secondary endpoints of interest were the association of CD8⁺ density, analysed as a continuous variable, with TTR of CRC (defined as the time from randomisation to CRC relapse, with censoring at last contact or death in case of no recurrence), and OS (measured as the time from randomization to death from any cause, with censoring at date of last contact in patients still alive) respectively. For the multivariable models, predefined covariates were included on the basis of known prognostic value, and included: age, sex, disease stage, pT4 primary and MMRd/*POLE*-mutation. Additionally, clinicopathological variables that were independently

associated with CD8⁺ cell density in this study population were also included: primary tumour location and *BRAF* mutation. Missing data were defined as data not stored for the variable of interest [244]. Overall, the rate of missing data in the QUASAR2/VICTOR cohort was low (a maximum of 11.6% for any single co-variable). Logistic regression was used to determine the pattern of missing data, and was suggestive of a pattern of missing at random (MAR). As such, it was possible to impute the missing data by multiple imputation by chained equations with predicted mean matching, using the MICE package in R [245].

3.2.7. Ethical approval

For the VICTOR trial, the protocol was reviewed by the Cancer Research Campaign, the multicentre Research Ethics Committee and research ethic committees at participating centre. Ethical approval for the QUASAR2 trial was granted by the West Midlands Research Ethics Committee (Edgbaston, Birmingham, UK; REC reference: 04/MRE/11/18). Both trials were conducted in accordance with the Declaration of Helsinki. Ethical approval for this biomarker study was provided by North West – Liverpool Central Research Ethics Committee (Liverpool, UK; REC reference: 17/NW/0252).

3.3.Results

3.3.1.QUASAR2 and VICTOR patient details

The CONSORT diagram showing the flow of cases included in this work is shown in Figure 3.1, and the baseline demographic variables of the QUASAR2 and VICTOR trial cases used in this study are shown in Table 3.4. In total, there were 1804 cases included in this study based on the availability of tissue in TMA format and clinical outcome data. Importantly, comparison of included cases with non-included did not reveal statistically significant differences in patient age, disease stage, sex or disease recurrence (Table 3.4). In QUASAR2 there was a significant difference in number of deaths between included cases and non-included cases, though this was not the case for VICTOR (Table 3.4).

The two trials did not differ significantly in terms of baseline patient characteristics (age and sex), molecular features (MMR, *POLE* mutation and CIN status) or outcome results (disease recurrence events and death) (Table 3.5). As such, the two trials were pooled for purposes of analysis. *POLE* mutations are rare in CRC (1-2% cases), and given that they are associated with an increased immune response and a similarly favourable prognosis to MMRd [90], these two molecular disease subtypes were combined into a ‘hypermutated’ group for analysis.

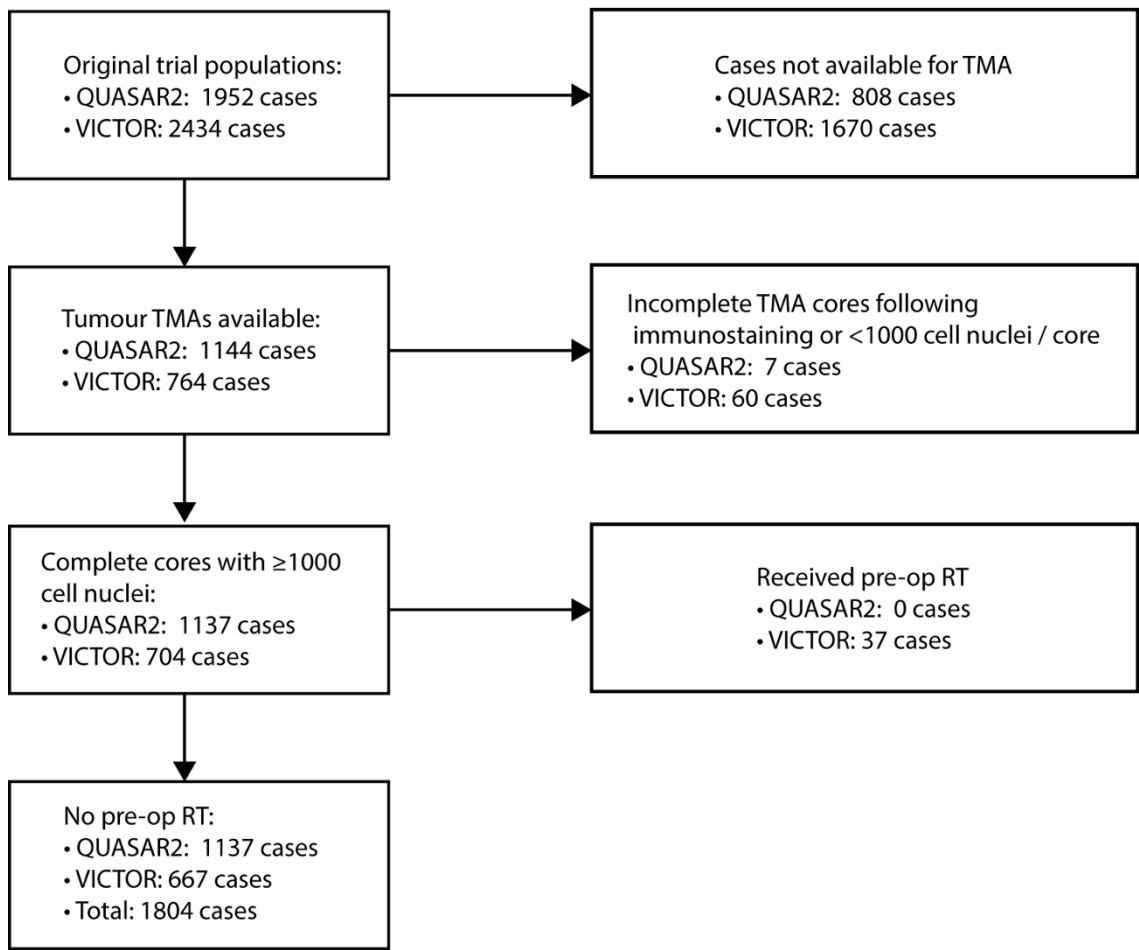


Figure 3-1. CONSORT diagram showing the inclusion and exclusion of CRC cases from QUASAR2 and VICTOR trials. RT: radiotherapy. TMA: tissue microarray (reproduced from Glaire et al, 2019)

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

	VICTOR		QUASAR2		Combined		P (VICTOR vs. QUASAR2)
	No.	%	No.	%	No.	%	
Total	667	100	1137	100	1804	100	
Age (years)							
Median (range)	64.9 (24.6 – 89.1)		65.0 (21.0 – 85.0)		65.0 (21.0 – 89.1)		0.2 ⁺
Sex							
Male	426	63.9	657	57.8	1083	60.0	0.12 ⁺
Female	241	36.1	480	42.2	721	40.0	
Unknown	0	0	0	0	0	0	
Disease stage							
II	314	47.1	394	34.7	708	39.2	2.1 x 10 ^{-7†}
III	353	52.9	743	65.3	1096	60.8	
Unknown	0	0.0	0	0	0	0	
Primary tumour stage							
pT1-3	520	78.0	719	63.2	1239	68.7	3.7 x 10 ^{-13†}
pT4	134	20.1	418	36.8	552	30.6	
Unknown	13	1.9	0	0	13	0.7	
Primary tumour location							
Right	284	42.6	446	39.2	730	40.5	0.7 ⁺
Left	368	55.2	604	53.1	972	53.9	
Unknown	15	2.2	87	7.7	102	5.7	
MMR status [†]							
MMRp	492	73.2	932	82.0	1424	78.9	0.81 [†]
MMRd	73	10.9	145	12.8	230	12.1	
Unknown	102	15.3	60	5.3	162	9.0	
POLE status [†]							
POLE wild-type	477	71.5	1090	95.9	1567	86.7	0.76
POLE-mutant	4	0.6	8	0.7	12	0.7	
Unknown	186	27.9	39	3.4	225	12.5	
Chromosomal instability							
CIN high	370	55.5	679	59.7	1049	58.2	0.3 [†]
CIN low	178	26.7	372	32.7	550	30.5	
Unknown	119	17.8	86	7.7	205	11.4	
KRAS							
Wild-type	377	56.5	697	61.3	1074	59.5	0.8 [†]
Mutant	186	27.9	335	29.5	521	28.9	
Unknown	119	17.8	105	9.2	209	11.6	
BRAF							
Wild-type	509	76.3	903	79.4	1412	78.3	0.1 [†]
Mutant	58	8.7	136	12.0	194	10.8	
Unknown	100	15.0	98	8.6	198	11.0	
Chemotherapy							
No	253	37.9	0	0.0	253	14.0	2.2 x 10 ^{-16†}
Yes	414	62.1	1137	100	1551	86	
Unknown	0	0.0	0	0.0	0	0.0	
Bevacizumab							
No	667	100	555	48.8	1222	67.7	2.2 x 10 ^{-16†}
Yes	0	0	582	51.2	582	32.3	
Unknown	0	0	0	0	0	0	
Rofecoxib							
No	330	49.5	1137	100	1467	81.3	2.2 x 10 ^{-16†}
Yes	337	50.5	0	0	337	18.7	
Unknown	0	0	0	0	0	0	
Disease Recurrence							
No	505	75.7	864	76.0	1369	75.9	0.9 [†]
Yes	162	24.3	273	24.0	435	24.1	
Unknown	0	0	0	0	0	0	
Death							
No	542	81.3	912	80.2	1454	80.6	0.6 [†]
Yes	125	18.7	225	19.8	350	19.4	
Unknown	0	0	0	0	0	0	

Table 3–4. Comparison of the baseline characteristics of the QUASAR2 and VICTOR trial cases used in this biomarker study. pT: pathological tumour (T) stage; MMR: DNA mismatch repair; MMRp: mismatch repair proficient; MMRd: mismatch repair deficient; POLE-mutant: pathogenic POLE exonuclease domain mutation; KRAS-mutant: KRAS mutation at codons 12,

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

1712 13 or 61. BRAF-mutant: BRAF hotspot V600E mutation. *determined by unpaired Student's t-test. †determined by Fisher exact
1713 test (in cases which marker status was determined). Reproduced from Glaire et al 2019 [1]

1714

	VICTOR (included in study)		VICTOR (not included)		P value (included vs. not- included)	QUASAR2 (included in study)		QUASAR2 (not included)		P value (included vs. excluded)
	No.	%	N	%		No.	%	No.	%	
Total	667	100	1767	100		1137	100	791	100	
Age (years)										
Median (range)	64.9 (24.6 – 89.1)		65 (25 – 89)		0.92*	65.0 (21.0 – 89.1)		65.0 (23.0 – 88.0)		0.46*
Sex										
Male	426	63.9	1132	64.1	1.0 [†]	657	57.8	442	55.9	0.35 [†]
Female	241	36.1	635	35.9		480	42.2	349	44.1	
Unknown	0	0	0	0		0	0	0	0	
Disease stage										
II	314	47.1	853	48.3	0.30 [†]	394	34.7	150	19.0	0.34 [†]
III	353	52.9	914	51.7		743	65.3	641	81.0	
Unknown	0	0.0	0	0		0	0	0	0	
Chemotherapy										
No	253	37.9	627	35.5	0.54 [†]	0	0.0	0	0	1.0 [†]
Yes	414	62.1	1140	64.5		1137	100	791	100	
Unknown	0	0.0	0	0		0	0.0	0	0.0	
Bevacizumab										
No	667	100	1767	100	1.0 [†]	555	48.8	412	52.1	0.14 [†]
Yes	0	0	0	0		582	51.2	379	47.9	
Unknown	0	0	0	0		0		0	0	
Rofecoxib										
No	330	49.5	875	49.5	0.47 [†]	1137	100	791	100.0	1.0 [†]
Yes	337	50.5	892	50.5		0	0	0	0	
Unknown	0	0	0	0		0	0	0	0	
Disease Recurrence										
No	505	75.7	1313	74.3	0.96 [†]	864	76.0	631	79.8	0.052 [†]
Yes	162	24.3	454	25.7		273	24.0	160	20.2	
Unknown	0	0	0	0		0	0	0	0	
Death										
No	542	81.3	1415	80.1	0.87	912	80.2	667	84.3	0.02 [†]
Yes	125	18.7	352	19.9		225	19.8	124	15.7	
Unknown	0	0				0	0	0	0	

1715 Table 3-5. Comparison of baseline characteristics of VICTOR and QUASAR2 cases included in biomarker
1716 study vs. cases not included. *determined by unpaired Student's t-test. †determined by Fisher exact test (in
1717 cases which marker status was determined). Reproduced from Glaire et al 2019 [1]

1718

1719 3.3.2. Immune cell infiltrate

1720 3.3.2.1. Quantification of T cell infiltrate

1721 Slides containing TMA cores were stained for expression of the pan-T cell marker CD3 or the
1722 cytotoxic T cell marker CD8 by immunohistochemistry as detailed in Section 3.2.3. Stained slides
1723 were scanned and transferred electronically to the laboratory of Havard Danielsen (Oslo) for analysis.
1724 For each TMA core, the total number of marker positive cells (CD3⁺ or CD8⁺ for QUASAR2; CD8⁺
1725 only for VICTOR) and marker negative cells were quantified by automated analysis by ImmunoPath
1726 software by Prof Danielsen's group (Figure 3.2). The quantification process begins with defining a

sub-area of interest within the tissue segment. Cells are then classed as either positive or negative on the basis of HSV (Hue, Saturation and Value) threshold. A process of refinement is conducted, which involves selecting and deselecting positive and negative cells until an acceptable level of true positives/negatives are achieved (determined by visual review). The automated counts were additionally validated by comparing to manual counting with Spearman rank correlation and Bland-Altman tests [241]. Case level data files were returned to my group for analysis by me. During initial review of the results, two possible ways of analysing the intensity of intratumoral T cell infiltrate were considered:

- (i) cell density (calculated as the number of marker positive cells/area of tissues)
- (ii) cell fraction (number of marker positive cells/total number of cells) (the total number of cells was determined from the total number of nuclei counted for each TMA sample)

There was a small but systematic difference in core size between the cores size in the two trial cohorts. This was corrected for by using cell density per mm² in the metric used for analyses.

Interestingly, during preliminary analysis of the data, I noted that the median CD8⁺ cell density differed substantially between cases from the QUASAR2 trial population (231.6 cells/mm² ± 288.4) and those from the VICTOR trial (145.0 ± 188.0) (KS test $P < 0.0001$). While the fraction of CD8⁺ cells also differed between the trial cohorts, this was far less marked (Kolmogorov-Smirnov test $D = 0.078$, $P = 0.0087$). The explanation for this discordance was not immediately obvious (eg shrinkage of VICTOR cores due to their greater age would be expected to increase marker density). In light of this information and to minimise variation between cohorts, I decided to use marker positive cell fraction for downstream analyses, bearing in mind this variation when considering the statistical models to be used.

Prior to beginning analysis of the CD3⁺ and CD8⁺ cell infiltrate data, I wanted to conducted baseline quality control. This firstly involved visual inspection of every TMA core to ensure that tissue loss had not occurred; those cores with significant tissue loss were excluded (27 cases from VICTOR trial and 32 cases from QUASAR2 trial) (Figure 3.1). Secondly, quantification results were checked to confirm

that each core had least 1000 cells, to ensure each TMA core was a representative sample. This resulted in the removal of an additional 70 from VICTOR trial and 58 cases from the QUASAR2 trial (Figure 3.1)

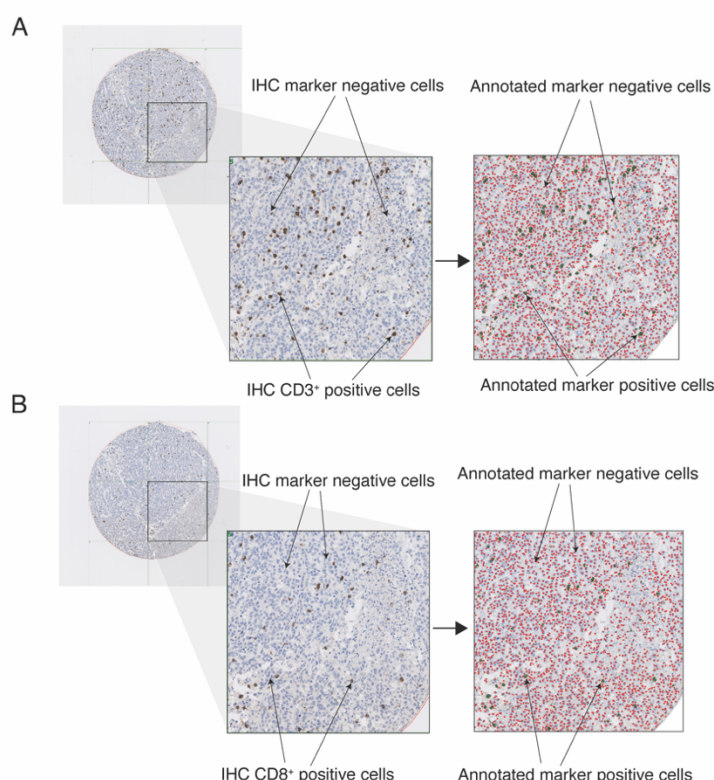


Figure 3-2: T-cell infiltrate quantification. Representative image showing IHC based quantification of CD3⁺ (A) and CD8⁺ (B) in TMA sections using the ImmunoPath automated software. For analysis, each histopathological image is divided into subsections, and the software is trained to recognise marker positive and negative cells. Total number of positive and negative cells are given per unit area

Variation in the CD8⁺ cell infiltration between cores of the same patient was not assessed in this piece of work. This presents an interesting opportunity for future research; first to understand the extent of the variation in immune cells infiltration within individual cancers, and how this variation impacts upon clinical outcomes. This will be assessed in the larger body of work investigating immune cell infiltrate in the QUASAR2 and SCOT trials. For these trials, three cores per patient and five cores per patient respectively have been analysed.

1768

1769 3.3.2.2. Transformation of markers for analysis

1770 Initial analysis of the data distribution revealed that the raw CD3⁺ and CD8⁺ cell fraction in the
 1771 QUASAR2 and VICTOR cohorts showed a significantly negatively skewed distribution
 1772 (skewness=2.46) (Figure 3.3A). Similar non-normality is well recognised in analyses of gene
 1773 expression, where data are routinely log₂ transformed prior to analysis. As such, cell fraction was log₂
 1774 transformed to prior to further analyses, as this approximated a normal distribution, and enabled
 1775 downstream statistical analysis (Figure 3.3B)

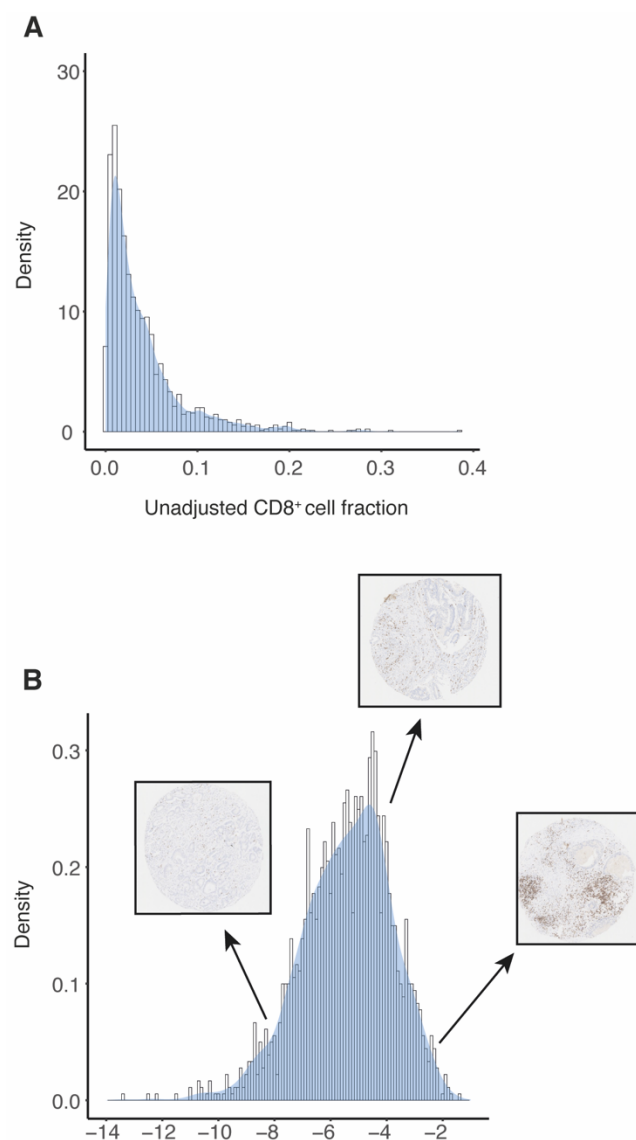


Figure 3-3 CD8⁺ cell infiltration in the VICTOR and QUASAR2 combined cohort. Histogram with overlaid kernel density plot showing unadjusted CD8⁺ cell fraction in the VICTOR and QUASAR2 combined cohort (A) and the Log₂ transformed CD8⁺ cell fraction with representative images (B)

3.3.2.3. Comparison of the prognostic utility of CD3⁺ and CD8⁺ cell quantification in QUASAR2 trial samples

The Immunoscore developed by Galon and colleagues arguably represents the most advanced effort to develop a T-cell based prognostic biomarker in CRC. In its latest iteration, a score is generated based on the combination of CD3⁺ and CD8⁺ positive cell densities at the invasive margin and tumour centre [148]. However, the individual contribution of each of these markers in predicting CRC prognosis is unclear. Furthermore, the importance of quantifying cell density in both of these tumour compartments has not been formally established. As such, initial analyses in this study sought to address these uncertainties. In the QUASAR2 TMA cores, which were all taken from the tumour centre, both CD3⁺ and CD8⁺ cell infiltrates were quantified. They were found to be strongly positively correlated (Spearman rho 0.65; $P < 2.2 \times 10^{-16}$). To investigate the relationship between T-cell infiltrate at the invasive margin and tumour centre a set of 51 full faced tumour sections from the QUASAR2 cohort were stained for CD8⁺ cells. Importantly, CD8⁺ cell density at the invasive margin and tumour centre was also strongly positively correlated (Spearman rho 0.73, $P = 1.38 \times 10^{-9}$). This raised the possibility that at least some of the information captured by Immunoscore's use of two compartments and two markers may be redundant.

3.3.2.4. Preliminary analysis of CD3, CD8 fraction and clinical outcome in QUASAR2 study

To investigate the association between intratumoral CD3 and CD8 infiltrate and clinical outcome, univariable analyses were conducted with CD3 alone and CD8 alone, and bivariable analyses preformed with both markers. In the univariable models, CD8⁺ cell fraction was significantly associated with time to recurrence (TTR), and this association remained significant in a bivariable model that included CD3⁺ (Table 3.6). However, CD3⁺ fraction was not significantly associated with TTR in either univariable or multivariable analyses. Furthermore, the CD8⁺ only model had a similar discriminatory ability and smaller Akaike's Information Criterion (AIC) value (an indicator that allows comparison of the quality of statistical models) compared to the bivariable model (Table 3.6).

Given that CD8⁺ cell infiltrate alone was a superior predictor of tumour recurrence than CD3⁺ cell infiltrate, and that the addition of CD3⁺ to the prognostic model provided no apparent improvement, it was decided that further analyses would concentrate on CD8⁺ cells alone. Furthermore, the strong correlation between CD8⁺ infiltrate at the invasive margin and tumour centre suggested that the use of TMAs cores taken from the tumour centre was likely to be acceptable for purposes of quantification.

Variable	HR (95% CI)	P	C index	AIC
Model 1. CD8⁺ cell density only (920 cases, 225 events)				
CD8 ⁺ cell density (log2 transformed)	0.90 (0.83–0.98)	0.011	0.549	2975.5
Model 2. CD3⁺ cell density only (920 cases, 225 events)				
CD3 ⁺ cell density (log2 transformed)	0.92 (0.83–1.03)	0.17	0.531	2980.1
Model 3. CD8⁺ and CD3⁺ cell density (920 cases, 225 events)				
CD8 ⁺ cell density (log2 transformed)	0.90 (0.81–0.99)	0.030	0.549	2977.5
CD3 ⁺ cell density (log2 transformed)	1.01 (0.88–1.16)	0.88		

Table 3-6 Preliminary analyses of CD3⁺ and CD8⁺ cell fraction in the QUASAR2 trial samples. AIC: Akaike's Information criterion. P value calculated by Cox proportional hazards model

3.3.2.5. Relationship between CD8⁺ cell infiltrate and other parameters in QUASAR2 and VICTOR cohorts

To identify variables to be included in the multivariable models, the association between CD8⁺ cell infiltrate and other features of the QUASAR2 and VICTOR CRCs was initially assessed by univariable analyses (Table 3.7). CD8⁺ cell fraction was significantly greater in stage II than stage III disease (3.0% positive cells vs 2.2, $P=1.5 \times 10^{-5}$ determined by non-parametric Mann-Whitney test), but not with pT stage (2.4% positive cells vs 2.7, $P=0.15$, non-parametric Mann-Whitney test). Furthermore, CD8⁺ was associated with right-sided tumour location (3.1 vs 2.2, $P=4.1 \times 10^{-6}$, Mann-Whitney test), MMRd/*POLE* mutation (2.4 vs 4.1, $P=4.9 \times 10^{-12}$, Mann-Whitney test), lack of chromosomal instability (CIN) (2.9 vs 2.4, $P=1.2 \times 10^{-3}$, Mann-Whitney test), *BRAF* mutation (2.4 vs 3.1, $P=3.2 \times 10^{-3}$, Mann-Whitney test), and absence of disease recurrence (2.7 vs 2.0, $P=7.3 \times 10^{-5}$, log rank test) and death (2.7 vs 2.0, $P=3.2 \times 10^{-3}$, log rank test). There was no significant association with age, sex, or *KRAS* mutation observed (Table 3.7).

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

1825
1826

	No.	Median CD8 ⁺ cell fraction, % (IQR)	P*
Total	1804	2.5 (0.01–38.5)	–
Age (years)			
<65	902	2.6 (1.1–5.2)	0.06
≥65	902	2.4 (1.1–4.8)	
Sex			
Male	1083	2.5 (1.1–4.9)	0.43
Female	721	2.6 (1.1–5.2)	
Disease stage			
II	708	3.0 (1.3–5.4)	1.5 × 10 ⁻⁵
III	1096	2.2 (1.0–4.7)	
Primary tumour stage			
pT1-3	1239	2.4 (1.1–4.9)	0.15
pT4	552	2.7 (1.2–5.4)	
Unknown	13	2.6 (0.9–3.9)	
Location			
Right	730	3.1 (1.3–6.0)	4.1 × 10 ⁻⁶
Left	972	2.2 (1.0–4.6)	
Unknown	102	2.5 (0.9–4.8)	
MMR & POLE status [‡]			
MMRp & POLE wild-type	1412	2.4 (1.0–4.7)	4.9 × 10 ⁻¹²
MMRd or POLE-mutant	230	4.1 (1.9–7.8)	
Unknown	162	2.6 (1.2–6.0)	
Chromosomal instability			
CIN low	550	2.9 (1.3–5.4)	1.2 × 10 ⁻³
CIN high	1049	2.4 (1.0–4.8)	
Unknown	205	2.8 (1.2–6.0)	
KRAS			
Wild-type	1074	2.5 (1.1–4.9)	0.70
Mutant	521	2.5 (1.1–5.0)	
Unknown	209	2.6 (1.2–6.0)	
BRAF			
Wild-type	1412	2.4 (1.1–4.8)	3.2 × 10 ⁻³
Mutant	194	3.1 (1.5–5.9)	
Unknown	198	3.1 (1.2–6.1)	
CRC recurrence			
No	1369	2.7	7.3 × 10 ⁻⁵ ∇
Yes	435	2.0	
Death			
No	1454	2.7	3.2 × 10 ⁻³ ∇
Yes	350	2.0	

Table 3-7. **Relationship between clinicopathological/molecular characteristics and tumour CD8⁺ cell density in the combined QUASAR2 and VICTOR trial population.** IQR: interquartile range. pT : primary tumour stage. POLE-mutant: pathogenic POLE mutation. MMR: mismatch repair MMR-p: mismatch repair proficient. MMR-d: mismatch repair deficient. KRAS-mutant: KRAS mutation at codons 12, 13 or 6. BRAF-mutant: BRAF hotspot V600E mutation * P value determined by non-parametric Mann-Whitney test. ∇Determined by log-rank test.

Given the complex interrelationship of these factors, I further examined this by multiple linear regression. This revealed that only tumour location (Effect = -0.25, $P=7.2 \times 10^{-3}$) and MMRd/*POLE* (0.52, $P=3 \times 10^{-4}$) mutation status were statistically significant predictors of CD8⁺ cell infiltrate (Table 3.8).

	Linear Regression Results	
	Effect (95% CI)	P
Age (continuous)	-0.11 (-0.27 – 0.04)	0.15
Sex		
Male	0.0 (ref)	–
Female	0.04 (-0.13 – 0.21)	0.63
Location		
Left	0.0 (ref)	–
Right	-0.25 (-0.43 – -0.07)	7.2×10^{-3}
Stage		
II	0.0 (ref)	–
III	-0.17 (-0.34 – 0.00)	0.05
Primary tumour		
pT1-3	0.0 (ref)	–
pT4	0.00 (-0.18 – 0.18)	0.98
<i>BRAF</i> mutation		
Wild-type	0.0 (ref)	
Mutant	0.0 (-0.29 – 0.27)	0.97
MMR & <i>POLE</i> status		
MMR-P & <i>POLE</i> wild-type	0.0 (ref)	–
MMR-D or <i>POLE</i> -mutant	0.52 (0.25 – 0.81)	3×10^{-4}
Chromosomal instability		
CIN low	0.0 (ref)	–
CIN high	-0.05 (-0.24 – 0.14)	0.59

Table 3-8. Multiple linear regression analysis of the relationship between clinicopathological/molecular characteristics and CD8⁺ cell density in the combined QUASAR2 and VICTOR trial samples. *POLE*-mutant: pathogenic *POLE* mutation. MMRd: mismatch repair deficiency. MMRp: mismatch repair proficient. CIN: chromosomal instability. CRC: colorectal cancer. *BRAF*-mutant: *BRAF* hotspot V600E mutation.

3.3.3. Missing Data Imputation

Missing data is an important consideration when analysing large cohorts as it can have a substantial impact on conclusions drawn [244, 246]. Of the 1804, QUASAR2 and VICTOR cases used for the final multivariable analyses, data were missing for covariables of tumour location in 102 cases (5.7%), primary tumour stage in 13 cases (0.7%), *KRAS* mutation in 209 cases (11.6%), *BRAF* mutation in 198 cases (11.0%), MMRd/*POLE* mutation in 162 cases (9.0%) and CIN in 205 cases (11.4%). Before deciding a strategy to handle these missing data it was important to determine the missing data pattern. Logistic regression analysis of the missing data using completely observed variables as covariates was consistent with a pattern of MAR. MAR implies that any systematic differences between the missing values and the observed values can be explained by differences in the observed data. This is in contrast to a pattern of missing not at random (MNAR), which suggests that there are systematic differences between the observed and missing values [247].

The MAR pattern of missing data in QUASAR2 and VICTOR cohorts meant that it was feasible to impute the missing covariate data using multiple imputation by chained equations with predicted mean matching,[248] using the MICE package in R.[245]. This method was chosen because it is able to impute missing variables of different data types (continuous, categorical, ordinal or binary) [249]. Furthermore, it has been suggested that it is the most appropriate imputation methodology for datasets with missing data at a similar frequencies to those observed in this study [248]. A considerable advantage of using multiple imputation, as opposed to exclusion of incomplete cases, is that it prevents loss of statistical precision and power [247]. The multiple imputation models included all available covariates, event status, and the Nelson-Aalen estimate of the hazard function.[249] For the final multivariable models, imputed datasets were pooled using the Hmisc package in R.

3.3.4. CD8⁺ cell infiltrate and clinical outcome

The relationship between CD8⁺ cell fraction and clinical outcome was next analysed by Cox regression in univariable and multivariable models. The use of multivariable models enables delineation of whether the prognostic association of CD8⁺ cell infiltrate in CRC is independent of

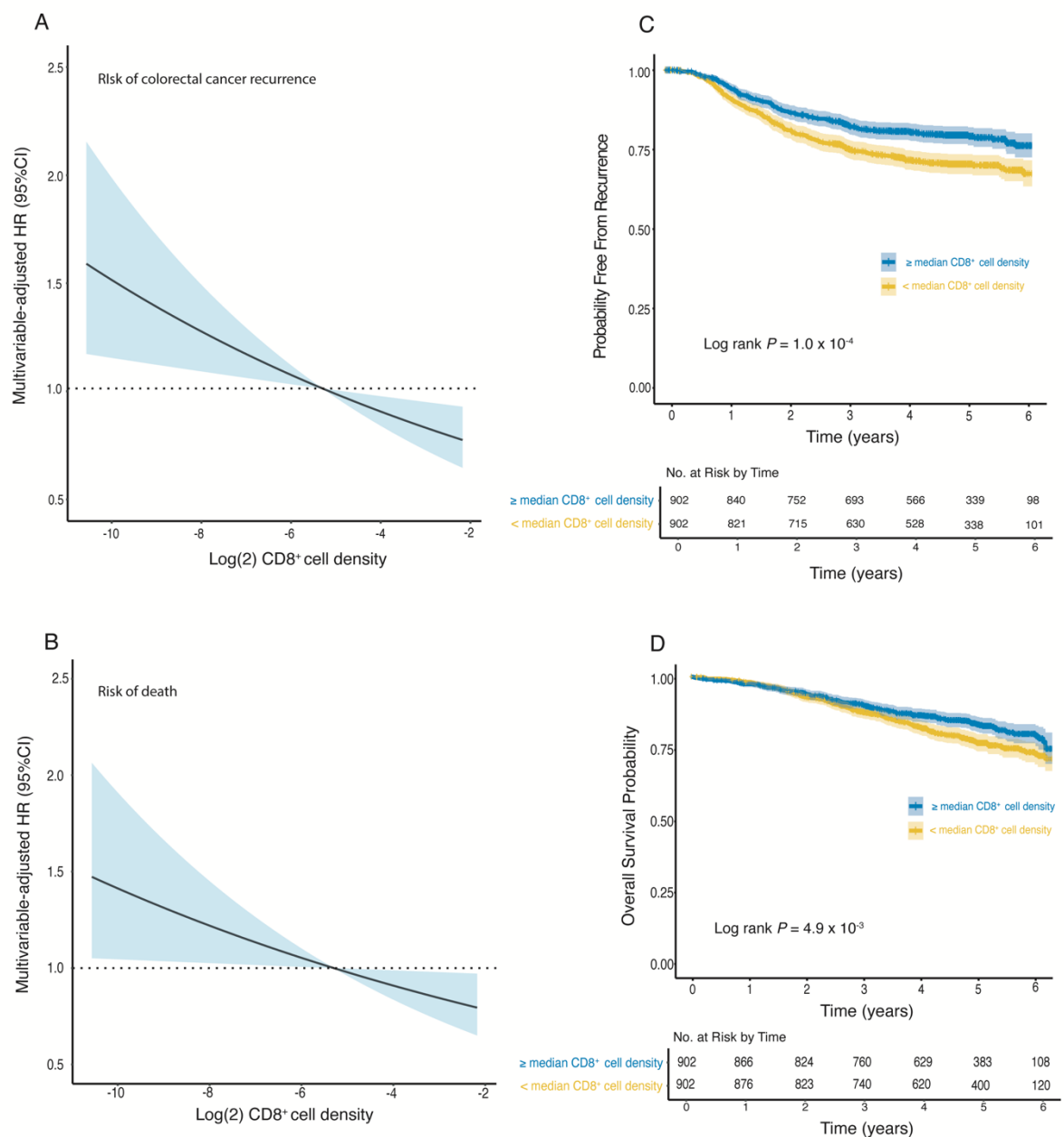
other variables known to be significantly associated with clinical outcome, CD8⁺ cell infiltrate or both. As such, the co-variables included in the multivariable models were those that have previously been shown to be associated with prognosis in CRC, or those which I had shown to be associated with CD8⁺ cell fraction (Table 3.7). Increasing CD8⁺ cell fraction was associated with a significantly longer TTR in univariable analyses (HR=0.90 for each two-fold increase (ie per doubling in CD8 fraction); $P=1.7 \times 10^{-4}$). In keeping with previous research, univariable analyses showed both MMRd and *POLE* mutations were also significantly associated with an improved TTR [90, 95] (HR= 0.68, 95% CI=0.49-0.94, $P=0.022$), and CIN status with a worse TTR (HR= 1.37, 95% CI=1.10-1.71, $P=0.0051$) (Table 3.9) [227].

To build the final multivariable models, the parameters included were either clinical variables of known prognostic significance (age, sex, disease stage, pT4 primary stage and MMRd/*POLE*) or those demonstrated to be significantly associated with CD8⁺ cell infiltrate (*BRAF* and tumour location) (Table 3.7). Interestingly, in the multivariable model, the association between TTR and CIN status, or TTR and MMRd/*POLE* mutations did not remain statistically significant. However, CD8⁺ cell density did continue to be significantly associated with TTR (HR=0.92; 95% CI=0.87–0.97, $P=3.6 \times 10^{-3}$) (Table 3.9 Figure 3.4A). Similarly, CD8⁺ cell infiltrate was significantly associated with overall survival in both univariable and multivariable models (HR=0.91; 95%CI =0.86-0.98, $P=3.2 \times 10^{-3}$ and HR=0.93; 95% CI=0.87–0.99, $P=0.024$ respectively) (Table 3.9; Figure 3.4B). This translated to estimated probabilities of being CRC recurrence-free at 3-years of 76.3% (95% CI= 73.4 – 79.4) and 83.8% (95% CI= 81.4 – 86.3) for tumours with CD8⁺ cell density at the 10th and 90th percentiles respectively.

To allow for comparison with the previously published “Immunoscore”, patients were dichotomised according to a median CD8⁺ cell density cut point and the association with both TTR and OS analysed. The results were plotted as Kaplan Meier curves and compared by the log-rank test. CD8⁺ cell high tumours (greater than the median density) compared with CD8⁺ cell low (less than the median density) were associated with a significantly reduced risk of recurrence (HR =0.71, 95%

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

1887 CI=0.59–0.87, $P=7.2 \times 10^{-4}$) (Figure 3.4C) and improved overall survival (HR =0.71; 95% CI=0.57–
1888 0.88, $P=1.6 \times 10^{-3}$) (Figure 3.4D). These results were comparable to previously published
1889 Immunoscore based analyses of clinical trials. In 600 patients from the NCTG N0147, CD8⁺ high
1890 tumours were at a reduced risk of recurrence (HR = 0.64; 95%CI=0.45-0.92, $P=0.015$) [144] and
1891 similarly in 1,322 patients from the IDEA trial (HR = 0.65; 95%CI=0.52-0.81, $P=0.0001$) [225].



1892 **Figure .3-4. The association between CD8⁺ cell infiltrate and clinical outcome in the VICTOR and**
1893 **QUASAR2 trials.** (A). Multivariable adjusted hazard ratio for risk of recurrence and CD8⁺ cell density in
1894 QUASAR2 and VICTOR (B). Multivariable adjusted hazard ratio for risk of death and CD8⁺ cell density in
1895 QUASAR2 and VICTOR (C). Kaplan Meier curves showing risk of recurrence with patient cohort divided at a
1896 median cut point. (D). Kaplan Meier curves showing overall survival with patient cohort divided at a median
1897 cut point. Reproduced from Glaire et al 2019 [1]

The association between tumour-infiltrating CD8+ lymphocytes and clinical outcome in stage II/III colorectal cancer

	No.	TTR events	OS events	Univariable analysis				Multivariable analysis			
				Time to recurrence		Overall survival		Time to recurrence		Overall survival	
				HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
Age (continuous)	1804	435	350	1.01 (1.00 – 1.02)	0.10	1.03 (1.02 – 1.04)	6.2 x 10 ⁻⁶	1.00 (0.99 – 1.01)	0.43	1.02 (1.01 – 1.04)	1.0 x 10 ⁻⁴
Sex											
Male	1083	281	222	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
Female	721	154	128	0.83 (0.68 – 1.01)	0.063	0.89 (0.72 – 1.11)	0.31	0.83 (0.68 – 1.02)	0.075	0.89 (0.71 – 1.11)	0.30
Location											
Left	972	167	172	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
Right	730	167	158	1.08 (0.93 – 1.31)	0.45	0.77 (0.62-0.96)	0.019	1.05 (0.85 – 1.31)	0.63	0.83 (0.65 – 1.06)	0.13
Stage											
II	708	115	91	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
III	1096	320	259	1.89 (1.52 – 2.40)	6.4 x 10 ⁻⁹	1.85 (1.45 – 2.35)	5.6 x 10 ⁻⁷	1.96 (1.57 – 2.44)	1.34 x 10 ⁻⁹	2.03 (1.59 – 2.60)	1.3 x 10 ⁻⁸
Primary tumour											
pT1-3	1239	248	186	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
pT4	552	191	158	1.96 (1.61 – 2.37)	1.0 x 10 ⁻¹¹	2.12 (1.70 – 2.61)	1.1 x 10 ⁻¹¹	2.14 (1.76 – 2.60)	3.0 x 10 ⁻¹⁴	2.20 (1.77 – 2.74)	1.3 x10 ⁻¹²
BRAF mutation											
Wild-type	1412	325	250					1.0 (ref)			
Mutant	194	55	52	1.33 (1.00 – 1.78)	0.048	1.68 (1.25 – 2.27)	6.5 x 10 ⁻⁴	1.59 (1.18 – 2.17)	2.7 x 10 ⁻³	1.55 (1.12 – 2.16)	8.0 x 10 ⁻³
MMR & POLE status											
MMRd & POLE wild-type	1412	357	272	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
MMRd or POLE- mutant	230	40	42	0.68 (0.49 – 0.94)	0.022	1.00 (0.73 – 1.39)	0.98	0.72 (0.50 – 1.05)	0.090	0.96 (0.65- 1.41)	0.85

1910 **Cont'd**

1911

1912

1913

1914

1915

1916

1917

1918

1919

1920

1921

1922

	No.	TTR events	OS events	Univariable analysis				Multivariable analysis			
				Time to recurrence		Overall survival		Time to recurrence		Overall survival	
				HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
Chromosomal instability											
CIN low	550	110	88	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
CIN high	1049	278	214	1.37 (1.10 – 1.71)	0.0051	1.27 (0.99 – 1.62)	0.063	1.17 (0.93 – 1.49)	0.18	1.21 (0.94 – 1.58)	0.14
Bevacizumab treatment											
No	1222	279	223	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–	1.0 (ref)	–
Yes	582	156	127	1.31 (1.03 – 1.67)	0.027	1.25 (0.96 – 1.63)	0.095	1.28 (1.00- 1.62)	0.047	1.23 (0.95 – 1.60)	0.12
CD8+ cell density (continuous, log2 transformed)	1804	435	350	0.90 (0.85 – 0.95)	1.7 x 10 ⁻⁴	0.91 (0.86 – 0.98)	3.2 x 10 ⁻³	0.92 (0.87 – 0.97)	3.6 x 10 ⁻³	0.93 (0.87 – 0.99)	0.024

1923

1924

1925 *Table 3-9. Univariable and multivariable analyses of time to colorectal cancer recurrence and overall survival in pooled VICTOR and QUASAR2 trial population. pT :*

primary tumour stage. POLE-mutant: pathogenic POLE mutation. MMR: DNA mismatch repair MMRp: mismatch repair proficient. MMRd: mismatch repair deficient. KRAS-

mutant: KRAS mutation at codons 12, 13 or 61. BRAF-mutant: BRAF hotspot V600E mutation. Reproduced from Glaire et al [1]

1926

1927

1928

1929

1930

3.3.5. Prognostic association of CD8⁺ cell infiltrate according to tumour and nodal risk strata

Whilst previous research has shown that there is a prognostic association between increased T cell infiltrate and clinical outcome, it has not been fully investigated whether this relationship holds true across all CRCs or whether it varies according to molecular and clinicopathological features [12, 148]. One particular advantage of using a large experimental trial cohort with detailed annotations is the capacity to investigate this. As such, I next investigated whether the association between the CD8⁺ cell infiltrate and clinical outcome was consistent across this group of 1,804 stage II/III CRCs. In univariable analysis, CD8⁺ cell density was significantly associated with CRC recurrence irrespective of patient age (<70 vs. ≥70 years), sex, tumour location (right vs. left sided), *KRAS* and *BRAF* mutation status. Whilst the association was not statistically significant in MMR-D/*POLE*-mutant or CIN-low subgroups, the hazard ratios were similar to those in MMR-P/*POLE*-wild-type and CIN-high subgroups respectively. This was likely due to the low numbers of tumours in those subgroups.

Univariable and multivariable analyses did show that there was a notable difference in this association between pT1-3 and pT4 tumours (multivariable-adjusted HR=0.96, 95% CI=0.89–1.04, *P*=0.35; and HR=0.87, 95% CI=0.80–0.95, *P*=1.9 x 10⁻³ respectively), and, between node negative and node positive tumours (HR=0.94, 95% CI=0.84–1.06, *P*=0.31; and HR=0.91, 95% CI=0.85–0.97, *P*=5.0 x 10⁻³ respectively). To further analyse this variation, tumours were grouped according their risk of recurrence. From the multivariable model (Table 3.9), pT4 and node positivity carried a similarly reduced TTR. Furthermore, analysis of recurrence probability showed that pT1-3, pN1-2 and pT4, pN0 tumours were at a similar risk of disease recurrence. Visualisation of the recurrence risk by Kaplan Meier curves suggested three distinct subgroups (Figure 3.5). As such, tumours were grouped into three risk strata: low risk (pT1-3, N0), intermediate risk (pT4, pN0 or pT1-3, pN1-2) and high risk (pT4, pN1-2).

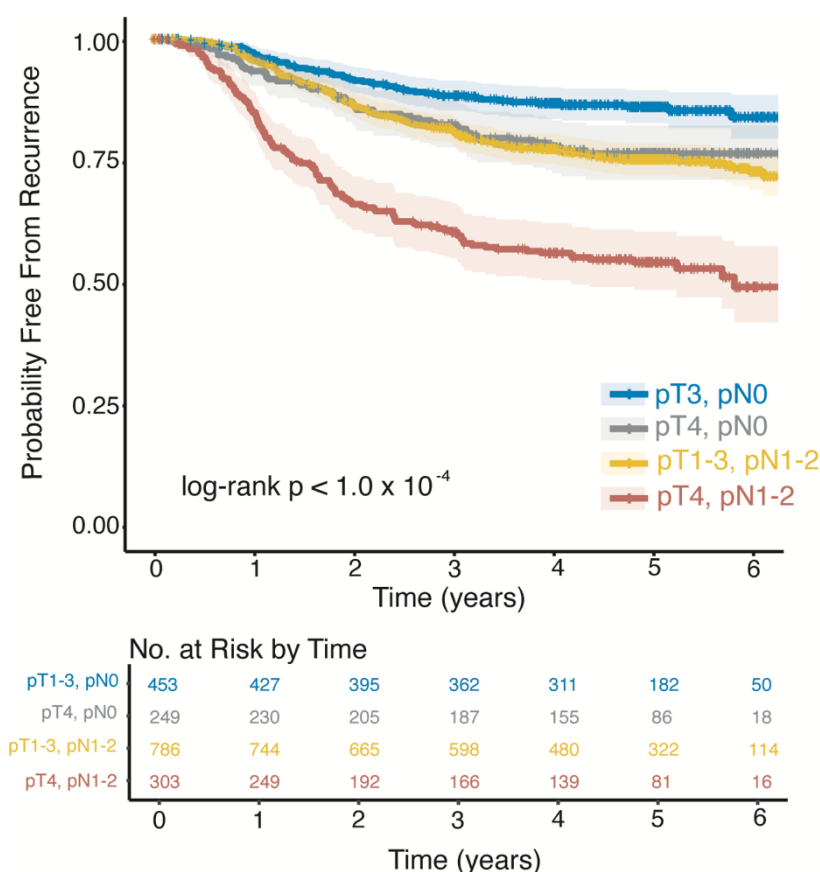


Figure 3-5. **Kaplan Meier curve showing the probability of recurrence according to primary tumour and nodal stage in QUASAR2 and VICTOR cohorts.** Reproduced from Glaire et al [1]

1957

1958 The association between CD8⁺ cell infiltrate and clinical outcome was examined across these groups.

1959 Interestingly, it revealed that the association was practically absent in low-risk tumours (HR=1.03,

1960 95% CI=0.87–1.21, $P=0.75$), modest in intermediate-risk cases (HR=0.92, 95% CI=0.86–1.00,

1961 $P=0.046$), and strongest in high-risk cases (HR=0.87, 95% CI=0.79–0.97, $P=9.4 \times 10^{-3}$). However,

1962 formal interaction testing between risk strata and recurrence probability was not statistically

1963 significant result using a two-sided alpha of 0.05 ($P_{\text{INTERACTION}}=0.090$) (Figure 3.6, Table 3.10). A

1964 similar variation was found for the relationship between CD8⁺ cell infiltrate and OS: low-risk tumours

1965 (HR=1.04, 95% CI=0.87–1.24, $P=0.69$), intermediate-risk cases (HR=0.94, 95% CI=0.86–1.03,

1966 $P=0.22$), and high-risk cases (HR=0.88, 95% CI=0.79–0.98, $P=0.022$), with a borderline significant P

1967 value for interaction ($P_{\text{INTERACTION}}=0.056$). The consequence of this variation is further highlighted by

1968 examining the absolute CRC recurrence risk at 3 years in each risk strata. In the low-risk group,

1969 tumours with a weak and intense T-cell response (defined as a CD8⁺ cell density at the 25th or 75th

centiles) had 3 year recurrence free probabilities of 90.2% (95% CI 87.8–92.7%) and 90.1% (95% CI 87.7–92.71%) respectively. In comparison, in the high-risk strata tumours corresponding tumours had 3 year recurrence free probabilities of 58.3% (95% CI=52.8–64.4%) and 68.6% (95% CI=63.1–74.6%) (Table 3.10; Figure 3.7).

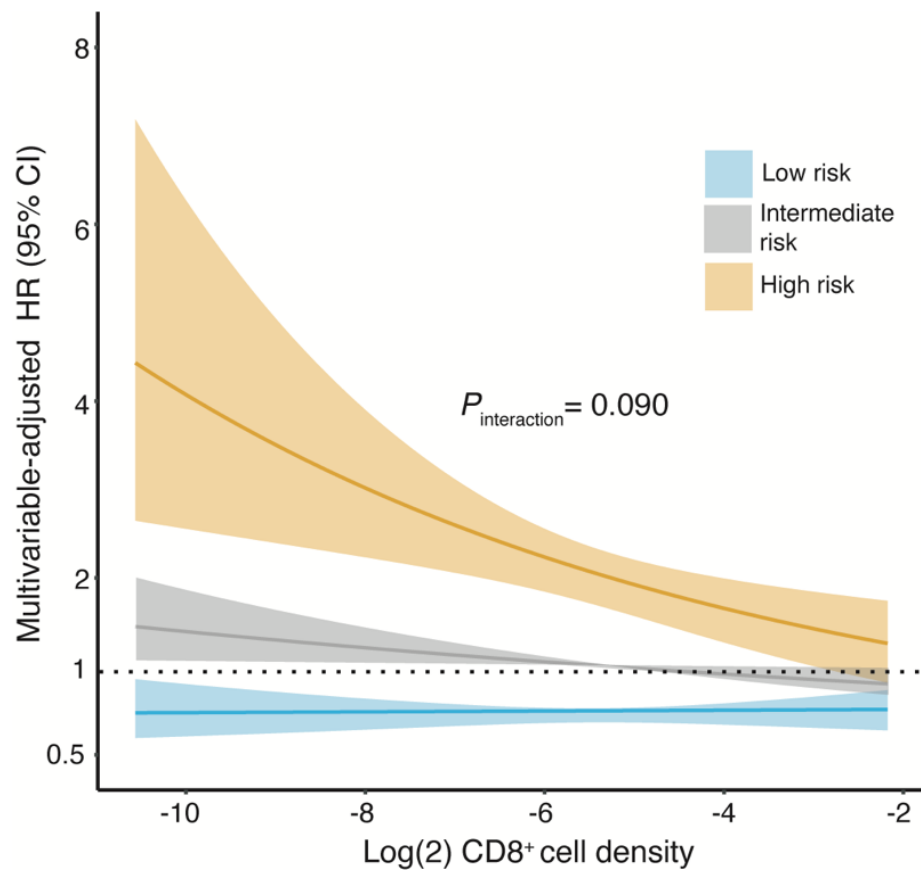


Figure 3-6. **Multivariable adjusted hazard ratio showing association of CD8⁺ cell infiltrate and recurrence risk divided into low risk (pT1-3), intermediate risk (pT1-3 & pN1-2, pT4 & pN0) and high risk (pT4 & pN1-2).** Reproduced from Glaire et al 2019 [1]

1976 **Table 3-10. Estimated probabilities of colorectal cancer recurrence and overall survival according to**
 1977 **tumour risk strata and CD8⁺ cell density dichotomized at sample median in the pooled QUASAR2 and**
 1978 **VICTOR trial population Glaire et al 2019 [1]**

	No.	No. events	Predicted proportion event free at 3 years		Univariable analysis			Multivariable analysis		
			25 th centile CD8 ⁺ cell density (95% CI)	75 th centile CD8 ⁺ cell density (95% CI)	HR (95% CI)	<i>P</i>	<i>P</i> _{INTERACTION}	HR (95% CI)	<i>P</i>	<i>P</i> _{INTERACTION}
Time to recurrence										
Low risk (pT3, N0)	453	60	0.90 (0.88 – 0.93)	0.90 (0.88 – 0.93)	1.02 (0.86–1.20)	0.85	0.072	1.03 (0.87–1.21)	0.75	0.090
Intermediate risk (pT4, N0 or pT1-3, N1/2)	1035	242	0.79 (0.76 – 0.82)	0.82 (0.80 – 0.85)	0.91 (0.85–0.98)	0.016		0.92 (0.86–1.0)	0.046	
High risk (pT4, N1/2)	303	132	0.58 (0.53 – 0.64)	0.69 (0.63 – 0.75)	0.86 (0.78–0.96)	4.7 x10 ⁻³		0.87 (0.79–0.97)	9.4 x10 ⁻³	
Overall survival										
Low risk (pT3, N0)	453	52	0.95 (0.94 – 0.97)	0.95 (0.93 – 0.96)	1.03 (0.87–1.23)	0.72	0.051	1.04 (0.87–1.24)	0.69	0.056
Intermediate risk (pT4, N0 or pT1-3, N1/2)	1035	177	0.89 (0.87 – 0.91)	0.90 (0.88 – 0.92)	0.94 (0.86–1.03)	0.20		0.94 (0.86–1.03)	0.22	
High risk (pT4, N1/2)	303	120	0.75 (0.71 – 0.80)	0.82 (0.78 – 0.86)	0.88 (0.79–0.97)	0.017		0.88 (0.79–0.98)	0.022	

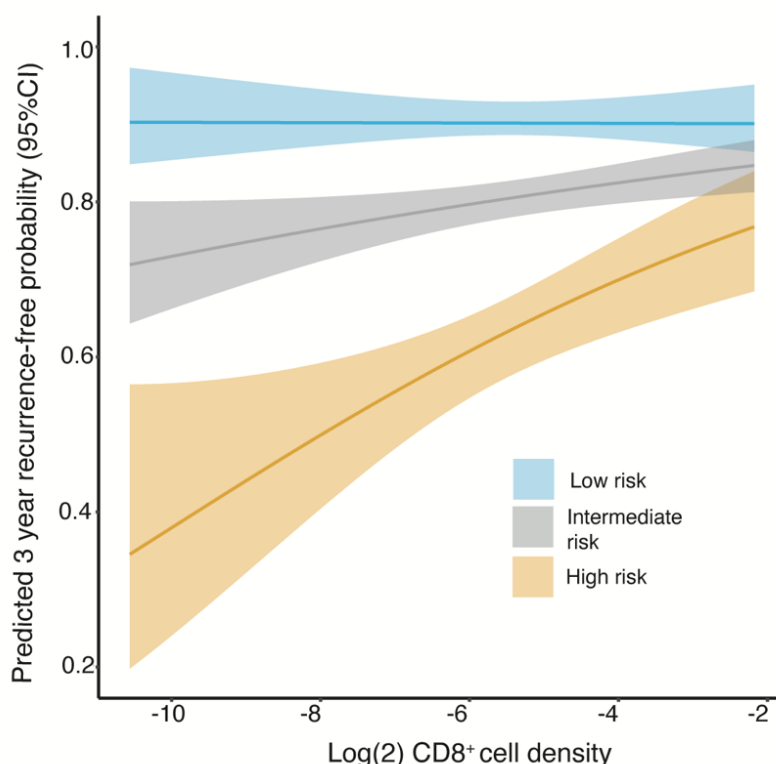


Figure 3-7. 3 year recurrence free probability according to risk strata: low risk (pT1-3), intermediate risk (pT1-3 & pN1-2, pT4 & pN0) and high risk (pT4 & pN1-2). Reproduced from Glaire et al 2019 [1]

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

To again facilitate comparisons with previous publications, patients were dichotomised according a median cut point and the recurrence probability analysed. In the low risk group there was no significant difference in the risk of recurrence between CRCs with high CD8⁺ infiltrate vs low (HR = 1.03, 95%CI = 0.54-1.72, $P=0.91$). In contrast, and consistent with the previous results, similar analysis of the intermediate and high risk groups revealed statistically significant differences (HR = 0.69, 95%CI = 0.51-0.93, $P=0.014$ and HR = 0.59, 95%CI = 0.39-0.89, $P=0.011$ respectively) (Figure 3.9; Table 3.11).

Table 3-11. Estimated probabilities of colorectal cancer recurrence and overall survival according to tumour risk strata and CD8⁺ cell density dichotomized at sample median in the pooled QUASAR2 and VICTOR trial population

Time point (years)	All cases		Low risk (pT3, N0)		Intermediate risk (pT4, N0 or pT1-3, N1/2)		High risk (pT4, N1/2)	
	CD8 ⁺ lo % (95% CI)	CD8 ⁺ hi % (95% CI)	CD8 ⁺ lo % (95% CI)	CD8 ⁺ hi % (95% CI)	CD8 ⁺ lo % (95% CI)	CD8 ⁺ hi % (95% CI)	CD8 ⁺ lo % (95% CI)	CD8 ⁺ hi % (95% CI)
CRC recurrence-free probability								
3	76.5 (73.7–79.4)	84 (81.6–86.6)	86.9 (82.2–91.8)	89.6 (85.8–93.6)	79.0 (75.5–82.7)	85.7 (82.6–89.0)	55.2 (47.9–63.7)	67.4 (59.7–76.1)
5	70.1 (67.0–73.4)	79.8 (77.0–82.7)	85.6 (80.0–90.8)	86.9 (82.4–91.5)	71.6 (67.6–75.8)	80.7 (77.1–84.5)	45.7 (38.2–54.6)	63.8 (55.8–72.9)
Overall survival probability								
3	86.4 (84.0–90.0)	89.9 (88.0–91.9)	95.0 (93.4–96.8)	95.0 (93.2–96.8)	88.4 (86.2–90.7)	90.6 (88.7–92.5)	74.0 (68.5–79.8)	82.6 (78.0–87.5)
5	75.8 (72.3–79.5)	81.8 (78.8–84.8)	90.8 (87.8–93.8)	90.6 (87.5–93.7)	78.9 (75.5–82.4)	82.6 (79.7–85.6)	55.9 (48.7–64.1)	69.2 (62.3–76.8)

The association between tumour-infiltrating CD8⁺ lymphocytes and clinical outcome in stage II/III colorectal cancer

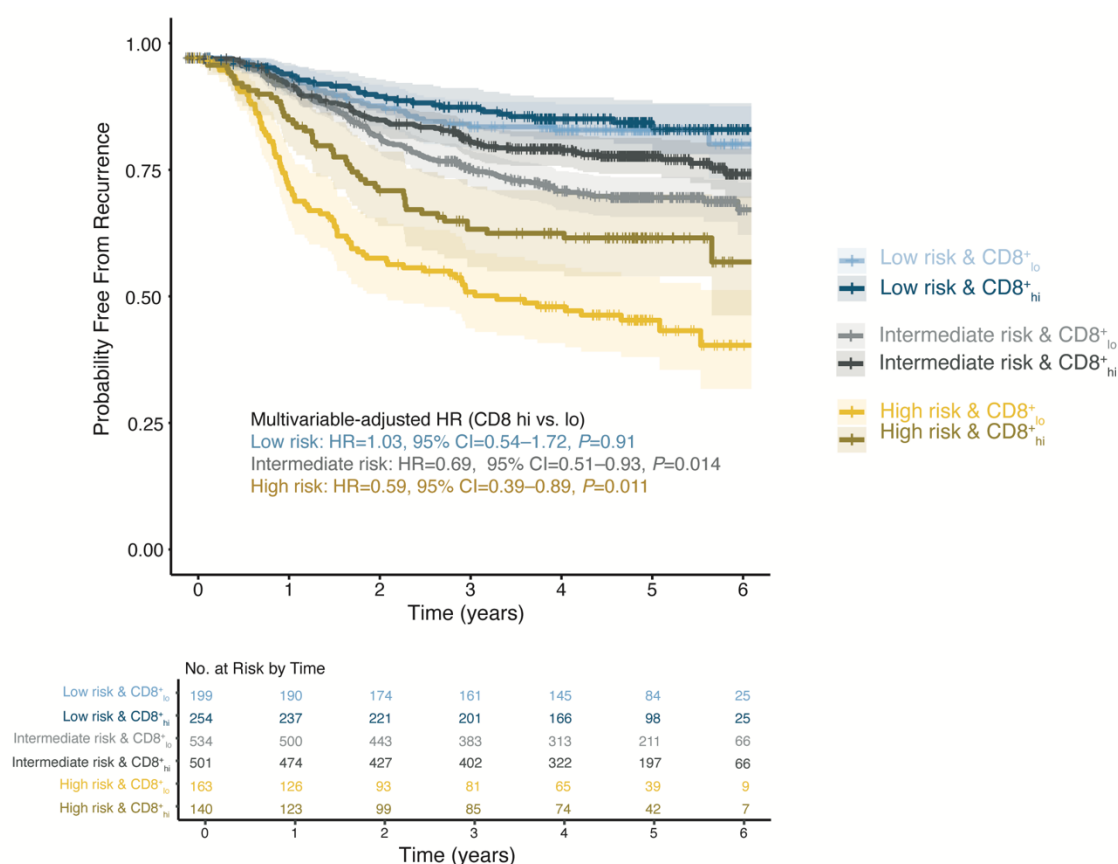


Figure 3-8. **Kaplan Meier curve showing probability of recurrence according CD8⁺ infiltrate high or low and by tumour risk strata.** Reproduced from Glaire et al 2019 [1]

3.3.6. Prognostic implication of *BRAF* mutation and genomic status on clinical outcome in the combined QUASAR2 and VICTOR population

To understand the impact of *BRAF* mutation on clinical outcome in stage II/III CRC patients in the VICTOR/QUASAR2 cohort, patients were stratified into four groups according to hypermutation status (MMR-D or POLE mutant) and *BRAF* mutational status. Examination of the Kaplan Meier curves showed that *BRAF* mutant, non-hypermuted CRCs had considerable worse time to recurrence and overall survival compared to *BRAF* wild-type, non-hypermuted CRCs (Figures 3.9 and 3.10). In comparison, there was no evidence of reduced time to recurrence or overall survival for *BRAF* wild-type, hypermutated CRCs and *BRAF* mutant, hypermutated CRCs. Furthermore, in multiple linear regression model *BRAF* mutational status did not affect the positive effect of hypermutation on CD8⁺ cell infiltration (Table 3.12). Taken together these results confirm the established finding that *BRAF* mutation may be a negative prognostic factor in non-hypermuted

2009 (predominantly MMR-P) CRCs. In contrast, *BRAF* mutation in hypermutated (predominantly MMR-
2010 D) CRCs does not influence CD8⁺ cell infiltration, nor clinical outcome.

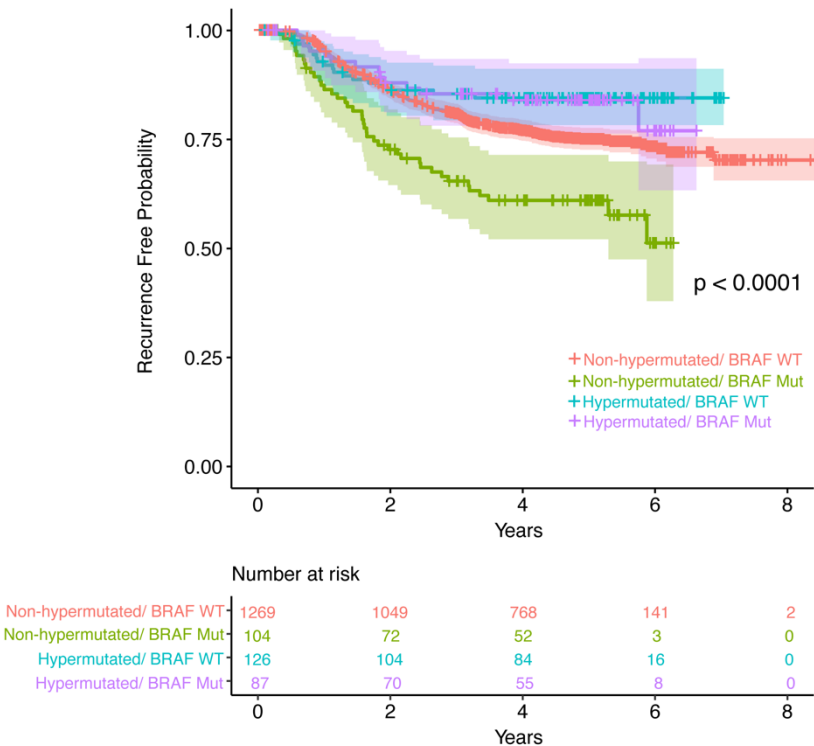


Figure 3-9 Kaplan Meier curve showing the probability of recurrence according to combined hypermutation (MMR-D or POLE mutation) and BRAF mutation status in QUASAR2 and VICTOR cohorts. P value calculated by log-rank test

2011

The association between tumour-infiltrating CD8+ lymphocytes and clinical outcome in stage II/III colorectal cancer

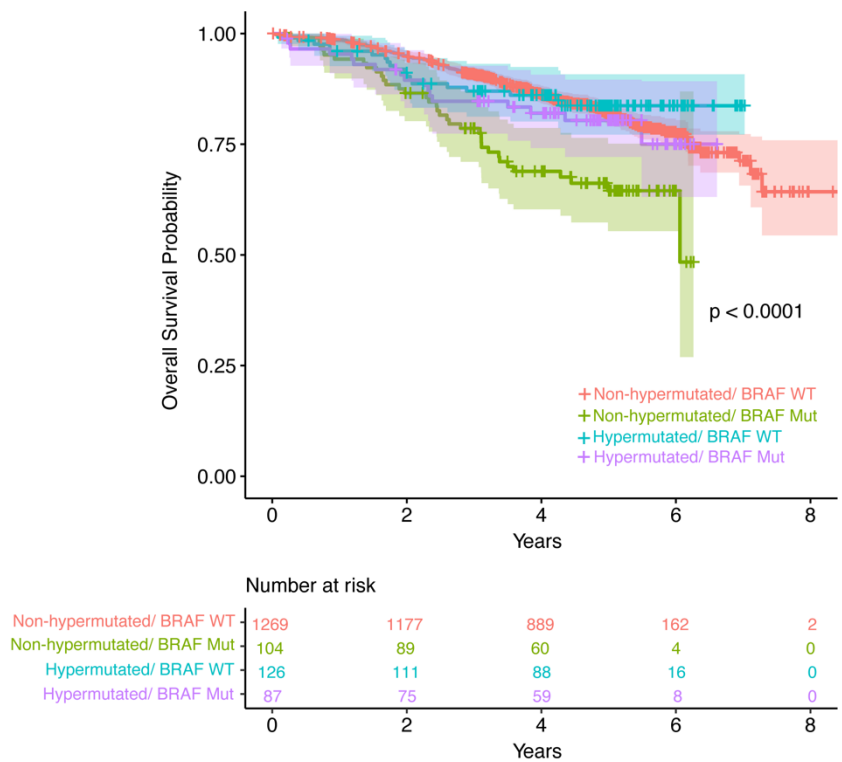


Figure 3-10. **Kaplan Meier curve showing overall survival probability according to combined hypermutation (MMR-D or POLE mutation) and BRAF mutation status in QUASAR2 and VICTOR cohorts. P value calculated by log-rank test**

2012

2013

Table 3-12. Multiple linear regression analysis of the relationship between clinicopathological/molecular characteristics and CD8⁺ cell density in the combined QUASAR2 and VICTOR trial samples. *POLE-mutant: pathogenic POLE mutation. MMRd: mismatch repair deficiency. MMRp: mismatch repair proficient. CIN: chromosomal instability. CRC: colorectal cancer. KRAS-mutant: KRAS mutation at codons 12, 13 or 6. BRAF-mutant: BRAF hotspot V600E mutation. Effect, or beta-coefficient, represents the effect independent variables have on the dependent variable (in this case, CD8⁺ cell density).*

	Linear Regression Results	
	Effect (95% CI)	P
Age (continuous)	-0.11 (-0.27 – 0.04)	0.15
Sex		
Male	0.0 (ref)	–
Female	0.04 (-0.13 – 0.21)	0.63
Location		
Left	0.0 (ref)	–
Right	-0.25 (-0.43 – -0.07)	7.2 x 10 ⁻³
Stage		
II	0.0 (ref)	–
III	-0.17 (-0.34 – 0.00)	0.05
Primary tumour		
pT1-3	0.0 (ref)	–
pT4	0.00 (-0.18 – 0.18)	0.98
MMR & POLE status		
MMR-P & POLE wild-type and BRAF wild-type	0.0 (ref)	–
MMR-P & POLE wild-type and BRAF mutant	-0.064 (-0.41 – 0.29)	0.71
MMR-D or POLE-mutant and BRAF wild-type	0.46 (0.12 – 0.80)	8.3 x 10 ⁻³
MMR-D or POLE-mutant and BRAF mutant	0.59 (0.20 – 0.98)	3.3 x 10 ⁻³
Chromosomal instability		
CIN low	0.0 (ref)	–
CIN high	-0.05 (-0.24 – 0.14)	0.59

3.3.7. External validation of results

Given the novelty of this finding, it was essential to validate these results in an external, independent cohort. Unfortunately, there were no publicly available datasets that have quantified CD8⁺ cell infiltrate in CRC by IHC. Instead, external validation was carried out on alternative, publicly available sources. As an alternative to CD8⁺ cell quantification by IHC, data series were sought out that had performed quantification of *CD8A* expression, which is expressed by CD8 T cells [250] and increasing expression levels of which have been previously associated with outcome in CRC [218]. Datasets were identified either through search of the NCBI GEO database, or were those that are well known (such as TCGA). To be included, the datasets had to have gene expression data for *CD8A*, detailed tumour and nodal staging, and clinical outcome information. In total, the validation set included 1375 patients with stage II/III CRC tumours from seven independent series (Table 3.13). Expression of *CD8A* was performed by either RNAseq or expression microarrays. Gene expression data from log2 transformed, if not already done, and Z-score normalised within each series prior to pooling.

Table 3-13. Details of the seven publicly available independent series included in the external validation cohort

Cohort	N. of patients included	Gene expression quantification methodology	MMR status available	Cancer stage	Cohort Type	Reference
GSE103479	152	Expression profiling by array	Yes	II & III	Non-trial	[251]
GSE12945	42	Expression profiling by array (Affymetrix)	Yes	I - IV	Non-trial	[252]
GSE33113	90	Expression profiling by array (Affymetrix)	No	II	Non-trial	[253]
GSE39582	440	Expression profiling by array (Affymetrix)	Yes	I - IV	Non-trial	[254]
GSE41258	96	Expression profiling by array (Affymetrix)	No	I-IV	Non-trial	[255]
Oslo	257	Expression profiling by array (Affymetrix)	Yes	II-III	Non-trial	[256]
TCGA	298	RNA sequencing (Illumina)	Yes	I - VI	Non-trial	[87]

Consistent with previous studies [218] and in keeping with analysis of QUASAR2 and VICTOR, increasing tumour *CD8A* expression was significantly associated with an improved TTR in multivariable analysis (HR for each 2-fold increase= 0.86, 95%CI= 0.76–0.97, $P=0.018$). Analysis of the association between tumour and nodal status and CRC recurrence revealed similar separation of cases into three risk strata as seen in the combined QUASAR2/VICTOR series (Figure 3.11)

I next examined whether the strength of the association between CD8A expression and tumour recurrence displayed similar variation across risk strata as in the discovery cohorts. Concordant with results from the VICTOR AND QUASAR2 trial cases, the strength of this association varied between low, intermediate and high-risk pT/N strata (HR=0.99, 95% CI=0.81–1.24, $P=1.0$ vs. HR=0.83, 95% CI=0.69–0.99, $P=0.034$ vs. HR=0.77, 95% CI=0.51–1.16, $P=0.21$ respectively; $P_{\text{INTERACTION}}=0.048$) (Figure 3.12). Again, this was reflected in substantial differences in 3 year recurrence risk. In the low risk groups, tumours at the 25th and 75th centiles had a probability of being recurrence free at 3 years of 85.5% (95% CI 82.3-89.7) and 86.0% (95% CI 82.4-89.7) respectively. In the intermediate group, the corresponding probabilities were 68.7% (95% CI 63.4-74.5) and 74.7%(95% 10.0-79.6) and in the high risk group 40.7% (95% CI 30.4-54.4) and 56.8% (95% 47.3-68.3).

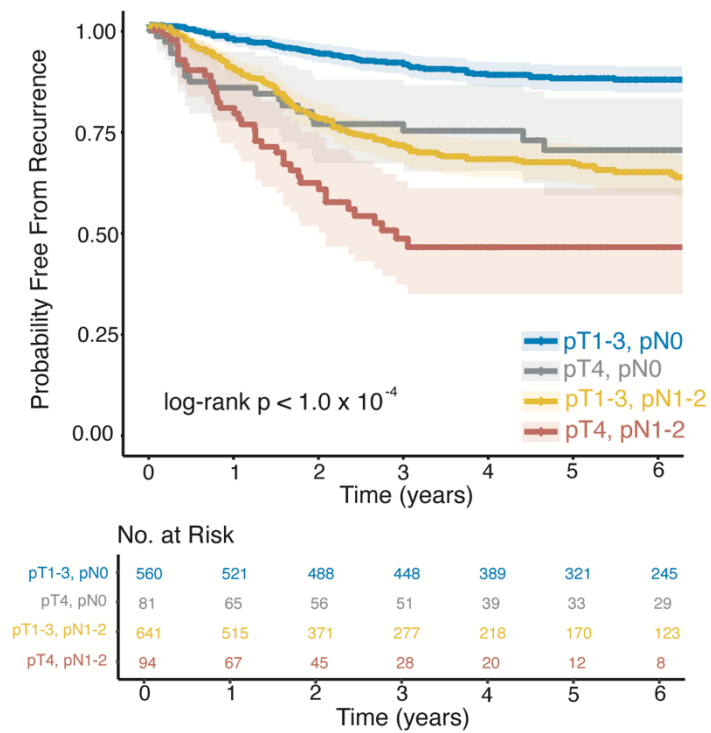


Figure 3-11.. **Kaplan Meier curve showing the probability of recurrence according to primary tumour and nodal stage in the external validation cohort. Reproduced from Glaire et al 2019 [1]**

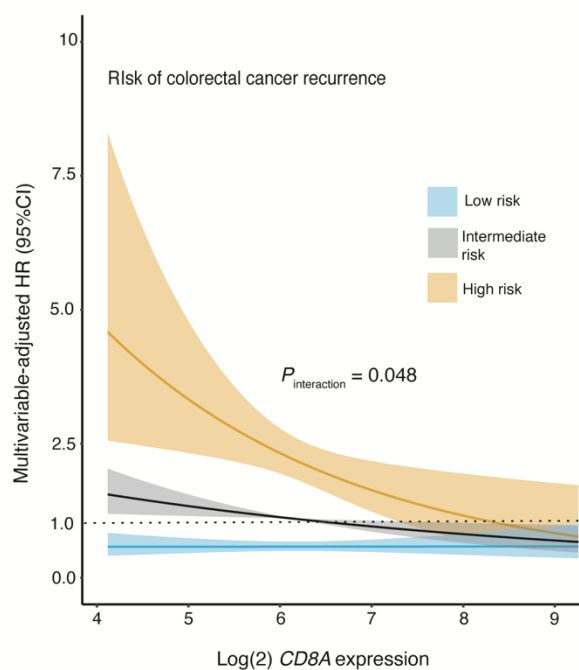


Figure 3-12 Multivariable adjusted hazard ratio showing association of CD8A expression and recurrence risk divided into low risk (pT1-3), intermediate risk (pT1-3 & pN1-2, pT4 & pN0) and high risk (pT4 & pN1-2) in the external validation cohort. Reproduced from Glaire et al 2019 [1]

3.3.8. Work planned but not completed due to COVID-19 pandemic

3.3.8.1. Multispectral immunoprofiling of QUASAR2 and SCOT trial cohorts

The data presented above show the important association between CD8⁺ cytotoxic T cells and clinical outcome in early-stage CRC. Furthermore, these data show that IHC quantification of CD3⁺ and CD8⁺ T cell populations may aid in the prognostication of CRC patients. However, IHC quantification of single markers is limited by the data that it can capture. In particular, it is not possible to quantify multiple immune populations in a single tumour section, and as such there is no opportunity to analyse the relationship between different cell types beyond those identified by a small number of markers. To this end, I wanted to conduct a more comprehensive analysis of the immune microenvironment in early-stage CRC. I intended to achieve this using a multi-spectral, immunofluorescence panel. This would be done using samples from the QUASAR2 trial (detailed above) and the SCOT trial. The SCOT trial was a phase III randomised, non-inferiority, trial comparing 6 months versus 3 months of combined oxaliplatin and fluoropyrimidine adjuvant chemotherapy for high-risk stage II/III CRC patients who had undergone curative resection [257]. The results of this trial showed the 3 months of adjuvant chemotherapy was not inferior to 6 months, but did result in reduced toxicity events and improved quality of life. The final pooled study population for the multispectral immune analysis includes ~4800 SCOT patients and ~1200 QUASAR2 patients, and similar to the work described above, utilises FFPE TMA slides.

At the time of writing, all slides for both cohorts have been stained and scanned, and are under analysis in collaboration with the group of Viktor Koelzer (University of Zurich). For most cases, triplicate tumour core samples (ie 3 tumour regions, including the tumour centre and invasive margin) have been stained and are under evaluation, however for a subset of SCOT trial patients, eight tumour cores (ie 8 regions) will be analysed. Details of the immune panel are listed in Table 3.14, and a representative image shown in Figure 3.13. To date, I have worked with the Translational Histopathology Laboratory at the University of Oxford to develop and optimise the staining panel and methodology. Having achieved this, I have also collaborated with the University of Glasgow to

organise samples for staining, and with the Kolzer laboratory, University of Zurich to begin developing an analytical pipeline. All tissue staining has now been completed, and image analysis is underway.

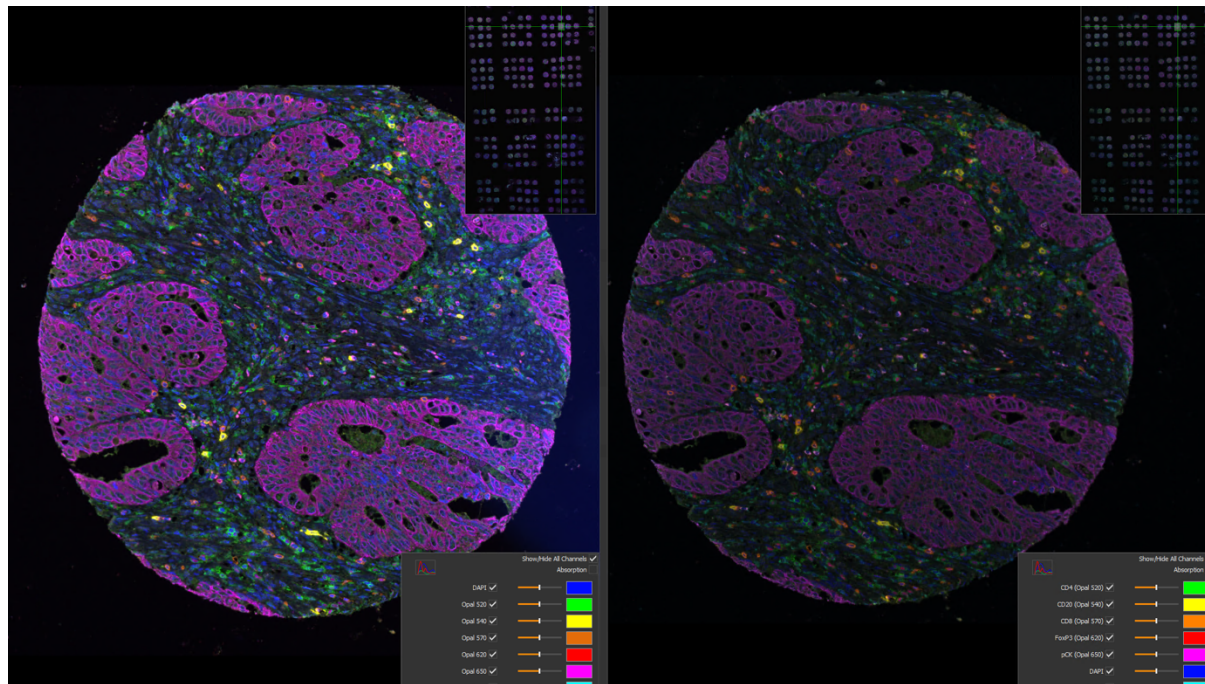


Figure 3-13. Representative images showing 7-plex co-immunofluorescence staining. These are example images of the analysis of QUASAR2 and SCOT trial samples that have been stained using the multispectral, immunofluorescence panel detailed in Table 3.12

3.3.8.2. Analysis of MMR status of SCOT trial cohort

To complement the analysis above, TMAs cores from the SCOT trial have also been stained for MMR protein (MLH1, PSM2, MSH2, MSH6) expression by immunohistochemistry and analysis will be conducted by the group of Viktor Koelzer.

Unfortunately, both projects have been delayed by the COVID-19 pandemic, meaning that completion has not been possible. I had aimed to completed part of this project, and particular areas that were of relevance to my work so far are:

- Confirmation of the variation in the prognostic association of CD8⁺ according to tumour and nodal status in the SCOT trial cohort

The association between tumour-infiltrating CD8+ lymphocytes and clinical outcome in stage II/III colorectal cancer

- 2107 • The prognostic association of the quantities of other immune cell types in stage II/III CRC:
- 2108 namely T-helper and B-cell
- 2109 • Preliminary analysis of the spatial interaction between immune cell subtypes, which may will
- 2110 give some insight into immune cell dynamics in the tumour microenvironment
- 2111 •
- 2112 Other analyses that were planned, but not to be performed by me included:
- 2113 • Estimation of the heterogeneity of the immune cell infiltrate in stage II/III CRCs
- 2114 • Estimation of the heterogeneity of MMR protein loss in stage II/III CRC and how this
- 2115 correlates with immune cell infiltrate
- 2116 •

2117 **Table 3-14. Details markers being using for multispectral, immunofluorescence immunoprofiling project**
 2118 **of ~5000 cancers from QUASAR2 and SCOT trial**

Marker	Description	Refs.
CD8	Marker for CTLs, which have been shown to strongly correlate with improved outcomes in CRC by the work above and numerous other studies	[1, 143, 147, 149, 153]
CD4	Defines T-helper cells, which are essential co-ordinators of the adaptive immune response. Associated with improved outcomes in diverse range of cancers	[150-154]
FoxP3	Defines T _{reg} cells, which regulate other adaptive immune cells and are associated with improved outcomes	[161, 163, 258-260]
CD20	B-cell marker; predominantly involved in humoral immunity but also contribute to antigen presentation. Role in CRC remains relatively under studied	[158, 261]
CD68	Pan-macrophage marker. Association with clinical outcome varies according the exact subset of macrophages examined	[117-119, 121-123]
Cytokeratin	Marker of epithelial cells within the tumour. Will allow for localisation of cells to epithelial or stroma	-

2119

3.4. Discussion

3.4.1. CD8⁺ cell infiltrate and clinical outcome in early-stage CRC

Current prognostication strategies are insufficient to capture the substantial variation in clinical outcome that is observed in early-stage CRC [49]. There is an urgent need for biomarkers to improve CRC patient stratification, and identify those most at risk of disease recurrence despite optimal treatment. Previous work has demonstrated the potential of immune based biomarkers, and in particular this has focused on quantification of tumour infiltrate T-cell populations. This clinical translation of this work has been limited by the use of non-experimental cohorts and a failure to adjust for key molecular factors of known prognostic significance [262]. In addition, the use of forced dichotomisation [148] for analysis presents a number of problems, including loss of important information and reduced sensitivity of a prognostic marker [235].

In this chapter, intratumoral CD8⁺ cell infiltrate was quantified by IHC in 1804 clinical trial patients with stage II/III CRC. Consistent with previous published research [147, 148, 153], this study showed that increasing CD8⁺ cell infiltrate was associated with an improved time to recurrence and overall survival in stage II/III CRC. In contrast to previous work, I performed these analyses with CD8⁺ cell infiltrate as a continuous variable. My results suggest this is a valid approach, which provides pertinent prognostic information: every doubling in CD8⁺ fraction was associated with a significantly reduced recurrence risk (HR=0.90, 95% CI=0.85–0.95). Furthermore, the relationship between CD8⁺ cell infiltrate and clinical outcome remained significant in multivariable models after adjustment for molecular variables. Importantly, these variables included MMR, *POLE* mutation and CIN status. These three forms of genomic instability have all been associated with clinical outcome in CRC, and differential tumour immune response. MMR status is routinely used in clinical work as a biomarker of improved prognosis [49, 95] and a predictive marker for novel immunotherapies [52]; whilst, experimental data has shown that *POLE* mutation [90] and CIN status [227] have been associated with improved and worse outcomes respectively. It was particularly interesting that the prognostic association between MMR status and clinical outcome was not significant in the full multivariable

models. Whilst most previous studies have not adjusted for MMRd, the latest Immunoscore publication (which had not been published at the time of completing this work) did include this as part of the multivariable model [148]. Results in this study showed that MMRd status remained significantly associated with TTR in multivariable analysis (HR = 0.63, 95%CI (0.44-0.90)). The discordance in findings could possibly be explained by the use of a clinical trial cohort in my work, and that I was able to adjust for other key molecular variables such as *BRAF* and *KRAS* mutations.

Interestingly, and not previously reported, the association between CD8⁺ cell infiltrate and clinical outcome was not uniform across all CRCs. The relationship appears to be weakest in low-risk CRCs (defined as pT1-3 and pN0), stronger in intermediate (pT1-3 and pN1-2; pT4 and pN0) and strongest in high-risk CRCs (pT4 and pN1-2). These novel findings were validated using an external, independent cohort. These results have important implications for the utilisation of CD8⁺ cell infiltrate as a prognostic biomarker. It would suggest that quantification of CD8⁺ cell infiltrate is of greatest utility in cancers defined at high risk by TNM staging, the current mainstay of prognostication. Patients who have high-risk cancers, and low CD8⁺ cell infiltrate, may prove to be ideal candidates for novel therapies that boost anti-tumour immunity, such as cancer vaccines or CAR-T therapy.

3.4.2.Strengths

There are several strengths of this work compared to previous publications. Firstly, this work used an experimental cohort formed with patients from two large clinical trials. In addition to providing detailed clinicopathological annotations and carefully curated outcome data, it also minimises the risk of bias [226]. Additionally, my work involved a large number of patients (1,804), approaching the size of the largest Immunoscore study to date (2,681 patients) and larger than other studies which have examined Immunoscore in clinical trials (600 and 1,322 patients from the NCTG N0147 and IDEA trials respectively) [144, 225] Secondly, this work has shown the association between CD8⁺ and clinical outcome remains significant after adjustment for MMR, *POLE* mutation and CIN status. The fact that the association of CD8⁺ infiltrate with clinical outcome is independent of these

molecular variables further strengthens the case for its use as a prognostic biomarker. At time of embarking on this study, there were few previous works that had adjusted for genetic and molecular confounding variables when assessing the association between T-cell infiltrate and clinical outcome. The latest iteration of Immunoscore was published while the manuscript reporting this work was under review, and involved validation on 2,681 patients (700 training set, 636 internal validation set and 978 external validation set) [148]. Results showed that the Immunoscore was strongly associated with time to recurrence, overall survival and disease-free survival; this relationship remained significant after adjustment for MMR status [148]. However, in contrast to my work, this study did not include multivariable adjustment for other key genetic and molecular variables in addition to MMR deficiency, such as *POLE* mutations, CIN, *KRAS* and *BRAF* mutations. Thirdly, analysing the relationship between clinical outcome and CD8⁺ cell infiltrate as a continuous variable avoided pitfalls of dichotomisation, which has been an approach commonly used in previous research efforts [148]. Fourthly, the novel finding of the variability of CD8⁺ cell infiltrate across CRCs was confirmed using an entirely independent, external validation set.

3.4.3.Limitations

There are limitations to this work. For logistic reasons, only tumour cores were stained and quantified. Future work should on assessing immune infiltrate across whole tumour sections, with particular attention paid to the role of immune cell infiltrate at the invasive margin compared with the tumour centre would be of benefit. It would seem that HRs from the dichotomised CD8⁺ Hi vs Lo comparisons are inferior to the whole tumour section tumour analysis in the largest Immunoscore study (HR 0.56, 95%CI 0.49-0.71 compared with HR =0.71, 95% CI=0.59–0.87, $P=7.2 \times 10^{-4}$ reported in my work) [148]. However, my results are comparable to studies assessing Immunoscore in clinical trial patients (HR = 0.64; 95%CI=0.45-0.92, $P=0.015$ [144] and HR = 0.65; 95%CI=0.52-0.81, $P=0.0001$ [225]). The reason for this discrepancy could be the contribution of hidden cofounders, which are more prevalent in non-trial cohorts. In addition, it was not possible to assess the intratumoural heterogeneity of the T-cell infiltrate. Additionally, in this work, total CD8⁺ (and CD3⁺) cell infiltrate was determined by IHC and automated quantification. This quantification methodology

did not allow for localisation of the immune cells to the stromal or epithelial compartment. Previous work has suggested that the exact localisation of T-cells is important in determining their use as predictive and prognostic markers in CRC. It was shown in stage II CRC that high levels of intraepithelial CD8⁺ cells confer a favourable prognosis, whilst high levels of intrastromal CD3⁺ predicted benefit from immunotherapy treatment [263]. The prognostic utility of intraepithelial immune infiltrate compared with intrastromal has been observed in endometrial cancer too. Analysis of PORTEC -1 and -2 trials showed that intraepithelial CD8⁺ cells were strongly associated with outcome, whereas it appears that intrastromal immune infiltrate had no prognostic value [264]. The capacity to assess intraepithelial vs intrastromal TILs will be essential in refining CD8⁺ as a prognostic biomarker, and is currently being analysed.

3.4.4.Future work

Work currently ongoing is seeking to undertake a more in depth immunoprofiling of two large clinical trials (QUASAR2 and SCOT) using multiplex immunofluorescence and machine learning based quantification. This work will provide an important opportunity to validate the findings in this study, and in particular confirm the variance in prognostic association seen across CRCs. Furthermore, it will be important to understand the relationship between other immune cell populations of interest and clinical outcome. Additionally, this study will not only be able to assess absolute immune cell infiltrate quantities, but also analyse the spatial relationship within the tumour and with other immune cell populations.

In summary, in this study of a large, experimental cohort of stage II/III CRCs it was shown that increasing CD8⁺ cell infiltrate, analysed as a continuous variable, is independently associated with improved clinical outcome. Importantly, this relationship is not a constant across all CRCs and appears to vary according to tumour and nodal status. These findings will have significant implications in the implementation of CD8⁺ cell infiltrate as a prognostic biomarker, and may provide important insights into the role of the adaptive immune response in CRC.

4. The consequences of DNA mismatch repair deficiency in colorectal cancer and endometrial cancer

4.1. Introduction

As discussed in the Introduction, cancers typically arise due to genetic and epigenetic aberrations that collectively lead to unchecked cell proliferation, and loss of normal homeostatic mechanisms [9]. In both CRC and EC, a substantial fraction of tumours acquire such mutations as a result of genomic instability caused by failure of one, or more, safeguards that prevent DNA damage and/or correct it should it arise [76]. The most frequent such safeguard, and an important aetiological factor in CRC and EC development, is loss of DNA mismatch repair.

4.1.1. The DNA Mismatch Repair System

The DNA mismatch repair system is critical for maintaining genomic integrity during genomic duplication [93]. A highly evolutionary conserved system, it is responsible for correcting insertion deletion loops (IDLs) and base-base pair mismatches that are introduced during genomic replication. Loss of MMR function, MMRd, results in a mutator phenotype, and a failure to repair errors caused by slippage of DNA polymerases at repetitive DNA microsatellites – a phenomenon referred to as MSI. Microsatellites are repeated sequences of 1-6 base pairs repeated throughout the human genome; they account for around 3% of the human genome [265]. These sites are particularly vulnerable to mutations due to DNA polymerase slippage during replication. The terms MSI and MMRd do not describe the same thing, but there is a considerable amount of overlap between them and they are often used interchangeably.

MMRd may result from germline mutations in MMR genes in the familial cancer condition Lynch syndrome, or as result of somatic alterations, most commonly epigenetic silencing of the *MLH1* promoter, but also due to biallelic somatic mutations of MMR genes. MMRd is found in 15-20% of CRCs and 20-30% of EC [94]. Its clinical relevance will be discussed below.

Much of our understanding of the MMR system comes from the study of *E. coli*, in which the process is driven by two homodimers. MutS and MutL are responsible for correcting errors produced by the proofreading proficient DNA polymerase III [93, 266]. In eukaryotic cells, the situation is more complex (Figure 4.1). Replication errors are introduced not just by one DNA polymerase, but rather three different Family B DNA polymerases [267]. In humans, several MMR proteins have been discovered based on their similarity to the *E. coli* counterparts. Whilst MutS and MutL are homodimers in *E. coli*, the human equivalents are heterodimers [93]. hMutS α is formed from hMSH2 and hMSH3, and is the principal component responsible for recognising and initiating repair of DNA mismatches [268-270]. It predominantly recognises single base-base and indel mismatches [271]. hMutS α contains two active ATPase sites, and binding with ATP and the DNA mismatch causes it to conformationally change [272-275]. This allows it to form a clamp like structure capable of moving along the DNA strand, and interacting with the protein PCNA (proliferating cell nuclear antigen) [276]. PCNA acts as a matchmaker and in turn recruits hMutL α , a heterodimer of MLH1 and PMS2 [277]. In an ATP dependent manner, hMutL α then nicks the nascent DNA strand, which allows for excision driven by EXO1 [278]. The DNA strand is repaired with DNA synthesised by polymerases ϵ or δ and the break closed by DNA ligase [279]. Variations on the canonical MMR pathway also contribute to correcting DNA replication errors. hMutS β is composed of MSH2 and MSH3, and participates in repairing larger indel mismatches [280]. Additionally, there exists hMutL α homologues that are comprised of either MLH1 and MLH2, or MLH1 and MLH3. It is worth noting that the mutator phenotype induced by loss of MLH1, MSH2, MSH6 or PMS2 is far more severe than that associated with MSH3, MLH2 and MLH3 defects [93]. As such, the exact contribution of the latter three proteins is not as well understood. MMR loss results in a 100-1000 fold increase in mutation burden [281, 282], and in most cases, MSI as discussed above.

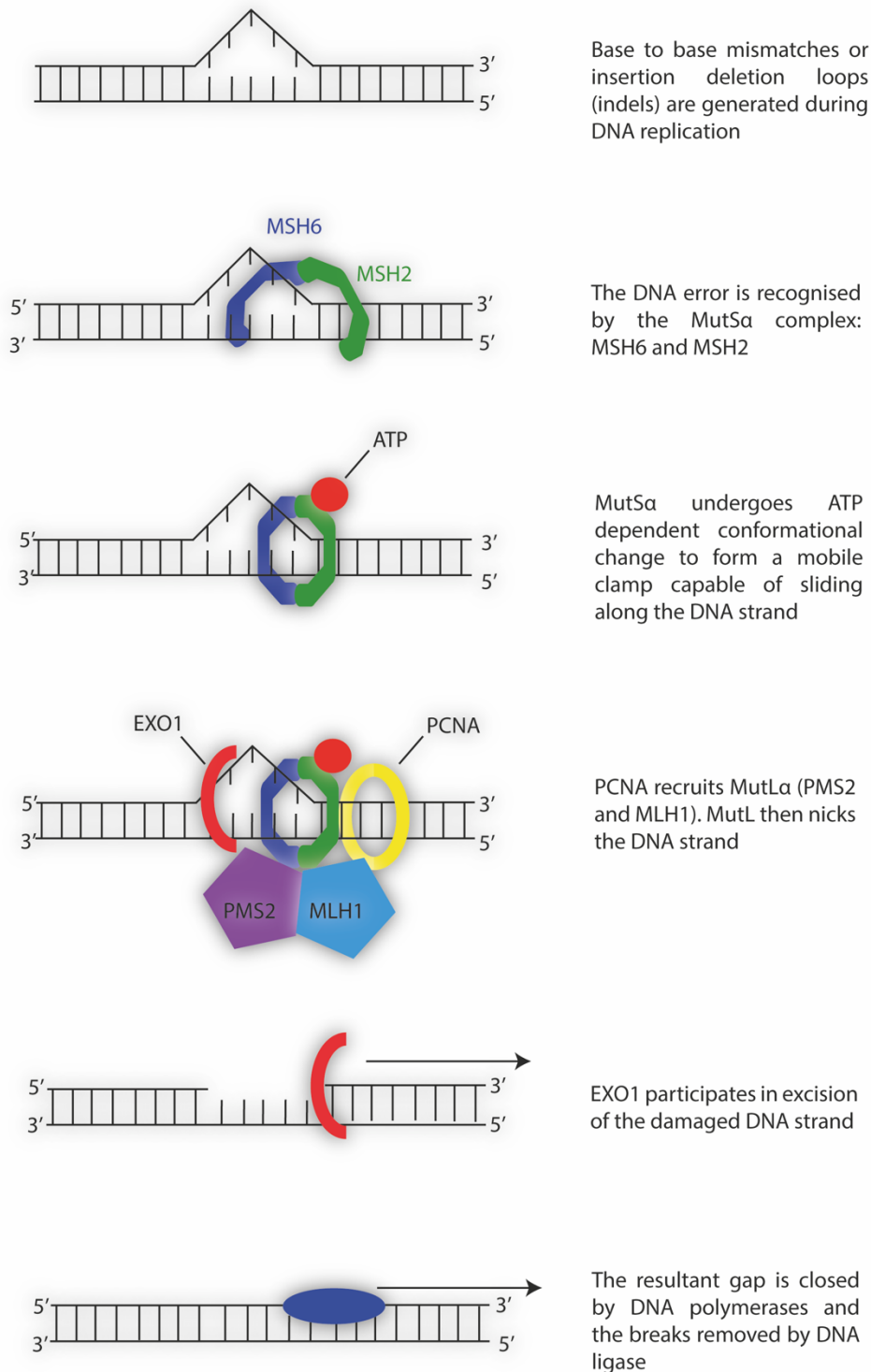


Figure 4-1. Diagrammatic representation of the MMR system in eukaryotic cells. The MMR system is responsible to correcting base-base mismatches, and insert-deletion loops that are introduced during genomic replication. Loss of the MMR system results in a mutator phenotype and genomic instability known as microsatellite instability (MSI). MMR: Mismatch Repair Figure reproduced from [6]

4.1.2. Mismatch Repair Deficiency in Endometrial and Colorectal cancer

Mismatch repair deficiency is a hallmark of several different cancers, including CRC, EC, gastric and ovarian [283]. As discussed above, it can arise either due to inherited germline DNA defects or from alterations that are incurred somatically.

4.1.2.1. Lynch Syndrome

Lynch syndrome (LS) is the most common hereditary cause of CRC and overall, one of the most common familial cancer syndromes [284]. It is an inherited disease of mismatch repair dysfunction, and sufferers are at increased risk of developing one or more of a spectrum of cancers [25]. Most commonly these are colorectal and endometrial, but gastric and ovarian cancers are also associated [284]. LS is an autosomal dominant disease in which sufferers inherit a defective allele of one of the four main MMR genes: *MLH1*, *MSH2*, *MSH6* and *PMS2* [285-288]. Additionally, germline mutations of the *EPCAM* gene can lead to epigenetic silencing of *MSH2* [289]. Associated cancers, which typically, but not universally exhibit MSI (cases due to *MSH6* defects less commonly display this), arise following somatic loss of the second MMR allele. Overall, LS has an incidence of 1:1000 in the general population and accounts for 2.5% of all CRCs [290-292]. LS results in a lifetime risk for developing CRC of 60-70% in men and 30-40% for women, who also have a 40-80% risk of developing EC [293]. The exact phenotype exhibited by individual patients varies, and is dependent on the underlying genetic alterations (Table 4.1) [284].

2301 **Table 4-1. Genetic alterations of Lynch Syndrome and the associated phenotype.** The exact phenotype
2302 exhibited by LS sufferers depends on the underlying genetic abnormality. Reproduced from [284]

Genetic Defect	Phenotype	Refs.
Heterozygous <i>MLH1</i> mutation	Predominance for CRC development; non-colon cancers less frequent than with <i>MSH2</i> defect	[293, 294]
Heterozygous <i>MSH2</i> mutation	Higher frequency of extracolonic cancers	[269, 293, 294]
Heterozygous <i>MSH6</i> mutation	Overall, lower cancer risk than <i>MLH1</i> and <i>MSH2</i> carriers. High frequency of endometrial cancers	[293]
Heterozygous <i>PMS2</i> mutation	Excess of colonic polyps, but lower frequency of cancer development	[295, 296]
Heterozygous <i>EPCAM</i> mutation	Results in silencing of <i>MSH2</i> expression. Associated lower risk of extracolonic cancers	[289]
Monoallelic <i>MLH1</i> epimutation	Similar phenotype to <i>MLH1</i> mutation. <i>MLH1</i> epimutations can be inherited, but most arise <i>de novo</i>	[297]
Biallelic mutation in any of the four MMR genes	Results in Constitutional Mismatch Repair Deficiency (CMMR-D) Syndrome: characterised by very early onset paediatric cancers	[298, 299]

2303
2304 A number of preventative strategies have been recommended for LS patients, including daily aspirin
2305 chemoprophylaxis and colonoscopy screening every two years [27]. Prophylactic hysterectomy and
2306 bilateral salpingo-oophorectomy may be advised for women who are past child bearing age [300].

2307
2308 **4.1.2.2. Sporadic Mismatch Repair Deficiency**
2309 Mismatch repair deficient colorectal and endometrial cancers can also arise due to somatic alterations
2310 [301]. Most commonly this is biallelic hypermethylation in the *MLH1* promoter region, which leads to
2311 epigenetic gene silencing [298, 302]. In CRC, but not EC, this is commonly associated with *BRAF*
2312 V600E hotspot mutation [303]. Whilst overall 15-20% of all CRCs are MMRd, the prevalence varies
2313 by stage: 20% in stage II disease, 12% stage III and relatively rare in stage IV (4%) [304]. Sporadic

MMRd CRCs tend to be proximally located, poorly differentiated and more commonly exhibit mucinous morphology [301].

Approximately 20-30% of ECs are MMRd, and 75% of these are of somatic aetiology. Similar to CRC, MMRd EC are most commonly found to be of mucinous histology [304]. Unlike CRC, there is not a similar association with *BRAF* mutations. Rather, high frequencies of truncating *PTEN*, *PIK3CA*, *ARID1A* and *RPL22* mutations are observed [67].

4.1.3. Mismatch repair deficiency and the tumour immune microenvironment

A key component of the anti-cancer immune response is T-cell recognition of specific tumour antigens [171, 305, 306]. During the process of malignant transformation, cells acquire TAAs [171]. These can be either due to genomic mutations, which cause the formation of novel peptides that are processed into neoantigens [307], or epigenetic alterations that cause aberrant expression of antigens normally only present during embryonic development [174, 308]. Expression of neoantigens results in the recruitment of tumour-specific adaptive immune cells, namely T and B cells. These are primed to recognise TAAs in draining lymph nodes, or within the tumour at tertiary lymphoid structures (TLSs) [309, 310]. Priming of T cells requires dendritic cells, which are activated in response to locally released proinflammatory cytokines [311]. Dendritic cells present neoantigens via MHC class I (MHC I) and/or MHC class II (MHC II) complex, and predominantly interact with T helper (CD4⁺) cells [312, 313]. T helper cells in turn activate cytotoxic T cells. Additionally, B lymphocytes differentiate into plasma cells, capable of antibody production, upon antigen recognition [314, 315]. There is also some evidence that B lymphocytes are able to participate in antigen presentation [316]. The immunogenicity of a given cancer, and hence the composition of the immune microenvironment, is in part determined by the tumour's mutational landscape [171]. It is recognised that a higher tumour mutational burden (TMB) is associated with an increased number of potential neoantigens [317]. In CRC, MMRd is associated with a 100-1000 fold increase in the mutation burden [87]. This accumulation of frameshift mutations results in an increased frequency of neoantigens, which induces a robust antitumour response. Consequently, MMRd CRCs are characterised by high numbers of

tumour-infiltrating lymphocytes (TILs) [48]. The active tumour immune response is counterbalanced by immunosuppressive immune checkpoints that include PD-1, PDL1, CTLA4, LAG-3 and Indoleamine 2,3-dioxygenase-1 (IDO1) [111].

There has been much work demonstrating the association between T lymphocytes and clinical outcome in cancer. CD8⁺ T cell infiltrate, as well as CD3⁺ and CD45⁺ cell infiltrate, has been correlated with improved overall survival in cancers of different histology and anatomical site, including colorectal (see Chapter 3) [1, 146, 153], breast [318, 319], ovarian [320] and gastric [258, 321]. In addition to being a prognostic marker, T lymphocytes may also have a predictive role for identifying patients who will respond to novel immunotherapies, such as ICIs. In melanoma, it was shown that response to anti-PD1 therapy was best predicted by CD8⁺ cell infiltrate at the tumour invasive margin [322, 323]. These results suggest that ICI response is dependent on an immunologically active tumour microenvironment.

4.1.4.MMR status as a biomarker in cancer treatment

4.1.4.1. Prognostic biomarker

In CRC, MMRd is a well-established marker of improved clinical outcome for early-stage disease: associated with both a reduced risk of recurrence and an improved overall survival [49]. Gryfe and colleagues were the first to report this. They demonstrated that MSI CRCs had a lower chance of metastasising to lymph nodes or distant organs [324]. These findings were later confirmed by Mouradov and colleagues, who showed that MSI is associated with improved disease-free survival (DFS) whilst CIN status is independently associated with worse outcomes [325]. In a large meta-analysis of 2,935 stage II and III CRCs, MSI was associated with a better overall survival (OS) compared to MSS with hazard ratio (HR) of 0.65 (95% CI, 0.59 – 0.85) [95]. Gustadenghi and colleagues confirmed this in a more recent meta-analysis of 12,782 patients showing that MSI was associated with improved OS odds ratio (OR) of 0.6 (95% CI 0.53 -0.69). and disease-free survival (DFS) (OR = 0.58, 95% CI 0.47-0.72) [326]. Whilst MMRd is associated with improved outcomes in

early-stage CRC, the converse may be true in the metastatic setting [304]. It has been suggested that this is because metastatic lesions far more commonly harbour *BRAF* mutations (35% of metastatic MMRd CRCs, compared with 24% of early stage MMRd CRCs), which have a detrimental effect on clinical outcome [327].

Despite the fact that MMRd is actually more common in endometrial cancer than in colorectal cancer, its prognostic value in this disease has received less attention, and there is a lack of data to show similarly associated improved outcomes. Indeed, studies have suggested that there is either no prognostic association [328], or possibly that MMRd is associated with a worse clinical outcome [329], though the small size of most precludes firm conclusions.

4.1.4.2. Predictive biomarker

In addition to use as a prognostic biomarker, MMRd has been proposed as a predictive biomarker for cytotoxic chemotherapy-based treatments, and is a proven biomarker for novel immunotherapies. It has been suggested that MSI predicts an absence of benefit from 5-fluorouracil (5-FU); a cytotoxic chemotherapy agent that works through inhibition of thymidylate synthase [330] and is the backbone of systemic therapy for CRC. Initially it was shown by Ribic and colleagues that 5-FU adjuvant chemotherapy was beneficial for stage II/III MSS CRCs, but not MSI CRCs [331]. These findings were initially supported by other studies [332]; however, the largest study to examine this found no predictive value of MMRd for fluoropyrimidine benefit [333]. In EC, it has been proposed that MMRd may confer a predisposition to responding to radiotherapy [334], but as of yet it has not been conclusively shown to predict benefit from chemotherapy agents.

The high mutational burden seen in MMRd, which could potentially induce a potent anti-tumour response, may account for its well-established utility as a predictive biomarker for immunotherapy. The advent of ICIs represent a considerable therapeutic advance in the field of oncology [13]. In particular, these therapies have produced dramatic results in melanoma and non-small cell lung cancer patients [335, 336], which are two cancers typified by very high mutation rate [78]. Initial trials

testing ICIs on unselected CRC patient cohorts produced disappointing results [337, 338]. However, further work that investigated stratifying patients according to tumour MMR status produced notable results. An initial phase 2 clinical trial was conducted to investigate the efficacy of PD-1 inhibition (pembrolizumab) on forty-one patients with metastatic, treatment refractory cancers; this included 11 MMRd CRC patients, 21 MMRp CRC patients and 9 MMRd non-CRC patients [208]. The results showed an objective response rate of 40% in the MMRd CRC and 71% of MMRd non-CRC patients. There was no evidence of response in MMRp patients [208]. The results of this trial, and others [339-341] led to the FDA to approve the PD-1 inhibitor for unresectable or metastatic MMRd cancers regardless of their anatomical site [342].

4.1.5. Rationale for study and aims for this chapter

Mismatch repair deficiency is a well described cause of genomic instability that is common in both endometrial and colorectal cancers [304]. In CRC, evidence conclusively demonstrates that MMRd portends an improved prognosis, which is probably at least in part a function of the dense T cell infiltrate associated with this particular form of genomic instability. However, whilst there is a lack of studies that have examined prognostic association EC, it would seem that MMRd does not convey the same survival benefits. I first sought to examine this using high-quality patient cohorts. I next examined the consequences of MMRd on the immune cell infiltrate in both CRC and EC. By doing it so, it may be possible to explain why there exists a discordance in the prognostic association in these two cancer types.

I aimed to first clarify the prognostic association of MMRd in early-stage CRC and EC by conducting a systematic review and meta-analysis. To explain the discordance in prognostic association, I wanted to quantify the T-cell infiltrate (given the depth of literature showing their association with prognosis in cancer) in MMRd CRC and EC tumours using multispectral immunofluorescence quantification on FFPE samples. Finally, I aimed to confirm and expand upon findings from FFPE tissue using a multi-omics approach with publicly available TCGA data.

4.2. Methods

4.2.1. Identification of studies for meta-analysis

For the meta-analysis, the included studies were found by search PubMed using the terms “endometrial” OR “colorectal” AND “mismatch repair” OR “microsatellite instability” AND “prognosis”. Included results were limited to clinical trials published between 2000 and 2021, so as to ensure consistency of MSI/MMRd classification and a period during which clinical treatment and patient outcomes remained broadly consistent for both these cancer types.

4.2.2. Multispectral immunofluorescence quantification of T cell infiltrate

T-cell infiltrate in EC and CRCs was quantified using a previously published and validated method; I conducted this work at University of Leiden with the group who had developed it [2]. The CRC and EC cases that underwent analysis were selected based on tissue availability, and no matching according to stage was attempted. The staining methodology is shown in Figure 4.2. Full face FFPE tumour sections of 4 μ m were deparaffinised and then underwent antigen retrieval using citrate buffer solution. The slides were then incubated overnight with primary antibodies for FoxP3 (236A/E7, 1:25, ThermoFisher Scientific) and CD8 (4B11, 1:50, ThermoFisher Scientific). Subsequently, they were washed in Tris-buffered Saline (TBS) before incubation for 1 hour with appropriate secondary fluorophore antibodies (CFF633 goat anti-mouse IgG1, 1:100, Avantar and CF555 goat anti IgG2b, 1:100, Sigma-Aldrich for FoxP3 and CD8 detection respectively). Following further washes in TBS, the slides were incubated with conjugated antibodies (primary antibody with fluorophore) for CD3 (D7A6E, 1:50, Cell Signalling Technology) and pan-cytokeratin (Pankeratin (C11), 1:50, Cell Signalling Technology and cytokeratin, AE1/AE3, 1:50, ThermoFisher Scientific) for six hours. The slides were then washed again, DAPI was applied as a nuclear stain and the slides were mounted using ProLong Gold (Thermofisher).

2448

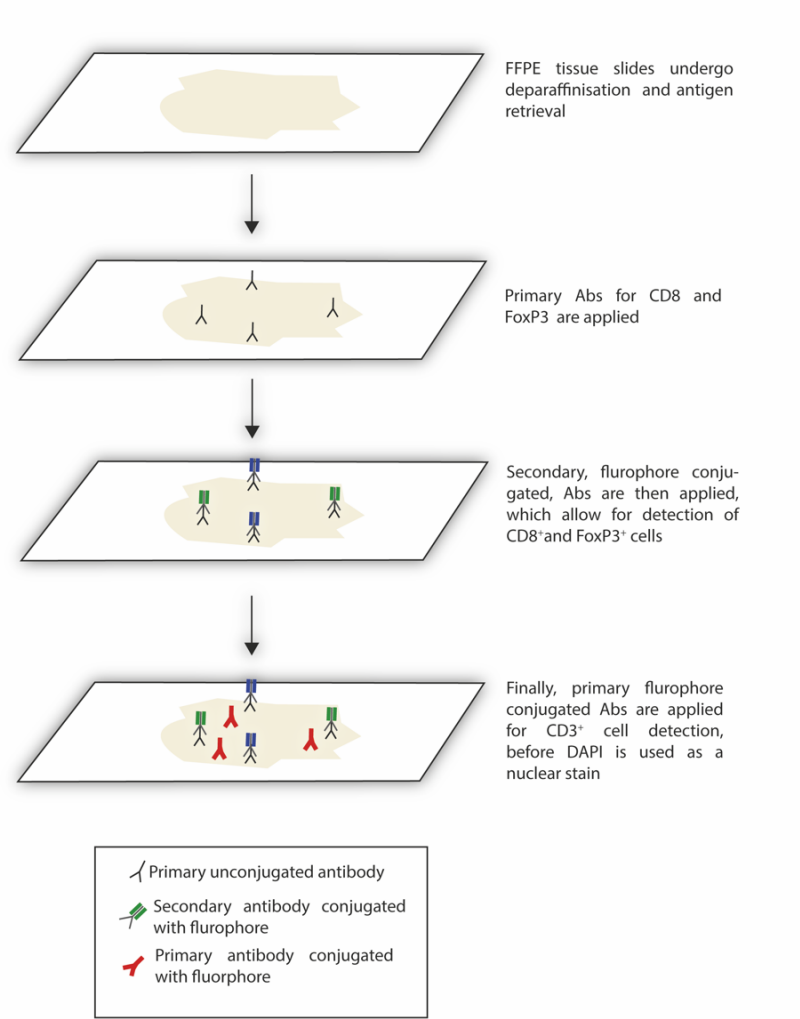


Figure 4-2. **T cell infiltrate quantification by multispectral immunofluorescence staining** in full faced FFPE tumour sections. This quantification method has been previously published and validated [2]. FFPE: formalin fixed, paraffin embedded

4.2.3. Image acquisition and analysis

Stained tissue slides were scanned using the Vectra Polaris Imagine system (Perkin Elmer). Initially, the entire tissue area was scanned and then sub regions were selected to be imaged in depth at 40X magnification. I selected the regions of interest using the Perkin Elmer software; two regions from the centre of the tumour and one from the invasive margin were chosen. After obtaining full slide scans, the channels representing fluorophores of stained immune cells were unselected and using images only showing staining of cytokeration (to allow for identification of epithelial and stromal regions) were using for region of interest selection. When selecting regions, I was blinded to the genomic status, i.e. whether they were MMRp or MMRd, Lynch syndrome or sporadic. Following image capture, quantification was conducted using the same software platform. First, the quantification software was trained to segment the tumour images into epithelium and stroma by the presence of the cytokeratin stain. Next, cell segmentation training was done using DAPI as a nuclear marker and an approach of refinement to define the cell membrane. The following cell populations were defined on the basis of their marker positivity: FoxP3⁺ T cells (CD3⁺ membrane marker and FoxP3⁺ nuclear marker), CD8⁺ T cells (CD3⁺ and CD8⁺ positive) and CD3⁺ cell positive T cells. For each image segment, the following metrics were obtained: total number of cells, number of negative cells and number of marker positive for each population defined above. Results were presented as cells per megapixel and from there it was possible to calculate the cell density (cells/mm²). Prior to analysis, all the images underwent visual inspection to ensure correct tumour areas had been captured and to remove those with obvious artefact effect that may disturb immune cell quantification.

4.2.4. TCGA analysis

MC3 MAFs containing curated TCGA colorectal (COADREAD) and endometrial (UCEC) cancer whole exome sequencing (WES) data [212] were downloaded from the NIH Genomic Data Commons (<https://gdc.cancer.gov/about-data/publications/mc3-2017>) along with tumor MSI status determined by testing of microsatellite markers[343, 344]. Germline MMR gene mutations were identified by analysis of controlled-access BAMs for pathogenic variants[345]. Controlled FASTq files for RNAseq data were downloaded from the NIH GDC data portal (<https://portal.gdc.cancer.gov/>) and

processed by Salmon [211] to obtain relative transcript abundance (Transcripts Per Million – TPM) for downstream analysis (Methods 2.1.1). Estimates of tumour-infiltrating immune cell populations and related data were downloaded from the supplementary material of Thorsson et al [346]. To calculate absolute estimates of immune cell populations, proportional estimates were multiplied by leucocyte fraction, the latter being an estimate of the total fraction of the bulk tumour due to immune cells. Methylation data for TCGA cases were downloaded from the NIH GDC. Probes mapping to *MLH1* which were most differentially methylated according to tumor MSI status were identified using the TCGAbiolinks package [347].

4.2.5. Cell lines

Human endometrial adenocarcinoma cell lines AN3CA, HEC-1A, HEC-108, EFE-184, EN, KLE and SNG-II cells were a kind gift from Konstantin Dedes (University of Zurich, Switzerland). EN, HEC-1B, Ishikawa, MFE-296 and NOU1 cells were shared by Britta Weigelt (previously Cancer Research UK, Lincoln's Inn Fields UK). MFE-280 cells were purchased from the European Collection of Authenticated Cell Cultures. HEC-251 and HEC-59 cells were obtained from the JCRB cell bank. All cell lines were culture in an atmosphere of 5% CO₂ at 37°C. SNG-II and AN3-CA were cultured in DMEM medium with 10% foetal bovine serum (FBS); EN in DMEM with 20% FBS; MFE-280 and MFE-296 in 40% RPMI/40% DMEM with 20% FBS; KLE and Ishikawa in DMEM/F12 with 10% FBS; HEC-1-A and HEC-1-B in MEM with 10% FBS; NOU1, HEC-59 HEC-251 and HEC-108 in MEM with 15% FBS, and EFE-184 in RPMI with 10% FBS. MSI status was determined by Promega MSI analysis system and JAK1 mutation by WES done in-house or from by the Cancer Cell Line Encyclopaedia: <https://depmap.org/portal/ccle/>

4.2.6. Interferon-gamma stimulation

Cells for interferon-gamma (IFN- γ) stimulation were plated at 1x10⁶ cells in 25cm² cell culture flasks, grown to 70-80% confluency, and serum starved for 24 hours before treatment with IFN- γ

(554617, BD Biosciences) at 75ng/ml or equivalent volume of PBS for 16 hours before cell lysis and protein collection.

4.2.7. Western blotting

Details of Western blot methodology in Methods 2.5

4.2.8. Gene Set Enrichment Analysis

Differential gene set enrichment analysis (GSEA) was preformed using the DESeq package [348].

The input for this package is normalised RNAseq data. GSEA is a three-step process: firstly, enrichment scores (ES) are calculated, which reflect the degree to a set of genes are overrepresented at the extremes (either top or bottom) of a ranked list of gene. Secondly, the statistical significance of the ES is calculated using an empirical, phenotype-based permutation, which preserves the complex correlations observed in gene expression data; the result is a nominal *P* value. Finally, the estimated significance for each set is adjusted for multiple testing [349].

4.2.9. Statistical analyses

The statistical significance of differences between immune cell infiltrates in FFPE samples was determined using the non-parametric Mann-Witney test, and interaction testing conducted using the Aligned Rank ANOVA test. Statistical tests were two-sided, and hypothesis testing was performed at the 5% significance level. The meta-analysis was conducted, and Forest plots generated in R using the “meta” package. All statistical analyses were performed R, Version 4.0.1 (<https://cran.r-project.org>), using packages ‘meta’[350], ‘ggplot2’[351], ‘DESeq2’[348], ‘clusterProfiler’[352], ‘TCGAbiolinks’[347], and ‘MAFtools’[353].

4.3. Results

4.3.1. Mismatch repair deficiency and clinical outcome in colorectal and endometrial cancer

A number of studies, including two meta-analyses [95, 326], have shown that MMRd is associated with more favourable prognosis in early-stage CRC this. However, only one meta-analysis, published in 2013 has examined the association between MMRd and prognosis in EC – its conclusion was that no firm inference could be drawn [328]. Given this paucity of research, I wanted to examine the prognostic association of MMRd in CRC and EC by systematic review and meta-analysis. To ensure high-quality data, I restricted the search to clinical trials, which have multiple advantages for biomarker research when compared to non-trial cohorts. In total, the search strategy identified 13 publications that reported clinical trials pertaining to MMRd/MSI and clinical outcome in CRC [90, 225, 229, 325, 333, 354-360] (only 7 of these reported the association between MMRd and OS). There were four studies that reported the association between MMRd and clinical outcome in EC [264, 361-363]. Consistent with previous publications, MMRd was associated with a statistically significantly lower risk of disease recurrence in CRC (HR=0.62; 95% CI=0.51–0.72%; $P<1e-04$) (Figure 4.3). In contrast, in EC there was no significant association between MMRd and risk of recurrence (HR=1.15; 95% CI=0.72–1.58%; $P=0.23$) (Figure 4.2). There was no evidence of significant heterogeneity in the studies for either CRC nor EC ($I^2 = 35\%$, $\tau^2 = 0.0097$, $P = 0.09$ and $I^2 = 31\%$, $\tau^2 = 0.0571$, $P = 0.23$ respectively) Formal testing revealed a statistically significant interaction between tumour type and the prognostic value of MMRd ($P_{\text{INTERACTION}} = 0.02$). Similar results were demonstrated for OS, which was better for MMRd/MSI cases in CRC (HR=0.70; 95% CI=0.57–0.83%; $P=1.2e-05$) but not in EC (HR=1.51; 95% CI=0.87–2.16%; $P=0.34$) (Figure 4.4). Again, there was no evidence of significant heterogeneity for either included CRC or EC studies ($I^2 = 33\%$, $\tau^2 = 0.0103$, $P = 0.17$ and $I^2 = 50\%$, $\tau^2 = 0.19$, $P = 0.11$ respectively). These results show that MMRd is associated with significantly reduced risk of recurrence and improved overall survival in early-stage CRC, but not in EC, with a statistically significant difference between the two cancer types. I next wanted to explore whether the divergent association with prognosis could be explained by differences in the T cell infiltrate of MMRd CRCs and ECs.

2553

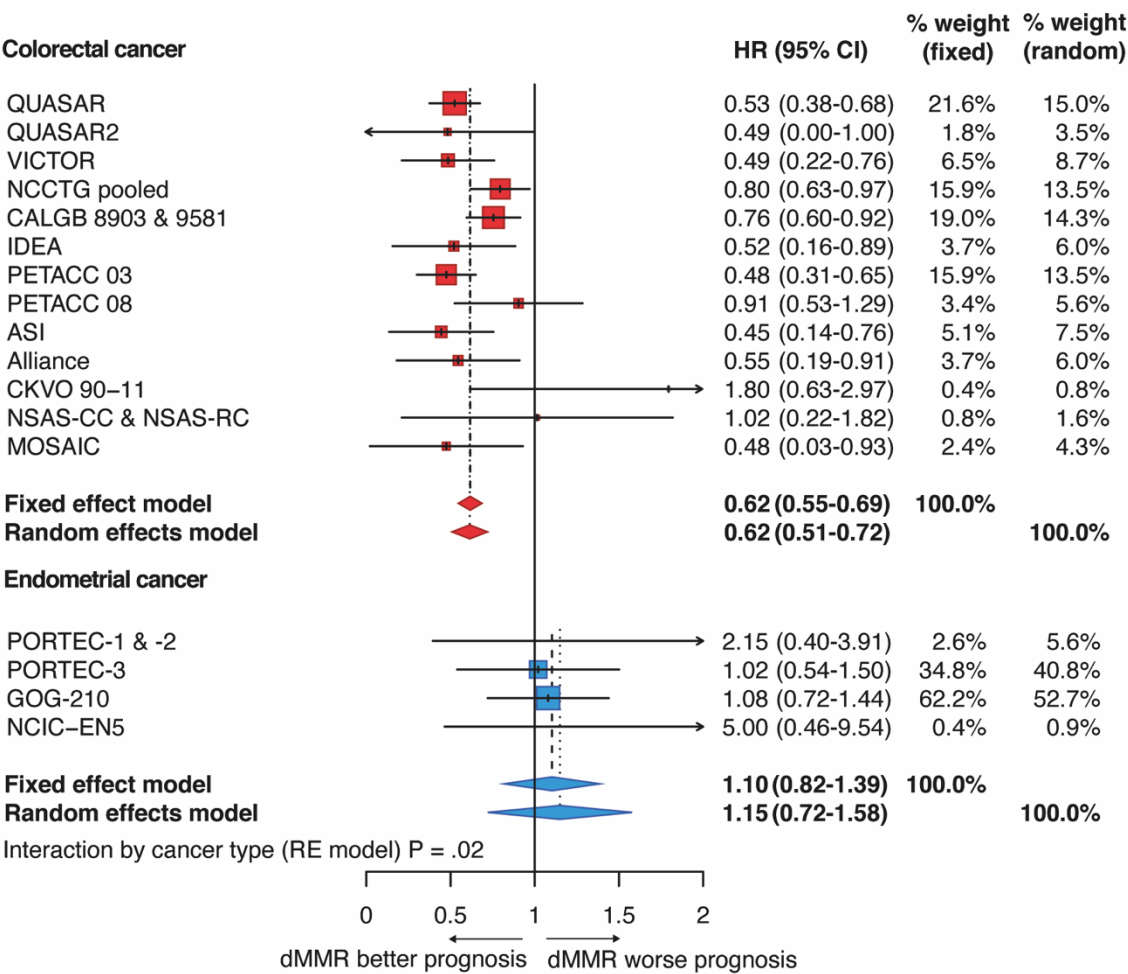


Figure 4-3. Forest plot showing the meta-analysis results for the association between MMRd and risk of disease recurrence in CRC and EC. MMRd was significantly associated with a reduced risk of disease recurrence in CRC (HR 0.62 95%CI 0.55-0.69) but not in EC (HR 1.10 95%CI 0.82-1.39). MMRd: mismatch repair deficient. MMRp: mismatch repair proficient. c. 95%CI: 95% confidence intervals

2554

2555

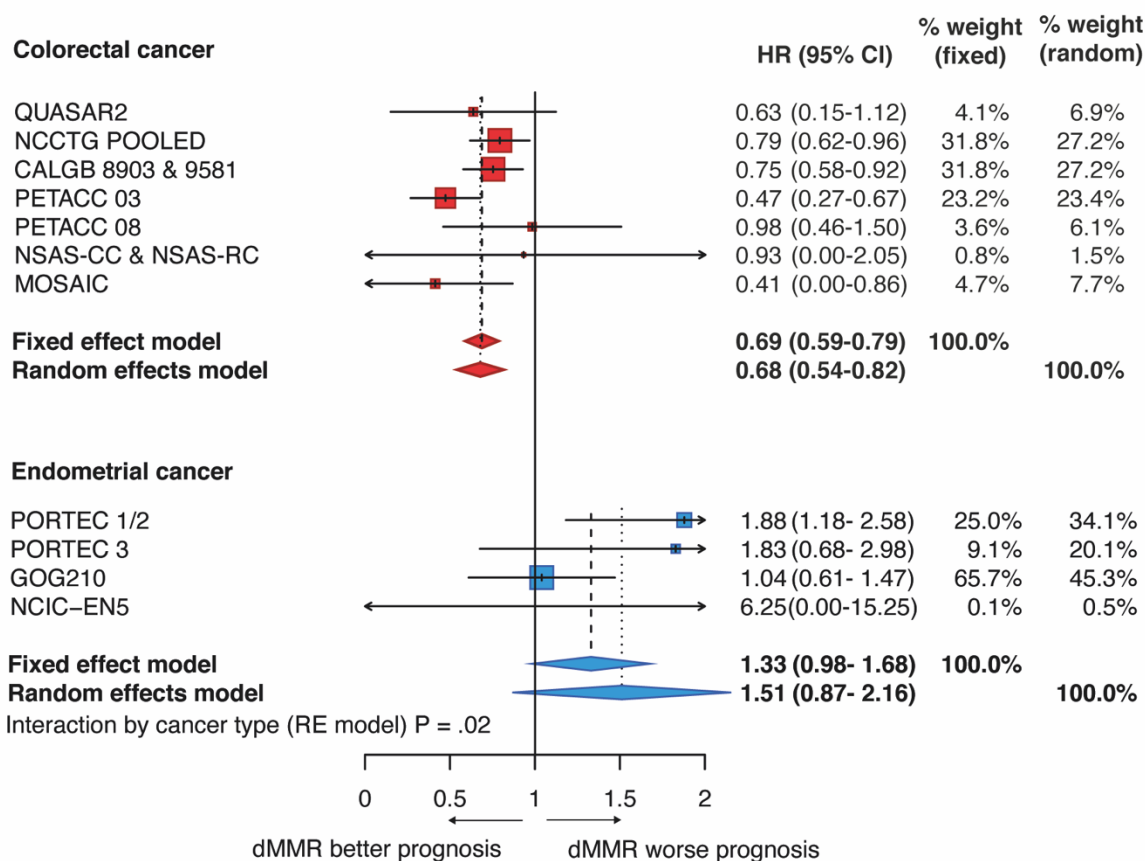


Figure 4-4. Forest plot showing the meta-analysis results for the association between MMRd and overall survival in CRC and EC. MMRd was significantly associated with a reduced risk of disease recurrence in CRC (HR 0.69 95%CI 0.59-0.79) but not in EC (HR 1.33 95%CI 0.98-1.68). MMRd: mismatch repair deficient. MMRp: mismatch repair proficient. HR: hazards ratio. 95%CI: 95% confidence intervals

4.3.2. T cell infiltrate in CRC and EC

As shown in the previous Chapter and in multiple other studies, MMRd colorectal cancers display increased lymphocytic infiltration [48], a feature which is widely thought to explain their better prognosis. However, whether MMRd has a similar impact on the immune infiltrate in endometrial cancer has not been conclusively determined. Given the discordant prognosis of MMRd colorectal and endometrial cancers shown above, I wished to examine this. To do so, as part of a collaboration between the University of Oxford, University of Manchester and Leiden University, I curated a cohort of CRCs and ECs of known MMR status, and, in the case of MMRd tumours, known aetiology (Lynch vs sporadic). The cases were selected based on tissue availability. There was no attempted matching based on stage of the EC and CRC. ECs are frequently diagnosed at an earlier stage than

CRC; this may have impacted immune cell quantification in FFPE tissue. Details of the cases used are provided in Table 4.2. Due to the non-experimental nature of the CRC samples, detailed pathological annotation was not available.

Table 4-2: Tissue samples used in the IF quantification of immune cells in CRC and EC FFPE Tissue samples

	CRC		EC	
	No.	%	No.	%
Total	90	100	177	100
MMR status [†]				
MMRp	33	73.2	69	39.0
MMRd (Sporadic)	33	10.9	51	28.8
MMRd (Lynch Syndrome)	24	15.3	57	32.2
Unknown	0	0	0	0
Disease stage				
I	0	0	85	48.0
II	0	0	21	11.9
III	0	0	23	13.0
IV	0	0	3	1.7
Unknown	90	100	45	25.4

These cases were analysed by multispectral immunofluorescence staining for immune markers CD3, CD8, CD68, FoxP3, the epithelial cell marker pan-cytokeratin (CK) and the nuclear stain DAPI (see section 4.2.2 details), to permit quantification of immune infiltrate within the intraepithelial and intrastromal compartments (Figure 4.5, A-C). In keeping with previous literature [90, 111, 364], both intraepithelial CD3⁺ and CD8⁺ cell infiltrates were significantly greater in MMRd cancers compared to MMRp cancers for CRC (Figure 4.6A, Table 4.3). This was true for measurements both at the tumour centre and invasive margin. Similar analysis of endometrial cancers also revealed greater density of infiltrating CD3⁺ and CD8⁺ cells in MMRd cases of both germline and sporadic cause; however, the difference was much smaller and not statistically significant. Formal testing for an MMRd*cancer type interaction confirmed a statistically significant difference for these comparisons

2589 (CD3⁺ cells tumour centre $P_{\text{INT}} = 0.0003$ and invasive margin $P_{\text{INT}} = 0.0003$; CD8⁺ cells tumour
2590 centre $P_{\text{INT}} = 0.0073$ and invasive margin $P_{\text{INT}} = 0.037$; all P_{INT} derived by Aligned Rank ANOVA).
2591
2592 In CRC, MMRd sporadic cases had significantly greater levels of intraepithelial CD8⁺ and CD3⁺ cell
2593 infiltrate compared to MMRp cases at the both the invasive margin and tumour centre (Table 4.3). In
2594 contrast, Lynch Syndrome MMRd CRCs had significantly greater intraepithelial CD8⁺ and CD3⁺ cell
2595 infiltrate at the tumour centre, but not at the invasive margin (Table 4.3). For EC, neither sporadic nor
2596 Lynch syndrome MMRd cases had statistically significant greater levels of intraepithelial CD8⁺ and
2597 CD3⁺ cell infiltrate compared to MMRp cases in either the tumour centre or invasive margin.
2598 Consistent with previous publications [365], my results showed increased levels of intrastromal and
2599 intraepithelial CD8⁺ cell levels in Lynch syndrome MMRd ECs compared to sporadic MMRd ECs,
2600 which was borderline statistically insignificant ($P=0.051$).

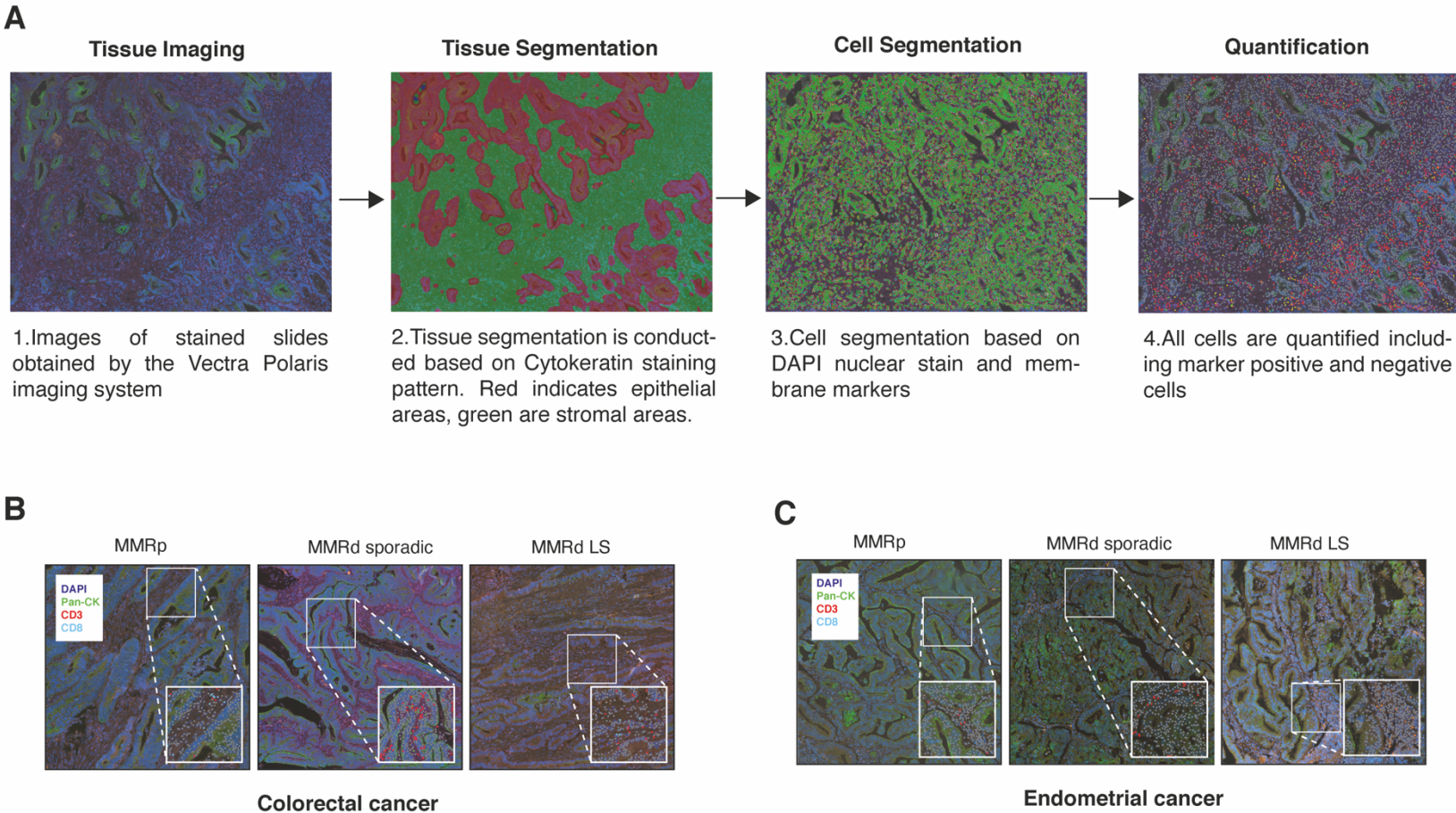


Figure .4-5. **Multispectral immunofluorescent quantification of immune cells in FFPE CRC and EC tissue** A. Shows the imaging and quantification methodology used to analyse multispectral IF stained FFPE tumour slides. B. Example images of CRC MMRp, MMRd Sporadic and MMR germline tumours. C. Example images of CRC MMRp, MMRd Sporadic and MMR germline tumours. C. Example images of EC MMRp, MMRd Sporadic and MMR germline tumours. MMRp: Mismatch repair proficient. MMRd: mismatch repair deficient. LS: Lynch Syndrome. Pan-CK: pan cytokeratin. DAPI: 4',6-diamidino-2-phenylindole

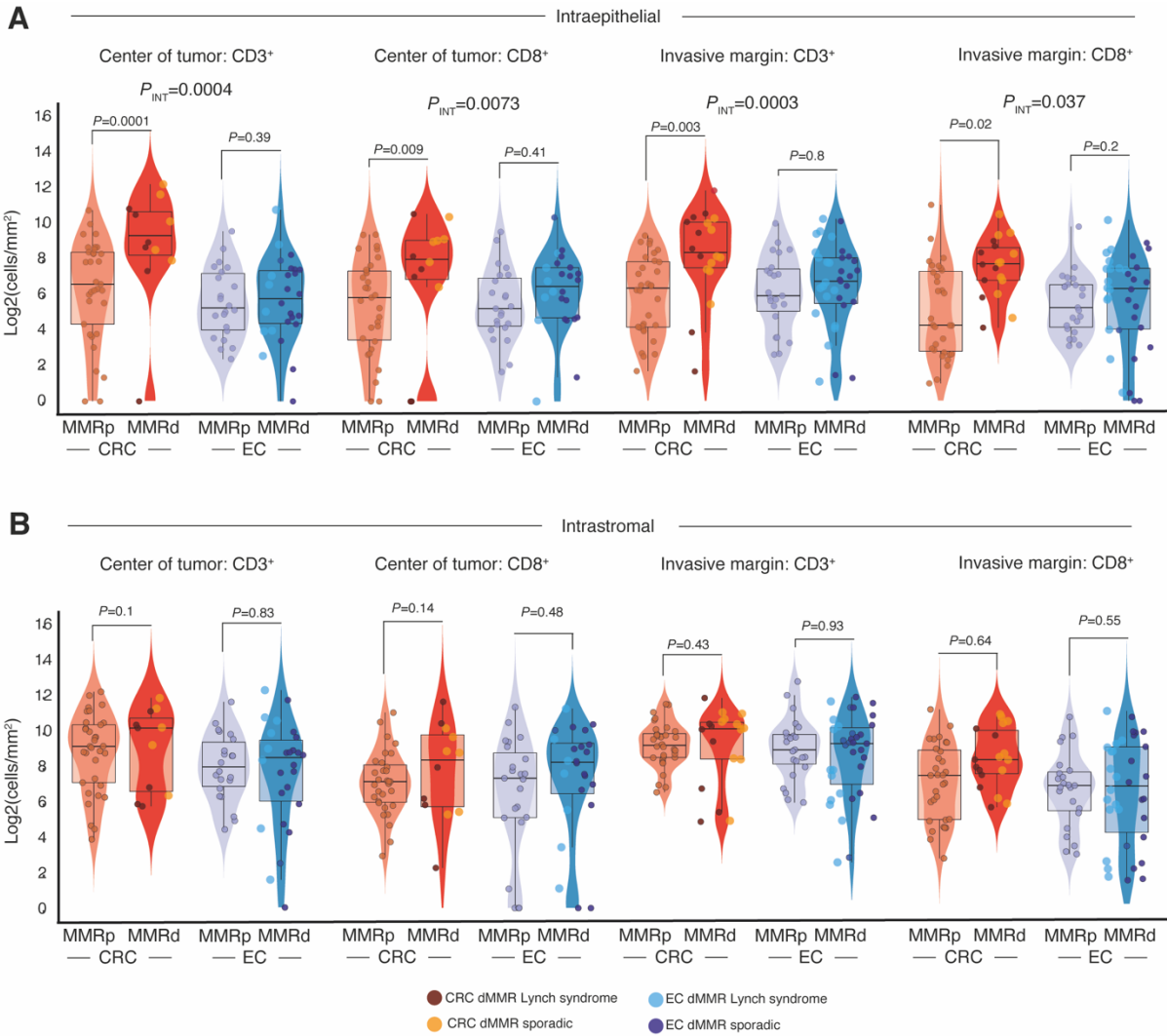


Figure 4-6. T-cell infiltrate in CRC and EC MMRp and MMRd tumours quantified by multispectral IF staining A) Intraepithelial CD3⁺ and CD8⁺ cell infiltrate. B) Intrastromal CD3⁺ and CD8⁺ cell infiltrate. Individual data points coloured according to whether MMRd was sporadic or Lynch syndrome. P values between MMR groups for each cancer type determined by non-parametric Mann-Whitney testing and P for interaction by aligned rank ANOVA, MMRp – mismatch repair proficient; MMRd – mismatch repair deficient. CRC: colorectal cancer. EC: endometrial cancer

The consequences of DNA mismatch repair deficiency in colorectal cancer and endometrial cancer

2615

2616

	Colorectal Cancer						Endometrial cancer						
Marker-compartment	MMRp	MMRd sporadic	MMRd LS	<i>P</i> (MMRp vs MMRd SP)	<i>P</i> (MMRp vs MMRd LS)	<i>P</i> (MMRp vs MMRd all)	MMRp	MMRp sporadic	dMMR LS	<i>P</i> (pMMR vs dMMR SP)	<i>P</i> (pMMR vs dMMR LS)	<i>P</i> (pMMR vs dMMR all)	<i>P</i> _{INTERACTION} (cancer type*MMR status)
Centre tumour													
CD3+ intraepithelial	110.5 (20.2–246.1)	738.6 (189.9–1099.1)	837.0 (270.1–1380.5)	3.37e-05	0.017	1.2e-04	74.1 (44.1–177.8)	168.5 (61.2–245.2)	144.9 (52.8–380.0)	0.50	0.32	0.39	4.3e-04
CD8+ intraepithelial	61.8 (12.4–190.6)	218.1 (125.1–586.8)	243.9 (113.1–665.2)	0.015	0.029	8.8e-03	55.6 (29.1–115.6)	61.5 (15.5–158.9)	135.1 (59.4–348.6)	0.98	0.059	0.41	7.3e-03
FOXP3+ intraepithelial	20.3 (10.2–73.9)	41.0 (34.7–54.8)	18.4 (9.9–88.9)	0.27	0.97	0.53	13.7 (3.4–31.2)	15.5 (2.2–44.2)	21.5 (13.0–43.9)	0.89	0.12	0.52	0.69
CD3+ intrastromal	571.4 (309.3–782.3)	926.7 (418.4–1229.1)	826.2 (127.1–1013.0)	1.8e-03	0.80	0.10	488.8 (213.5–782.7)	508.7 (330.9–1051.2)	566.0 (214.1–1141.4)	0.61	0.87	0.83	0.01
CD8+ intrastromal	176.6 (34.0–542.7)	341.7 (208.0–1473.5)	287.4 (252.7–529.3)	0.08	0.20	0.14	112.4 (49.5–230.8)	116.0 (21.7–244.2)	206.6 (95.8–801.2)	0.83	0.15	0.48	0.73
FOXP3+ intrastromal	181.4 (77.0–487.8)	283.5 (139.6–393.4)	169.2 (116.1–288.4)	0.75	0.96	0.86	46.8 (30.3–102.7)	42.5 (16.9–170.5)	87.6 (21.7–244.2)	0.63	0.68	0.90	0.95
Invasive margin													
CD3+ intraepithelial	95.6 (19.3–373.9)	1114.3 (366.0–3185.6)	448.7 (221.0–1205.3)	1.0e-03	0.12	3.3e-03	37.3 (15.2–146.2)	44.9 (21.3–168.5)	93.4 (14.8–240.9)	0.94	0.48	0.78	3.1e-04
CD8+ intraepithelial	56.3 (10.0–172.3)	515.0 (297.6–556.0)	150.3 (32.0–257.0)	4.6e-03	0.70	0.02	36.4 (17.9–122.5)	91.9 (25.1–154.7)	89. (48.3–217.5)	0.16	0.26	0.20	0.037
FOXP3+ intraepithelial	22.8 (6.9–53.1)	32.2 (7.2–51.7)	295.0 (41.7–599.0)	1.0	0.036	0.25	14.7 (2.9–39.2)	6.1 (2.8–41.4)	11.1 (3.33–19.5)	0.56	0.49	0.51	0.03
CD3+ intrastromal	546.0 (131.3–129.5)	1143.1 (573.7–2397.7)	614.2 (69.1 – 1231.1)	0.018	0.69	0.43	246.4 (113.9–656.4)	160.3 (87.4–431.2)	497.2 (21.0–1488.8)	0.86	0.62	0.93	0.073
CD8+ intrastromal	140.3 (62.1–271.0)	449.5 (140.7–721.9)	358.4 (59.6–1092.0)	0.32	0.89	0.64	262.1 (33.5–446.6)	300.2 (119.1–563.9)	254.8 (36.6–816.9)	0.31	0.48	0.55	0.35
FOXP3+ intrastromal	174.4 (62.0–407.6)	164.8 (35.2–288.1)	476.6 (253.0 – 541.8)	0.47	0.078	0.52	45.5 (33.9–180.2)	9.4 (0.0–63.9)	60.8 (7.2–180.4)	0.06	0.64	0.15	0.22

2617

2618

2619

2620

Table 4-3. Tumour infiltrating lymphocyte density by marker and compartment according to cancer type and MMR status. Values shown are sample medians with interquartile rangers (IQR) in parentheses. *P* values calculated by non-parametric Mann-Whitney *U* Test. Test for interaction were performed by aligned rank ANOVA. MMRp: Mismatch repair proficient. MMRd: Mismatch repair deficient

2621

There was no significant difference seen in FoxP3⁺ cell infiltrate between MMRd and MMRp cancers in either EC or CRC. This was true for all comparisons for intrastromal and intraepithelial, and for tumour centre and for invasive margin (Figure 4.7, Table 4.3).

These results suggest that MMRd has significantly different consequences for the T cell based immune response in these two cancer types. These findings have not been reported previously, and may provide some explanation as to why MMR status does not have the same significant association with prognosis in EC that it does in CRC. I wanted to confirm these findings, and explore possible explanations for it.

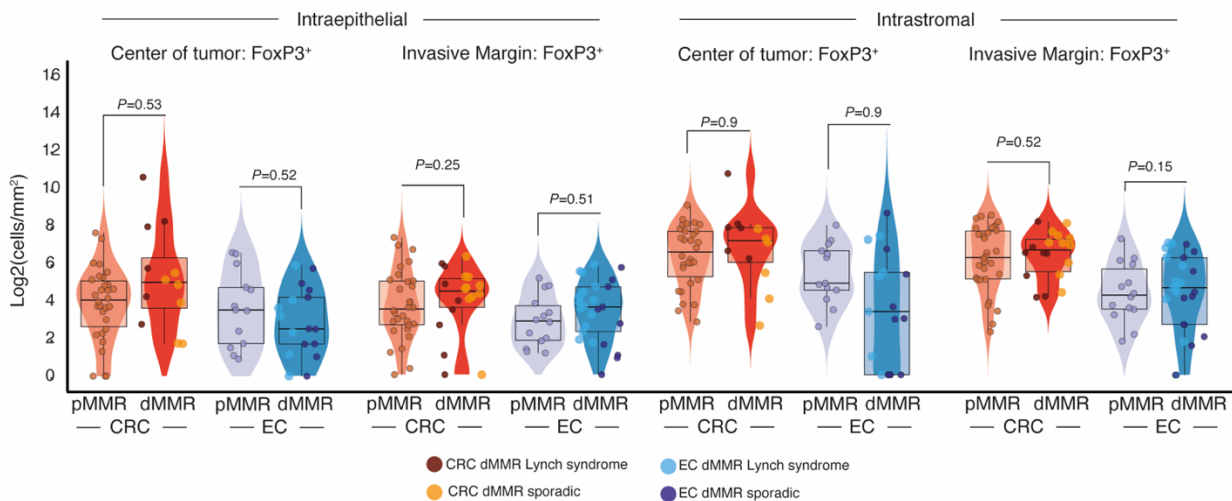


Figure 4-7. Intrastromal and intraepithelial FoxP3⁺ cell infiltrate in CRC and EC MMRp and MMRd tumours; quantified by multispectral IF staining. *P* values between MMR groups for each cancer type determined by non-parametric Mann-Whitney testing and *P* for interaction by Aligned Rank ANOVA. MMRp – mismatch repair proficient; MMRd – mismatch repair deficient. CRC: colorectal cancer. EC: endometrial cancer

4.3.3. Validation in TCGA series

The Cancer Genome Atlas was a large US-based collaborative initiative to perform molecular profiling of multiple cancer types, with results reported in a series of landmark publications [67, 87]. For colorectal and endometrial cancers, TCGA studied 633 and 560 cases respectively, using a combination of whole exome sequencing (WES), RNAseq, methylation array, SNP array and reverse phase protein array (RPPA). Following the landmark publications for both of these cancers data continued to be updated and are publicly downloadable for research purposes. While comprehensive, the TCGA datasets do not include *in situ* immunohistochemical or immunofluorescence-based immune cell quantification as used in the previous section. However, an alternative method to estimate infiltrating leukocytes is the determination of cell populations from bulk RNAseq data. Several methods have been published which perform this, one of the most widely-used of which is CIBERSORT [366]. Using input matrix of reference gene expression signatures, CIBERSORT generates estimates of the relative proportion of cells of interest. CIBERSORT analyses had previously been conducted on CRC and EC samples from TCGA and published by Thorsson and colleagues in a pan-cancer immune study [346]. To classify the TCGA cancers according to MMR status I used tumour microsatellite stability, as this is widely accepted as a surrogate for MMR loss (with exception of some endometrial cancers with germline *MSH6* defects).

Analysis of immune cell proportions estimated by CIBERSORT according to cancer types and MMR status was consistent with the above results. Comparison of selected immune cell populations, immune checkpoint expression and IFN- γ pathway activation are represented by the heatmap in Figure 4.8. MSI CRC and ECs both had statistically significantly greater proportions of total lymphocytes, CD8⁺ T cells and activated CD4⁺ cells compared to their MSS counterparts; however, the magnitude of increase was far larger in CRC, again with statistically significant interactions between MMR status and cancer type evident ($P_{\text{INT}}=0.002$ and $P_{\text{INT}}=0.005$ respectively; Figure 4.9). In addition to a greater enrichment of T cell infiltrate in MSI CRCs compared with ECs, M1 macrophages showed a greater upregulation in CRC than in EC. This result was notable given the

known association of M1 macrophages with a Type 1 inflammatory response driven by interferon gamma (IFN- γ) signalling pathway [116].

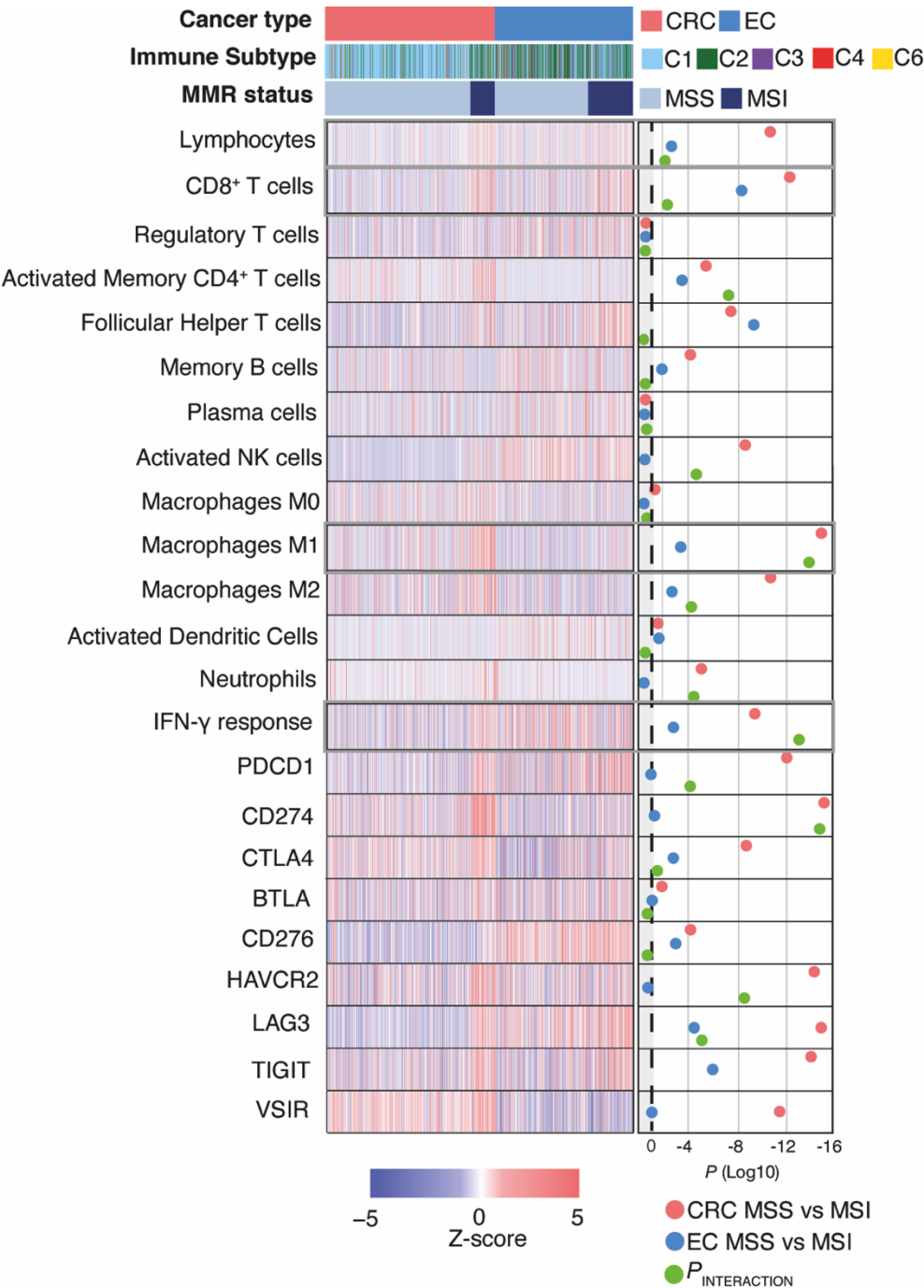


Figure 4-8. Heatmap showing relative tumour proportion of tumour-infiltrating immune cells, IFN γ response score and expression of immune checkpoints. The cancer type, MMR status and immune subgroup (C1-C6). MSI: microsatellite unstable. MSS: microsatellite stable. CRC: colorectal cancer EC: endometrial cancer. NK cells: Natural killer cells. The right column graphically represents the Log10 transformed P value for comparisons of CRC MSS Vs MSI (red), EC MSS. Vs MSI (blue) (both determined by Mann-Whitney U Test) and P_{INTERACTION} value (green) (aligned rank ANOVA test).

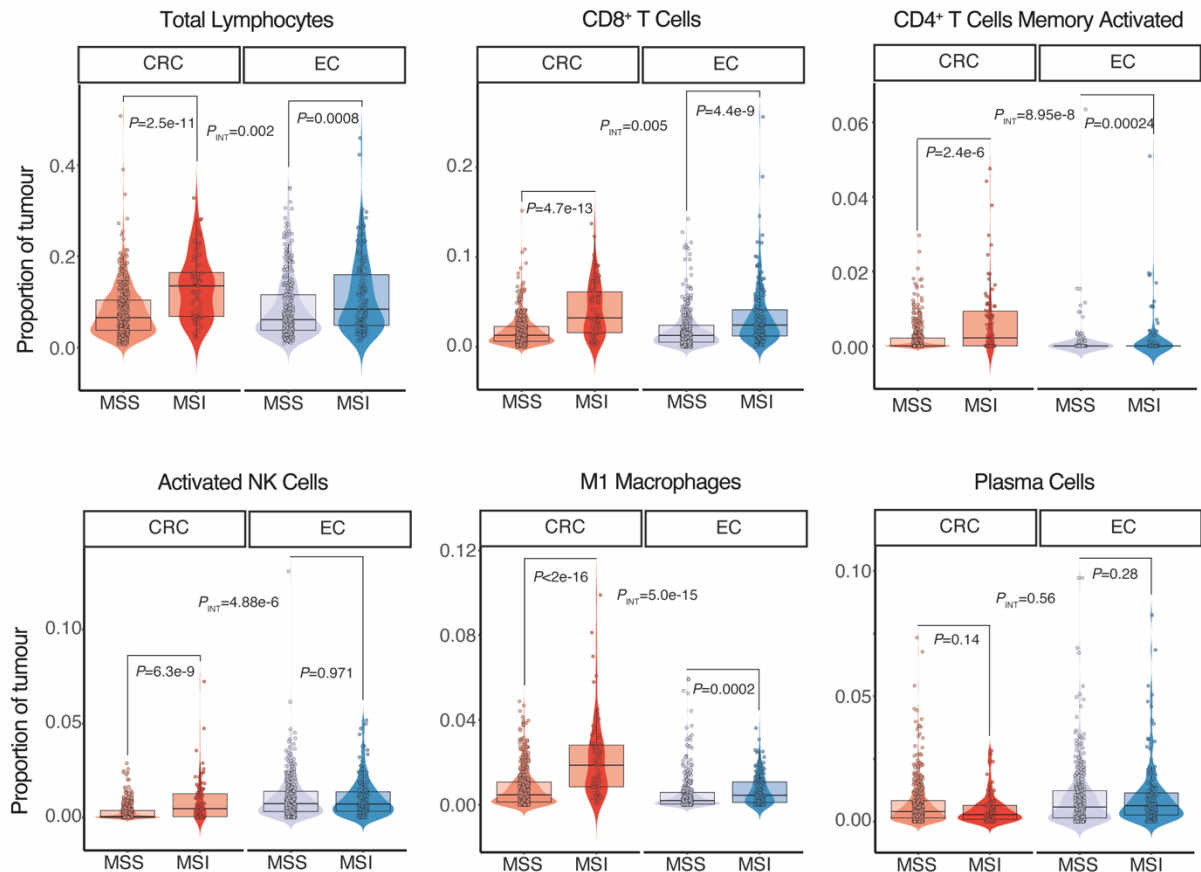


Figure 4-9. **Immune cell infiltrate in TCGA CRC and EC cancers** determined by CIBERSORT processing of bulk RNAseq data. P values between MMR groups for each cancer type determined by non-parametric Mann-Whitney testing and P for interaction by Aligned Rank ANOVA. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable. NK cells: natural killer cells

4.3.3.1. Immune checkpoint expression

It has previously been documented that the active immune microenvironment observed in CRC is counterbalanced by an upregulation of inhibitory signalling molecules, such as PD-1, PD-L1, CTLA4, LAG3 and IDO-1 [111]. This corresponds to the favourable response of MMRd CRCs to immune checkpoint inhibition therapy. As anticipated the immune checkpoints *PDCD1* (PD1), *CD274* (*PDL1*), *CTLA4*, *HAVCR2*, *LAG3*, *TIGIT* and *VSIR* were all substantially and significantly upregulated in in MSI CRCs. Interestingly, similar comparison in EC revealed a more mixed picture; with significantly upregulated markers (*CTLA-4*, PD-1, *LAG3* and *TIGIT*) showing less marked upregulation in MSI cases when compared to CRC, and some markers either (*PD-L1*, *HAVCR2* and *VSIR*) showing borderline significant or non-significant differences between groups. This variation

between the two cancers was reflected in statistically significant interaction tests for all immune checkpoints examined (Figure 4.10).

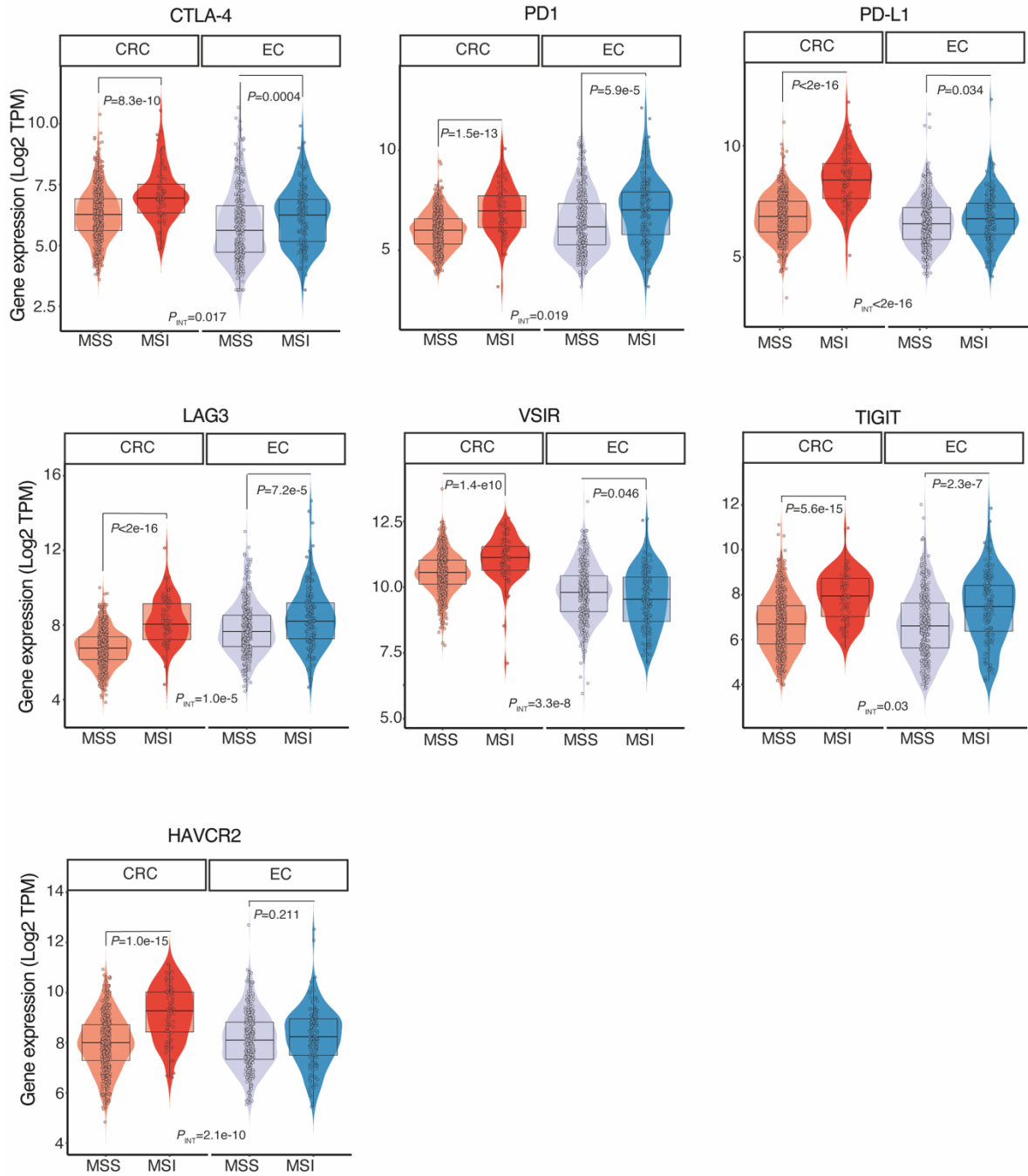


Figure 4-10. **Immune checkpoint expression in TCGA CRC and EC RNAseq data.** P values between MMR groups for each cancer type determined by non-parametric Mann-Whitney testing and P for interaction by Aligned Rank ANOVA. LAG3: Lymphocyte Activation Gene 3. VSIR: V-set Immunoregulatory Receptor. TIGIT: T-cell Immunoreceptor with Immunoglobulin and ITIM domain. HAVCR2: Hepatitis A Virus Cellular Receptor 2

2688 **4.3.3.2. Gene set enrichment analysis**

2689 Both IF analysis of FFPE tumours, and analyses of TCGA RNAseq suggest that the consequences of
2690 MMR loss on the T cell immune infiltrate differs significantly between CRC and EC. Next, I wanted
2691 to explore the consequences of MMR deficiency on cellular processes in the two cancers. Using an
2692 agnostic, hypothesis free approach, gene set enrichment analysis (GSEA) was conducted on RNAseq
2693 data specifically looking to identify dysregulated pathways in MSI tumours compared with MSS in
2694 both CRC and EC. Strikingly, while the gene sets examined included all biological processes curated
2695 in the MSigDB database all of the top ten most dysregulated gene sets in CRC MSI tumours were
2696 immune related (Figure 4.11). Of these, five were directly related to IFN- γ signalling – a process of
2697 critical importance to the antitumour immune response –, whilst a further three were related to a cell
2698 mediated immune response. In contrast, none of the most significantly upregulated pathways in EC
2699 MSI tumours were related to immune response but rather involved basic cellular metabolic processes.
2700 Indeed, it was noticeable that four pathways related to interferon alpha (IFN- α) response and type 1
2701 interferon signalling were significantly downregulated. The top 30 up- and downregulated gene sets
2702 for CRC and EC MSI Vs MSS GSEAs are shown in Appendix 7.1. These results support that MMR
2703 deficiency is associated with an active immune microenvironment in CRC, but not in EC. And
2704 furthermore, the immune response in MSI CRC is driven by IFN- γ signalling.

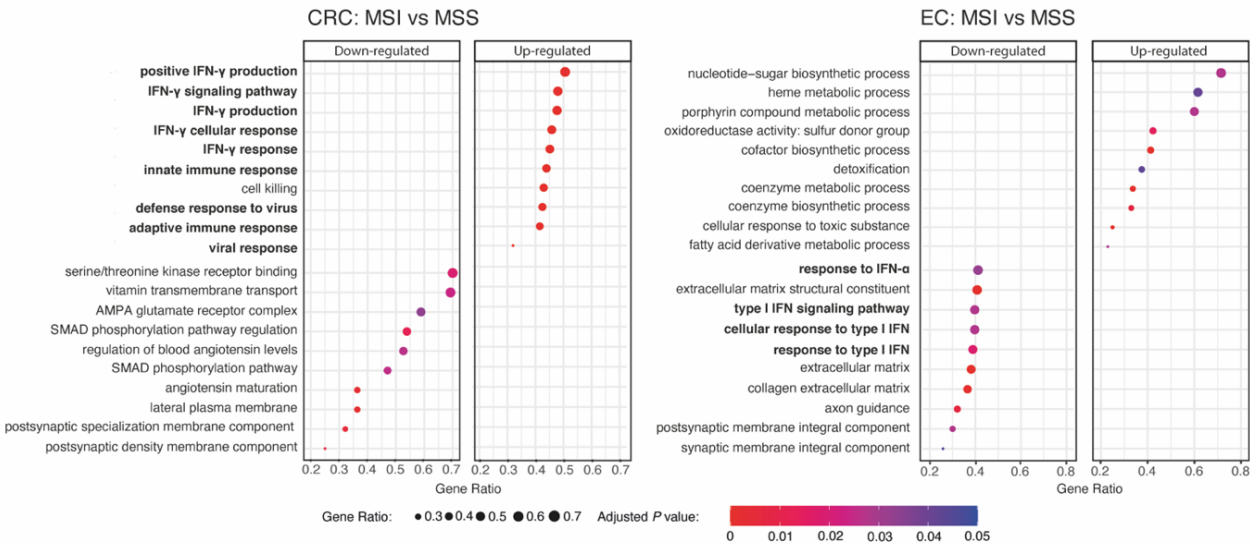


Figure 4-11. **Gene set enrichment analysis** results comparing CRC MSI cancers with CRC MSS cancers, and EC MSI cancers with EC MSS cancers. Notably, in CRC MSI cancers a number of the upregulated pathways are related to IFN- γ signalling.. CRC: Colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable

4.3.3.3. Mutational landscape in MMRd CRC and EC

As discussed previously, a substantial body of previous work indicates that the immune microenvironment in cancers is at least partly determined by a tumour's underlying mutational landscape. A higher mutational burden leads to increased frequency of neoantigens that may stimulate an adaptive immune response, and this believed to explain the high levels of TILs seen in MMRd CRCs [92]. Given the above results, I next sought to analyse whether the differences in the adaptive immune response seen between CRC and EC could be explained by differences in their mutational landscape. As expected MSI CRCs had a greater overall number of mutations, as well as indel mutations and SNV mutations specifically, than MSS CRCs ($P < 2.2 \times 10^{-16}$ for all three). Similar comparison in EC also revealed greater tumour mutational burden (TMB), SNV and indel mutations in MSI tumours, however the magnitude of these increases was considerably smaller, with significant interactions evident for each (Figure 4.12).

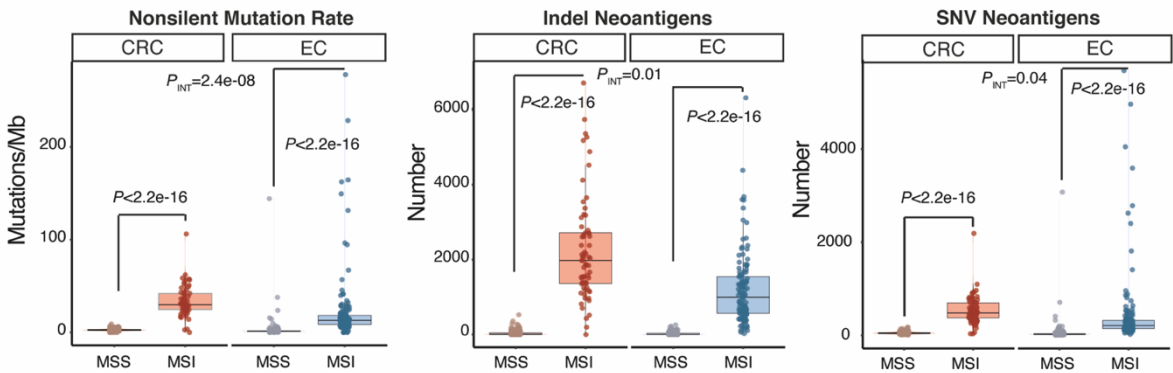


Figure 4-12. **The total non-silent mutation rate, indel neoantigens and SNV neoantigens for TCGA CRC and EC MSS and MSI cancer.** P value for comparisons within cancers generated by non-parametric Mann-Whitney testing and P for interaction by Aligned Rank ANOVA. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable. SNV: single nucleotide variants

As noted previously, the most common cause of sporadic MMRd among cancers of the colorectum and endometrium is *MLH1* loss due to promoter hypermethylation. I next wanted to explore whether this difference in mutational burden was possibly due to different levels of *MLH1* promoter methylation and consequently gene expression between MSI CRC and EC. As expected, the *MLH1* promoter was more methylated, and gene expression lower, in MSI cancers of both types (Figure 4.13A,B). However, *MLH1* expression levels were actually higher in the MSI CRCs vs MSI ECs, and *MLH1* promoter methylation relatively less. These data suggest that differences in *MLH1* expression levels are not the explanation for the discordance in SNV and indel mutation burden seen between MSI CRC and EC.

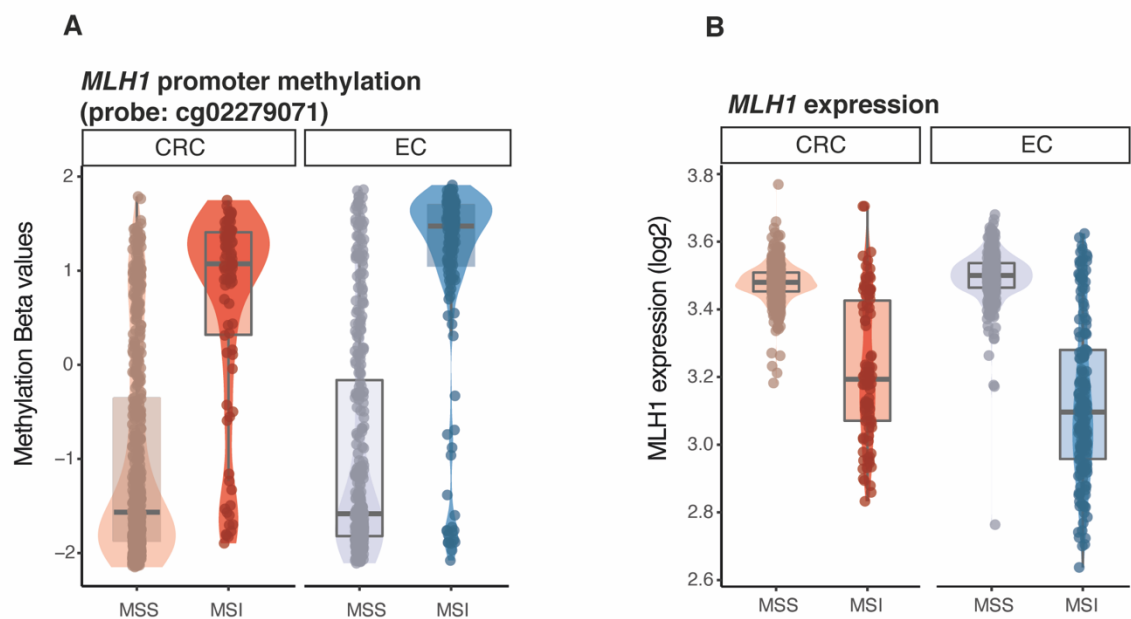


Figure 4-13. **MLH1 methylation and expression** A. *MLH1* promoter methylation in TCGA CRC and EC MSS and MSI cancers B. *MLH1* expression in TCGA CRC and EC MSS and MSI cancers. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable

4.3.4. Immune escape mutations

Increased mutational load typically results in an increase of neoantigens capable of stimulating an adaptative antitumour immune response [92]. However, cancers may acquire mutations that disrupt antigen presentation and consequently reduce immunogenicity. Collectively, these are known as immune escape mutations. For example, previous work has shown that *Beta-2-microglobulin (B2M)* mutations are enriched in MMRd CRCs. *B2M* is the light chain component of MHC Class I and as such is integral in this antigen presentation pathway; loss of function mutations in *B2M* consequently lead to reduced antigen presentation [183]. I wanted to explore whether immune escape mutation frequency could explain the differential immune response between CRC and EC.

A curated list of candidate immune escape mutations was obtained from the publication by Grasso and colleagues, who had analysed 1,211 CRCs [183], and mutation data for CRCs and ECs obtained from TCGA MC3 MAFs. These MC3 MAFs files were produced from a seven mutation caller pipeline that was designed with the intention to unify results across the TCGA cancer sets and reduced artefact risk [212]. Comparison of MSI CRC vs MSI EC showed an enrichment of loss of

function mutations (defined as truncating or frameshift mutations) in genes involved in antigen presentation, namely: *HLA-B*, *B2M*, *NLRC5* and *TAP2* ($P=0.01$ to $P=2.3e-06$, Fisher exact test) (Figure 4.14). The enrichment for this type of immune escape mutation possibly reflects selection pressure exerted by the tumour immune response. Interestingly, in MSI ECs, these genes were less commonly mutated, with analysis instead demonstrating enrichment of *JAK1* and *RPL22* loss of function mutations (mutational frequencies of 23.8% versus 4.2% of cases $P=3.0e-04$; 31.9% versus 8% of cases $P=0.02$ respectively for MSI EC vs MSI CRC; P values calculated by Fisher exact test). It was not possible to tell this whether the *JAK1* mutations were bi-allelic from the MAF data. This would require copy number analysis and estimation of tumour purity. However, data from COSMIC and CCLE showed that *JAK1* p.K860Nfs*16 in EC are most commonly heterozygous. The observed enrichment for *RPL22* mutations in MSI ECs is consistent with previous publications [67, 367]. *RPL22*, Ribosomal Protein L22, is a component of the 60S ribosomal subunit that is capable of RNA binding. In mice models, heterozygous *RPL22* loss leads to accelerated T-cell lymphomagenesis [368]. In human CRC, *RPL22* loss is associated with increased expression of the highly homologous paralog, *RPL22L1*. In turn, increased *RPL22L1* is associated with poor prognosis and 5-FU resistance. Interestingly, in cell line models induction of *RPL22L1* was found to increase O6-methylguanine-DNA-methyltransferase (MGMT) and decreased expression of *MLH1* [369]. Previous publications have confirmed the high frequency *RPL22* mutations in ECs, and suggested no association with stage, grade or clinical outcome [370]. Its role in immune escape is currently undefined, although loss of other ribosomal subunits has been shown to alter MHC class I peptide generation [368]. Interestingly, the majority of *JAK1* mutations were a recurrent frameshift at codon 860 (p.K860Nfs*16), the site of a DNA homopolymer run, potentially explaining its mutation frequency. *JAK1* (Janus-Kinase 1) is a critical component of the JAK-STAT signalling pathway, responsible for mediating cellular responses to more than 50 cytokines and growth factors [369]. In particular, *JAK1* is involved in cellular responses to IFN- γ signalling [370]. Given the absence of IFN- γ pathway enrichment seen in MSI ECs compared with MSI CRCs (Figure 4.8 above), the enrichment for *JAK1* mutations in this tumour subgroup was of particular interest.

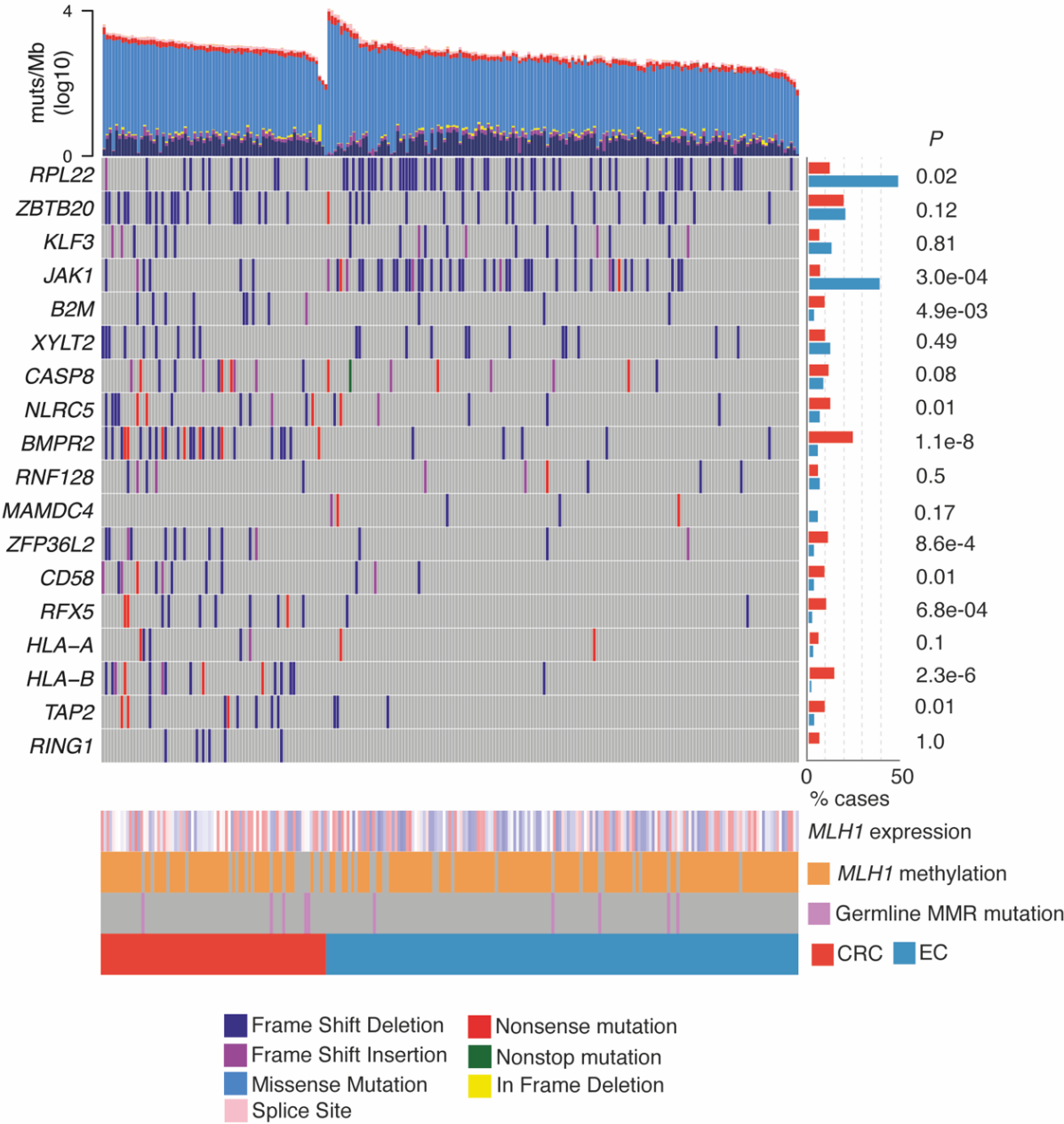


Figure 4-14. The frequency and type of immune evasion associated mutations in TCGA MSI CRC and EC. P values for comparison of mutation frequency between cancer types obtained using Fisher's exact test. CRC: Colorectal cancer. EC: endometrial cancer

4.3.5. *JAK1* LOF mutations are associated with reduced IFN- γ pathway activation in MSI ECs

In the previous section, CD8⁺ cells, CD4⁺ cells and M1 macrophages were found to be increased in MSI CRCs compared to MSI ECs, and enrichment for IFN- γ related pathways in MSI CRC but not EC was found by GSEA. Given these data, and the differences in immune escape mutations noted above, I wanted to further analyse IFN- γ pathway activation in CRC and EC. To do this I used a published IFN- γ pathway activation score, which is a composite measure of 28 genes that are upregulated in response to IFN- γ signalling [371].

CRC MSI cancers were associated with significantly increased IFN- γ pathway activation compared to CRC MSS cancers ($P=1.7\text{e-}10$) (Figure 4.15A). Contrastingly, EC MSI cancers exhibited significantly reduced IFN- γ pathway activation compared to EC MSS cancers ($P=2.3\text{e-}3$, with highly statistically significant interaction ($P_{\text{INTERACTION}}=1.4\text{e-}13$) (Figure 4.15A). To understand the relationship of immune escape mutations, including *JAK1*, MMR status and cancer type on IFN- γ pathway activation, I performed multiple linear regression (Figure 4.15B, Table 4.4). Using MSS CRCs as the referent, and IFN- γ score as the response variable, results showed that the presence of MSI in EC and CRC was associated with increased IFN- γ response score, with MSS EC also showing greater pathway activity than CRC (Figure 4.15B). Strikingly, of the published immune escape mutations included in the linear regression model, only *JAK1* LOF was significantly associated with reduced IFN- γ pathway activity ($\beta=-0.72$, $P=2.48\text{e-}07$). Stratification of MSI CRCs and ECs by *JAK1* mutational status showed the truncating LOF mutations were associated with significantly lower *JAK1* expression levels (Figure 4.15C $P=0.027$ and $P=5.0\text{e-}12$ for CRC and EC respectively). In MSI ECs, *JAK1* truncating mutations were also associated with a significant reduction in IFN- γ pathway activation. Non-truncating *JAK1* mutations did not appear to have a similar impact on the IFN- γ pathway in MSI ECs (Figure 4.15D $P=0.003$ and $P=1.7\text{e-}8$ for CRC and EC respectively).

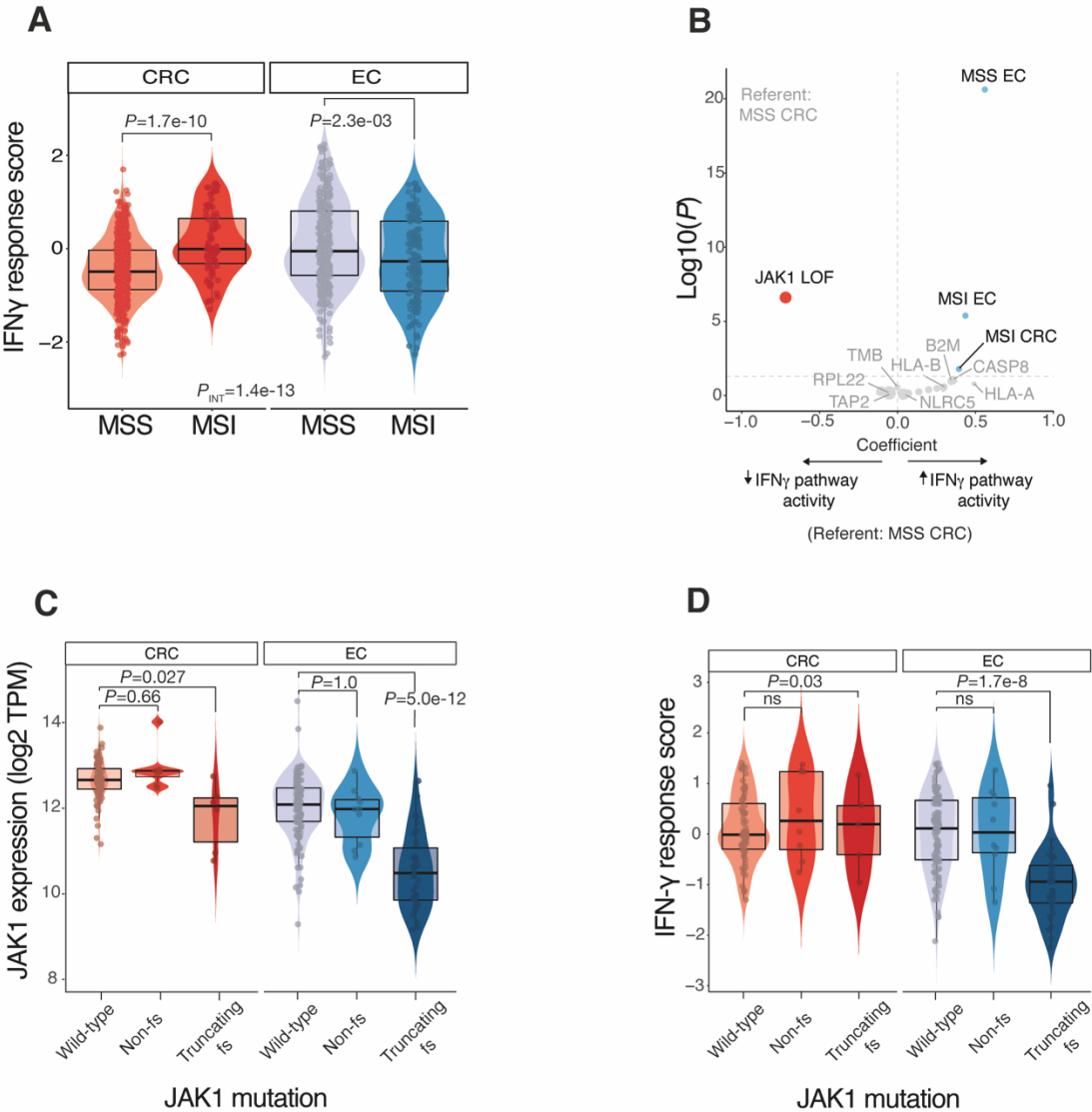


Figure 4-15. IFN- γ and JAK1 loss of function (LOF) mutations in TCGA CRC and EC. A IFN- γ pathway response in TCGA CRC and EC MSS and MSI cancers B. Linear regression examining the association of MMR status, mutational load and immune escape mutations on IFN- γ signalling C. JAK1 expression in TCGA CRC and EC MSI cancers according to JAK1 mutational status. D IFN- γ response in TCGA CRC and EC MSI cancers according to JAK1 mutational status. P values for within cancer comparison generated by Mann-Whitney test and P for interaction by aligned rank ANOVA. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite stable. MSI: microsatellite unstable. fs: frameshift mutation

Table 4-4. **Multiple linear regression analysis** of the effect of MMR status, mutational load and immune escape mutation on the IFN- γ response in TCGA CRC and EC cancers. LOF – Loss of Function. CRC: colorectal cancer. EC: endometrial cancer. MSS: microsatellite unstable. MSI microsatellite stable

Predictor	Linear regression result	
	Effect (beta)	P value
CRC MSS	Ref	NA
CRC MSI	0.39	0.016
EC MSS	0.56	<2e-16
EC MSI	0.43	4.16e-6
Non-silent mutation rate	0.00	0.27
<i>B2M</i> LOF	0.35	0.11
<i>BMPR2</i> LOF	0.04	0.81
<i>CASP8</i> LOF	0.36	0.09
<i>CD58</i> LOF	0.20	0.40
<i>HLA-A</i> LOF	0.49	0.16
<i>HLA-B</i> LOF	0.30	0.28
<i>JAK1</i> LOF	-0.72	2.48e-7
<i>KLF3</i> LOF	0.07	0.73
<i>MAMDC4</i> LOF	-0.12	0.71
<i>NLRC5</i> LOF	0.05	0.84
<i>RFX5</i> LOF	0.25	0.33
<i>RING1</i> LOF	0.29	0.44
<i>RNF128</i> LOF	0.14	0.56
<i>RPL22</i> LOF	-0.05	0.67
<i>TAP2</i> LOF	-0.05	0.86
<i>XYLT2</i> LOF	0.04	0.81
<i>ZBTB20</i> LOF	0.04	0.80
<i>ZFP36L2</i> LOF	-0.10	0.51

4.3.6. *JAK1* loss of function mutations cause protein loss and abolishes IFN- γ response *in vitro*

The analyses of TCGA RNAseq and mutational data presented above suggested that JAK1 LOF mutations attenuate IFN- γ pathway activation in MSI ECs, however confirmation of causality was not possible with these data. To address this, I next performed functional analysis of cancer cell line models *in vitro*. Our laboratory has assembled a large panel of endometrial cancer cell lines for such work. I first determined the *JAK1* mutation status of all of these using either publicly available data from the Cancer Cell Line Encyclopaedia: (<https://depmap.org/portal/ccle/>), or from whole exome sequencing (WES) conducted in our laboratory. In particular, I wanted to understand the relationship between *JAK1* mutations and MSI status in endometrial cancer cell lines. Next, I correlated these data with JAK1 protein levels by Western Blot analysis and found that six of eight cell lines with JAK1 LOF mutations had complete loss of JAK1 protein (Figure 16A). A selection of these EC lines, both MSI and MSS and JAK1 WT and mutant were selected for functional analysis. As expected, cell lines with wild-type JAK1 and normal JAK1 protein levels demonstrated clear evidence of IFN- γ pathway activation following exogenous IFN- γ stimulation, with induction and phosphorylation of STAT1 evident. In contrast, EC cell lines harbouring *JAK1* p.K860fs with loss of JAK1 protein, there was no effect on components of the JAK1-STAT pathway following IFN- γ stimulation: no upregulation of the downstream target STAT1, nor evidence of STAT1 phosphorylation. Additionally, in response to IFN- γ stimulation *JAK1* WT cell lines showed an increase in MHC Class 1 protein expression. In contrast, *JAK1* mutant HEC-59 and Ishikawa there was very low or absent MHC Class 1 expression in both unstimulated and stimulated cell lines. In MFE-296 and HEC-108 cells there was no apparent effect of IFN- γ stimulation on basal MHC Class 1 expression. These *in vitro* results suggest that *JAK1* p.K860Nfs mutations disrupt IFN- γ response in EC MSI cancers (Figure 16B). Taken in conjunction with TCGA analyses, these results support that the enrichment of recurrent *JAK1* LOF frameshift mutations found in MSI ECs may in part explain the attenuated T cell immune response in comparison to MSI CRCs through disruption of the IFN- γ signalling pathway. IFN- γ is essential for effective cytotoxic immunity: it directly induces CD8⁺ T cell differentiation and activation, as well as promoting other elements of cell-based immunity [372, 373]. Furthermore, IFN- γ has been shown to

2849 effect upregulation of MHC Class I and II molecules on APC, which may encourage the recruitment
2850 of other immune cell types, such as CD4⁺ T helper cells [374-376].

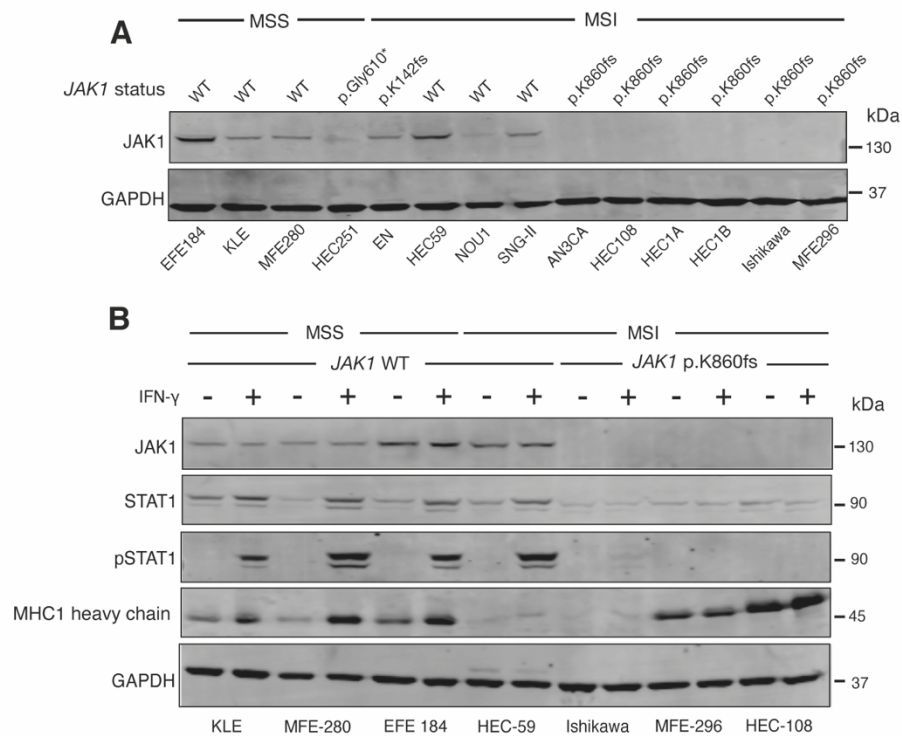


Figure 4-16. **JAK1 mutations and IFN-γ pathway activation in EC cancer cell lines in vitro.** A. JAK1 expression in EC cell lines under resting conditions. B. IFN-γ stimulation of JAK1 WT and mutant cell lines. JAK1 mutation was associated with reduced activation of the JAK-STAT pathway in response to IFN-γ

2851

4.4. Discussion

4.4.1. The consequences of MMR loss in colorectal and endometrial cancer

Genomic instability characterises most, if not all, cancers [40]. Mismatch repair loss is one particular example, and is on that is critically important in the development of colorectal and endometrial cancers. In CRC, MMRd tumours are associated with a vigorous immune response [48], improved prognosis in early stage disease [95, 326] and better response to immune checkpoint therapy in the metastatic setting [208]. In EC, the prognostic association with MMRd has not been clearly demonstrated [328], and it is unclear whether they are similarly responsive to immune checkpoint inhibition. This chapter sought to formally confirm the apparent discordance in the prognostic association of MMRd between the two cancer types and to explore whether this could be explained by differences in the T-cell based immune response and tumour genome. This was achieved using multispectral, IF quantification of tumour infiltrating T-cells in in FFPE tissue. The novel results produced by this were then further explored using TCGA data, and *in vitro* cancer cell models.

Systematic review and meta-analysis confirmed that MMRd was associated with reduced recurrence risk and improved overall survival in CRC but not in EC, with statistically significant interaction between cancer types. Tumour infiltrating T cell quantification in FFPE tumours samples showed that MMRd was associated with an enhanced T cell response in CRCs, but a far weaker difference in endometrial cancers with a significant interaction between tumour type and MMR status for both intraepithelial CD3⁺ and CD8⁺ cell densities. The discordance in immunophenotype was confirmed in TCGA analyses, where MSI CRCs were found to have greater increases in lymphocytes and CD8⁺ cytotoxic lymphocyte infiltrate than MSI ECs when compared to their MSS counterparts. Additionally, similar differences were also observed for other immune cell populations such as M1 macrophages. Agnostic gene set enrichment analysis revealed robust activation of the IFN- γ pathway in MSI CRCs, but not MSI ECs, which correlated with the frequent occurrence of *JAK1* p.K860Nfs mutations in the latter. Functional analysis confirmed these mutations are associated with absent IFN- γ pathway activation in both TCGA data and *in vitro* cell line experiments.

In combination, these results suggest that the lack of prognostic benefit associated with MMRd in EC may be in part explained by a weaker T cell immune response compared to that seen in MMRd CRCs. From the data presented above it is possible to suggest some explanations for this. The weaker immune response in MMRd ECs could be a reflection of a lower mutational burden seen in these cancers compared to CRCs. A lower mutational rate results in fewer possible neoantigens, and less immunogenic tumours [92]. One possible reason for the discordance in TMB between EC and CRC is the role of gut microbiome, which is recognised as a mutagenic agent and contributor to CRC development [75]. This would be an interesting question to explore in functional mouse models.

It was also observed that MMRd ECs exhibited lower levels of IFN- γ pathways activation, both in comparison to MMRd CRCs, but also MMRp ECs. The IFN- γ gene signature is positively associated with improved clinical outcome in many cancer types [109, 377], and has been shown to predict response to anti-CTLA4 and anti-PD-1 therapies [378-380]. IFN- γ exerts its tumour suppressor roles through direct effects on tumours cells (inhibiting proliferation, whilst increasing MHC expression and antigen presentation), and by increasing the activity of tumour infiltrating T cell populations, such as cytotoxic CD8⁺ T cells and T_H1 T helper cells [113, 114]. It has also been suggested that IFN- γ may stop metastases formation by remodelling the tumour architecture [112].

These findings have important implications for the use of MMRd as both a prognostic and predictive biomarker in EC. An active immune microenvironment appears to be a prerequisite for effective ICI treatment, and a reduced T-cell infiltrate suggests that the expected response in MMRd ECs would be less to this type of treatment. Additionally, *JAK1* LOF mutations have been shown as potential mechanism of PD-L1 blockade resistance in melanoma [381, 382].

4.4.2.Strengths

This piece of work had a number of strengths. Firstly, a comparative analysis of the interplay between the immune microenvironment and MMRd in colorectal and endometrial cancer has not been

conducted before. MMR status has been approved as a pan-cancer predictive marker for immune checkpoint treatment by the FDA in the USA. However, this has been done without a full understanding of whether MMRd has the same impact on tumour immunity in different cancers. This is clearly an important area to clarify. Secondly, the differing impact of MMRd on the immunophenotype in CRCs Vs EC was demonstrated using two data sets, and two different quantification methodologies. In particular, the data available from TCGA has been produced in accordance to rigorous standards. This reduces the possibility that observed difference is a methodological artefact, or the result of batch effect. Thirdly, the use of multispectral IF quantification allowed for localisation of T-cell infiltrates to either the stromal and tumour compartment. As discussed above, this may be particularly important in determining their influence on clinical outcomes.

4.4.3.Limitations

There were some limitations to this study. Due to available methodologies, immune cell quantification in FFPE tumour samples focused on T-cell quantification. As described above, IFN- γ has pleiotropic effects on the adaptive immune response and promotes the activity of many different immune cell types [4]. In particular, the role and influence of lymphocytes such as NK cells and B-cells was not explored. It has been suggested that these two immune cell types are important in determining immunotherapy response [383, 384]. Confirming with multispectral IF quantification whether there was similar depletion of other IFN- γ related immune cell populations would have been pertinent. Furthermore, an additional step in the FFPE tissue analysis would have been to include exome sequencing for *JAK1* mutations, and/or IHC for JAK1 protein loss to further analyse the association between this putative immune escape mutation and a diminished immune response. Due to the analysis methodology, it was only possible to analyse snap shots of tumour. I would have liked to examine the immune cell infiltrate across the whole tumour section, with a view to analysing possible heterogeneity of the anti-tumour immune response in MMRd CRCs and ECs. Particularly, if this could be correlated with variations in MMR and JAK1 protein expression within tumours.

2933
2934
2935
2936
2937
2938
2939
2940
2941
2942
2943
2944
2945
2946
2947
2948
2949
2950
2951
2952
2953
2954
2955

Additionally, whilst cancer cell line models provided a useful tool for exploring the effect of *JAK1* mutations on IFN- γ , they are an oversimplification of complex biological processes. The use of *in vivo* animal models will provide a robust, functional, model to explore the interplay between genomic instability, IFN- γ signalling and T cell immune response.

4.4.4.Future Work

Future work should focus to address key questions raised in this study. In the first instance, expanding the *in situ* tissue analysis to include more elements of the adaptive immune system will be important. This should include not only quantification of other immune cell types, but also analysis of the spatial interact between different populations. Additionally, further efforts should be directed to understanding why the immune response differs between CRC and EC. This study has highlighted the possible contribution from a difference in tumours total mutational burden. This difference could not be explained by differences in MLH1 loss, and it should be considered whether additional mutagenic processes play a role in CRC and EC MMRd cancers. As well as differences in the overall number of mutations, MMRd results in the selective enrichment of different particular gene mutations. This study demonstrated an enrichment for *JAK1* LOF mutations in MMRd ECs. Understanding why MMR loss results in the predilection for particular gene mutations will help further our knowledge of the biology of the disease, and illuminate potential therapeutic targets. And finally, work is already underway to develop a mouse model to study MMRd in colorectal and endometrial tumorigenesis, and to further examine the role of immune escape mutations such as *B2M* and *JAK1*, data from which are presented in Chapter 5.

5. Functional analysis of mismatch repair loss in intestinal cancer

5.1. Introduction

As discussed in the Introduction and shown in Chapter 4, MMRd is an important driver of carcinogenesis in the endometrium and colon. In colorectal cancer (CRC), I have shown, in agreement with previous studies [95, 326], that MMRd cancers are associated with reduced recurrence risk and improved overall survival compared to MMRp cancers (Chapter 4). I have presented data which suggest this is at least partly explained by the vigorous immune response MMRd tumours elicit (Chapter 4); particularly increased CD8⁺ cytotoxic T-lymphocyte infiltration, which I have demonstrated is strongly associated with improved outcomes in CRC irrespective of other molecular or genomic variables (Chapter 3). Whilst these data offer important insights, they remain correlative by their nature. Functional studies are required to further our understanding of the interplay between tumour genomics and the immune response. Specific issues to address are why MMRd has discordant effects on the immune microenvironment in colorectal compared with endometrial cancer (EC) (Chapter 4). My data suggests that this may be explained by enrichment for *JAK1* loss of function mutations in MMRd ECs. In contrast, MMRd CRC rarely acquired *JAK1* mutations, but frequently exhibited *B2M* mutations. This raises the question of why immune escape mutations are differentially enriched in cancers with the same underlying genomic instability, but arising at different anatomical sites. More generally, our understanding of how immune escape mechanisms, such as *B2M* loss, shape the immune microenvironment is relatively limited. This may prove to be an important area for future therapeutic targeting.

5.1.1. The mouse as a tool for cancer research

Genetically modified mice (*Mus musculus*) models have been invaluable *in vitro* experimental tools in cancer research, and have contributed to a number of important discoveries [385]. The earliest work studying cancer in mice was driven by the use of xenograft models, which involve the transplantation of animal or human tumour lines into immunocompromised mice [386]. However, xenograft tumours are inferior to autochthonous murine tumours in several important ways [387,

388]. Of particular relevance to my research is the observation that the tumour microenvironment of xenograft cancers lacks a number features that would be normally expected, including the presence of immune cells [387, 388]. These shortcomings have been somewhat overcome by the use of genetically engineered mice. Generally speaking, there are two main strategies of genetic modification employed: loss of function changes, which includes knockout or knockdown models, and gain of function, which includes transgenic or knock in models [385]. The first reported murine cancer models involved were transgenic knock in experiments, which involve the introduction of foreign DNA sequences into the pronuclei of fertilized zygotes [389, 390]. Initial experiments concentrated on the overexpression of viral and cellular oncogenes, which were found to induce carcinogenesis [389, 390]. Whilst the transgenic approach benefits from a simple methodology that allows for rapid generation of experimental mice, it is disadvantaged by inaccurate control over the level and pattern of transgene expression [385]. In contrast, studying tumour suppressor loss requires a more targeted approach [386]. By utilising homologous recombination in mouse embryonic stem cells, it is possible to either delete specific gene coding regions (knockout) or introduce exogenous sequences (knock in) [385]. Seminal early experiments investigating tumour suppressor genes focused on germline knockout of key tumour suppressors such as *Rb* [391], *p53* [392] and *Brca1* [393, 394]. Whilst these mice have provided great insight into familial cancer syndromes (such as *p53* mutant mice and Li-Fraumeni) [392], they are limited by the fact that it is not possible to control the timing of the mutation, nor tissue location [385]. The latter can prove particularly problematic as many germline tumour suppressor mutations are embryonically lethal [393, 394]. One potential path around this is the use of conditional or inducible systems. Primarily, these are the Cre-*loxP* and the flippase (FLP) -*flippase recognition target* (*FRT*) systems [395]. Of the two, the Cre-*loxP* system is the more commonly used. The enzyme Cre recombinase undertakes targeted DNA recombination between two *loxP* sequences [396]. Whether this recombination results in excision, inversion or translocation is dependent of the relative orientation of the *loxP* sites [396]. It is possible to achieve tissue specific effects by introducing transgenes with *Cre* expression under the control of promoters known to only be expressed at certain sites [397]. For example in *Villin*-Cre transgenic mice, the enzyme Cre recombinase is only present in intestinal cells that express *Villin* [397]. Inducible Cre systems allow

temporal control of murine genetic alterations [398]. The two most commonly used are tamoxifen inducible and the tetracycline (Tet)-inducible Cre. In tamoxifen inducible system, a mutated oestrogen receptor is joined to Cre-recombinase (Cre-ER^T or Cre-ER^{T2}); binding with tamoxifen (4-hydroxytamoxifen) results in activation of the Cre recombinase enzyme [399, 400]. Thus, the use of a conditional, inducible Cre-system allows for temporal and spatial control over genetic alterations in mice, which makes it possible to more closely model human cancer development [385].

5.1.2. Mouse models of intestinal cancer

Murine models have led to a number of key discoveries in cancer research. These include the recognition that key developmental pathways, such as Wnt/B-catenin or sonic hedgehog, can be aberrantly expressed in tumours and are central to carcinogenic processes [401-403]. The demonstration that multiple genetic alterations are required for tumour development (for example, in mice ovarian cancer models *p53* loss must be accompanied by disruption of a further two pathways to cause tumour formation) [404]. Or that specific gene combinations cooperate to promote tumorigenesis, such as *Brcal* and *p53* loss accompanied by *c-myc* activation drives malignant transformation of mice ovarian cells [405]. Correlatively, *BRCA1* mutant human tumours typically also have *p53* mutations and *c-myc* amplification [406].

Cancer formation in the gastrointestinal tract has also been the subject of considerable functional research in the mouse; selected key models are summarised in Table 5.1. The first murine model developed to investigate intestinal tumour specifically was the *Apc*^{Min}. This model carries a truncating mutation, discovered in mice exposed to the mutagen ethylnitrosourea, at codon 850 of the *Apc* gene [407]. The term Min (multiple intestinal neoplasia) is a reflection of the phenotype - heterozygote *Apc*^{Min} mice typically develop around 30 small intestinal polyps and die from anaemia by 120 days age. Further intestinal mice models were developed as variations of the original Min model. For instance, the *Apc*Δ⁷¹⁶ mice carry a truncating mutation at the 716 codon, and this results in an approximately 10 fold increase in intestinal polyp frequency [408]. In contrast, *Apc*^{I638N} mice

3038 harbouring a chain-terminating mutation in the exon 15 have substantially reduced polyp formation
3039 [409]. A model with greater specificity for colonic tumorigenesis was generated by combining *Apc*^{Min}
3040 with heterozygote deletion of *Cdx2* [410]. This results in colonic polyposis that more closely
3041 resembled human familial adenomatous polyposis (FAP) disease [410]. Mutagen exposure is another
3042 way inducing murine intestinal carcinogenesis on its own [411], or can exaggerate the effects of
3043 genetic modifications such as *Apc*^{Min} [412, 413]. Commonly used is azoxymethane (AOM), which is
3044 metabolised to methylazoxyformaldehyde, a compound capable of alkylating the DNA bases guanine
3045 and thymine [414]. The advantage of using an AOM-based mouse model is the reliability in
3046 mimicking human CRC morphological features [415]. And in fact, the addition of AOM treatment to
3047 *Apc*^{Min} mice increases colonic, as opposed to small intestine, polyp formation [412, 413].

Table 5-1. Summary of the key murine models for the study of intestinal carcinogenesis. APC: adenomatous polyposis coli. AOM: azoxymethane. Cdx2: caudal type homeobox 2

Model description	Phenotype	Ref.
<i>Apc</i> ^{Min/850} model. Heterozygous <i>Apc</i> stop mutation at codon 850	Multiple adenomas and a few invasive carcinomas in large and small intestine,	[407]
<i>Apc</i> ^{1638N} : heterozygous <i>Apc</i> chain-terminating mutation in exon 15	Multiple colonic adenomas and adenocarcinomas	[409]
<i>Apc</i> Δ ⁷¹⁶ : heterozygous <i>Apc</i> stop mutation at codon 716	Multiple intestinal polyp formation from 3 rd week of life.	[408]
<i>Apc</i> Δ ¹⁴ : <i>Apc</i> inactivation by Cre-loxP deletion of exon 14	Multiple polyp formation (predominantly in the colon; more so than <i>Apc</i> ^{Min} mice)	[416]
<i>Apc</i> Δ ²⁴² : gene trap cassette between exons 7 and 8 - results in truncated <i>Apc</i> product with no armadillo repeat domain	Increased intestinal adenomas of smaller size compared to <i>Apc</i> ^{Min} mice,	[417]
<i>Apc</i> ^{Min/+} <i>p53</i> ^{-/-} : <i>Apc</i> Min model combined with complete p53 knockout	No increase in adenoma formation or carcinoma development compared with <i>Apc</i> ^{Min} models alone	[418]
<i>Apc</i> Δ ⁷¹⁶ <i>Cdx2</i> ^{+/-} mice model	Greater concertation of colonic polyps than <i>Apc</i> ^{Min/+}	[410]
Wild type mice exposed to AOM carcinogen for 6 weeks	Multiple colonic adenomas, but at 24 week no evidence of carcinomas	[411]
<i>Apc</i> ^{Min/+} mice exposed to AOM	AOM increases intestinal adenoma formation 1.5-6 fold. Furthermore, resulted in a greater number of colonic polyps compared with untreated <i>Apc</i> ^{Min/+}	[412, 413]

5.1.3. Modelling tumour-immune interaction

In addition to delineating the role of genetic alterations in carcinogenesis, mouse models have also provided highly useful insights into the immune response and cancer. Perhaps the most important of these was the conclusive demonstration that the adaptive immune response plays an essential role in tumour suppression [107]. Seminal early work demonstrated that immunodeficient mice, with either absent IFN- γ signalling or complete lymphocyte depletion, spontaneously developed tumours [419]. These initial results were confirmed in subsequent publications; a selection of which are summarised in Table 5.2.

Table 5-2. Summary of the key murine models for the study of the immune response and cancer development. *Ifng*: Interferon gamma. *Pfp*: perforin. *TRAIL*: TNF- α related apoptosis inducing ligand. *B2m*: beta-2 microglobulin. *Stat1*: signal transducer and activator of transcription 1

Genetic modification	Immune defect	Description	Ref.
<i>Rag2</i> ^{-/-}	Inability to produce mature T or B cells	Spontaneous development of intestinal adenomas, and lung and colonic adenocarcinomas.	[419]
<i>Rag2</i> ^{-/-} <i>Stat1</i> ^{-/-}	Inability to produce mature T or B cells. <i>Stat1</i> loss results in resistance to IFN- α , β and γ	Increased carcinogenesis compared with <i>Rag2</i> ^{-/-}	[419]
<i>Ifng</i> ^{-/-}	IFN- γ deficient	Spontaneous development of lymphoma and lung adenocarcinomas	[420, 421]
<i>Pfp</i> ^{-/-}	Lacks perforin: a cytolytic molecule crucial for the activity of NK and T-cells	Developed spontaneous B and T cell lymphomas. Also, more susceptible to lymphomas transplanted from <i>p53</i> null mice	[420, 421]
<i>Trail</i> ^{-/-}	Loss of NK, B-cell and T-cell driven apoptosis via the TRAIL signalling pathway	Spontaneous development of lymphoid malignancies	[422]
<i>Pfp</i> ^{-/-} <i>Ifng</i> ^{-/-}	IFN- γ and perforin deficient	Accelerated lymphoma formation compared to <i>Pfp</i> ^{-/-} mice alone	[420, 421]
<i>Pfp</i> ^{-/-} <i>B2m</i> ^{-/-}	Lack perforin and <i>B2m</i> , an essential component of the MHC class I pathway	Accelerated lymphoma formation compared to <i>Pfp</i> ^{-/-} mice alone	[423]

In addition to the revelation that immunodeficiency enables spontaneous tumour formation, it was also been shown that immune function loss exacerbates carcinogenesis resulting from other genetic alterations (select studies are described in Table 5.3). Of particular interest are the studies reporting abrogation of cytotoxic immunity (*Pfp* knockout) or IFN- γ signalling (*Ifng* or *Stat1* knockout) as these are directly relevant to the results reported in Chapters 3 and 4 respectively.

Table.5-3. Summary of the key murine models for the study of the immune deficiency and cancer genetic modification. *Ifng*: Interferon gamma. *Pfp*: perforin. *Stat1*: signal transducer and activator of transcription 1

Genetic modification	Immune defect	Description	Ref.
<i>Ifng^{1-/-}Trp53^{1-/-}</i>	IFN- γ deficiency combined with <i>p53</i> knockout	IFN- γ loss increases the rate and spectrum of carcinogenesis compared to <i>p53</i> loss alone: mice developed teratomas, haemangiomas and chondrosarcomas	[424]
<i>Stat1^{1-/-}Trp53^{1-/-}</i>	<i>Stat1</i> loss	<i>Stat1</i> loss increases the rate and spectrum of carcinogenesis compared to <i>p53</i> loss alone: mice developed teratomas, haemangiomas and chondrosarcomas	[424]
<i>Pfp^{1-/-}Trp53^{1-/-}</i>	Perforin and <i>p53</i> null mice	Perforin loss accelerated lymphoma formation compared to <i>p53</i> null mice alone	[420]
<i>Pfp^{1-/-}Mlh1^{1-/-}</i>	Perforin and <i>Mlh1</i> null mice	Perforin loss accelerated B-lymphoma development in <i>Mlh1</i> null mice	[425]

Mouse models have also been vital for driving treatment advances, such as immune checkpoint inhibition. The therapeutic potential anti-CTLA4 [426], anti-PD1 [427] and anti-PDL1 [203] antibodies was first demonstrated in murine models.

5.1.4. Mouse models of mismatch repair deficiency and cancer

Characterisation of the genetic alterations that lead to MMRd in human cancers has enabled the development of murine models [428]. Similar to other eukaryotic organisms, mismatch repair in mice is carried out by the MutS (Msh2 and Msh6) and MutL complexes (Mlh1 and Pms2). Mutations resulting in loss of functions of these genes can result in microsatellite instability (MSI). Table 5.4

summaries some of the previous publications that have reported mouse lines harbouring genetic defects emulating those found in human disease. The spectrum of cancers, and hence the validity of the model, depends on which genetic alteration is introduced and its organ specificity. For instance, systemic homozygous knockout of *Msh2* or *Pms2* does not result in intestinal or endometrial cancers that are typical of MMR deficiency in humans [429]. Instead these mice develop lymphoma, which is not a malignancy associated with Lynch Syndrome [284]. Additionally, the severity of the phenotype meant that mice commonly have to be culled at early time points for humane reasons, at which point a GI phenotype may not be evident. In contrast, *Mlh1* and *Msh6* knockout mice do exhibit GI tract tumorigenesis [430]. In light of these shortcomings, other groups have sought to use conditional genetic modification to achieve a intestinal specific phenotype; these have included constitutive *Villin-Cre* or inducible *Lrg5-ERT2-Cre* [431, 432]. Mouse models have also been developed which perturb MMR genes less frequently involved in human disease. *Msh3* knockout alone results in a mild phenotype [433], whilst *Pms1* knockout mice were indistinguishable from wild type controls [430]).

A key feature of human MMRd features is a vigorous anti-tumour immune response. To date, there has been little work to characterise the immune microenvironment of tumours arising in MMRd mouse models. This is somewhat surprising, given the importance of the immune response in defining tumour behaviour and development. Moreover, one of the biggest areas that requires addressing is the role of immune escape mutations. To the best of my knowledge, there has been no functional work to investigate the role of the immune escape mechanisms in shaping the tumour microenvironment of MMRd.

Functional analysis of mismatch repair loss in intestinal cancer

Genetic alteration	Median survival (months)	MSI detected	Tumour type	Male fertility affected	Immune phenotyping	Ref.
Non organ specific MMRd models						
Homozygous Msh2 knockout (<i>Msh2^{-/-}</i>)	6	+	Lymphoma	-	-	[429]
Homozygous Msh2 knockout (<i>Msh2^{-/-}</i>)	6	+	Lymphoma	-	-	[434]
Homozygous Msh2 knockout (<i>Msh2^{-/-}</i>) + <i>Apc^{+/-}/Min</i>	2-3	NR	Multiple intestinal tumour formation	-	-	[435]
Homozygous Msh6 knockout (<i>Msh6^{-/-}</i>)	6	-	Lymphoma; intestinal tract (predominantly small bowel; adenoma and adenocarcinomas), liver, lung, skin and endothelial tumours	-	-	[436]
Homozygous <i>Mlh1</i> knockout	6	+	Intestinal adenomas and adenocarcinomas (predominantly small intestine). Lymphoma and cutaneous malignancies	+	-	[430]
Homozygous <i>Pms2</i> knockout	6	+	Lymphoma and sarcoma	+	-	[430]
Homozygous <i>Pms2</i> knockout	Nr	+	Lymphoma and sarcoma	+	-	[437]
Hetero- and homozygous <i>Mlh1</i> knockout with/without <i>APC^{+/-}</i>	7	+	Intestinal adenomas and adenocarcinomas; lymphomas. Phenotype exacerbated by combination with <i>APC^{+/-}</i>	+	-	[99]
Homozygous <i>Mlh1</i> missense mutation: <i>MLH1^{G67R/G67R}</i>	5	+	Intestinal adenomas and adenocarcinomas; lymphomas. Less tumour formation compared with <i>MLH1^{-/-}</i>	+	-	[438]
<i>Pms2</i> endonuclease deficient: <i>Pms2^{EK/EK}</i>	12	+	Lymphoma	+	-	[439]
Missense <i>Msh6</i> mutation: <i>Msh6^{TD/TD}</i>	12	+	Lymphoma; intestinal adenomas	-	-	[440]
<i>Msh2</i> point mutation (<i>Msh2^{G674A/G674A}</i>)	9	+	Lymphoma formation; infrequent intestinal adenocarcinomas (19% of mice)	+	-	[441]
<i>MLH1^{-/-}/Pms2^{-/-}</i> Vs <i>MLH1^{-/-}</i> or <i>Pms2^{-/-}</i>	NR	+	<i>MLH1^{-/-}/Pms2^{-/-}</i> similar mutation phenotype to <i>MLH1^{-/-}</i>	+	-	[442]

3104 **Cont'd**

3105

Genetic alteration	Median survival (months)	MSI detected	Tumour type	Male fertility affected	Immune phenotyping	Ref.
Intestine specific MMRd models						
<i>Vil-Cre Msh2^{LoxP/null}</i> (homozygous knockout) or <i>Vil-Cre Msh2^{LoxP/G674D}</i> (missense mutation knock in)	12	+	Intestinal adenomas and adenocarcinomas (predominantly small intestine). <i>Msh2^{LoxP/G674D}</i> genotype resulted in higher tumour rate than <i>Msh2^{LoxP/null}</i>	NA	-	[431]
Inducible <i>Msh2</i> loss: <i>Lgr5-CreERT2-Msh2^{flow/-}</i>	19	-	Predominantly intestinal adenoma and adenocarcinoma formation. Occasional extra-intestinal neoplasms: lung, liver and testis.	NA	-	[432]

Table 5-4. **Summary of selected studies that have investigated mismatch repair loss in mice**

5.2. Rationale for study and aims of chapter

The use of genetically modified mice models has greatly increased our understanding of the mechanisms that drive cancer in humans. There have been a number of published mouse models that have attempted to recapitulate MMRd-driven carcinogenesis. However, as noted above, there are a lack of studies describing the immune microenvironment in these models, or how it interacts with the underlying genomic instability.

My aim was to develop a mouse model for MMRd intestinal carcinogenesis. Specifically, this would be done using intestinal specific, inducible loss of *Mlh1* driven by Cre-loxP recombination. *Mlh1* loss was chosen as the focus because as shown in Chapter 4 it is the most common cause of sporadic MMRd in CRC. The results from the correlative analyses presented in Chapter 4 raised several hypotheses I wanted to investigate. To date, there has been little immunophenotyping of murine MMRd tumours reported. I wanted to investigate with *Mlh1* loss resulted in similar vigorous immune response, particularly increased T-cell infiltrate, that is observed in human cancers. Furthermore, parallel work is investigating *Mlh1* in the endometrium. This will allow for exploration of the findings presented in Chapter 4, and address the question of why MMRd has different effects in CRC versus EC. Additionally, mechanisms of immune escape remain poorly understood. I wanted to investigate the contribution of these in shaping the immune microenvironment of MMRd CRCs.

5.3. Materials and methods

5.3.1. Murine housing and breeding

Animals were housed in a conventional facility with 12 hours of light and dark cycles. Normal chow and water were provided *ad-libitum*. The genotype of adult animals intended for experimental use was determined following ear clipping, which took place at around post-natal day 10. All mouse work was conducted under Project licence no. PDF0B94C3 and Personal licence: 12D1D1096. The breeding schemes employed are summarised in Figure 5.1

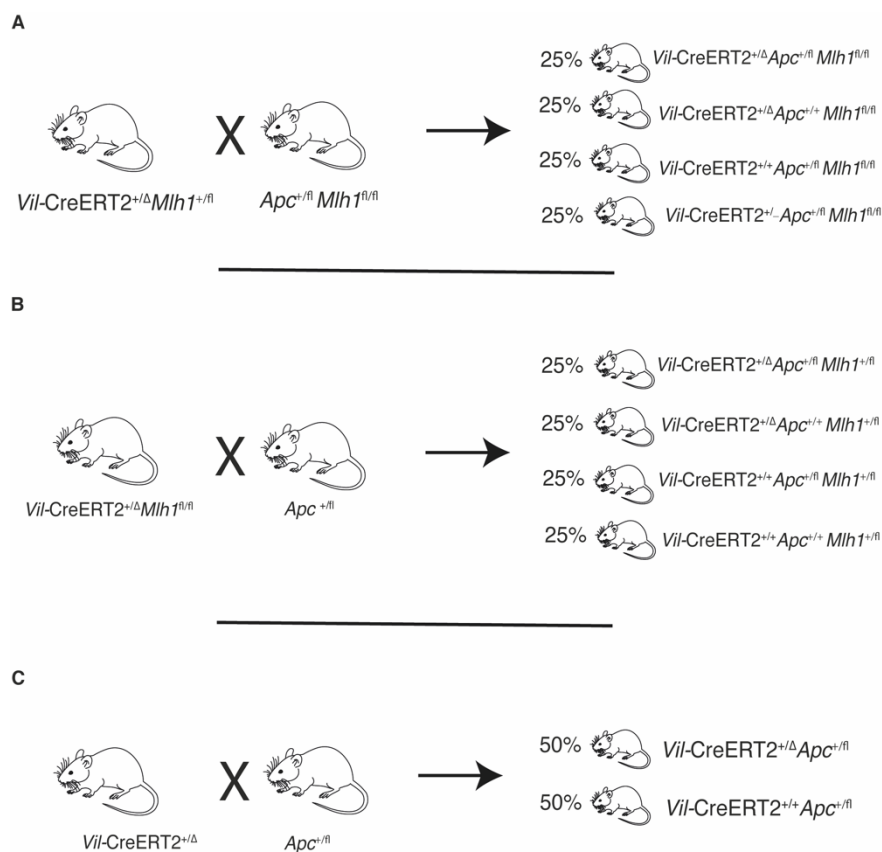


Figure 5-1. Outlines the breeding plan used to obtain experimental mice for this work. A. breeding to achieve $Vil-CreERT2^{+/Δ} Apc^{+/fl} Mlh1^{fl/fl}$. B. breeding to achieve $Vil-CreERT2^{+/Δ} Apc^{+/fl} Mlh1^{+/fl}$. C. breeding to achieve $Vil-CreERT2^{+/Δ} Apc^{+/fl} Mlh1^{+/Δ}$

5.3.1.1. Tamoxifen (4-hydroxytamoxifen) induction of Vil-Cre recombinase

Intraperitoneal (IP) tamoxifen induction to achieve *Vil-CreERT2* recombination was carried out in accordance with previously published protocols [443]. Administration of 4-hydroxytamoxifen began when mice were 6 weeks of age. The dosing regime was 1mg of tamoxifen (prepared as 10mg/ml

solution dissolved in corn oil) injected into alternating flanks for 5 consecutive days. The mice were monitored for approximately 30 minutes post injection, and then weighed daily for the following five days to monitor for ill effects of IP injection.

5.3.2. Genotyping PCR

Primers, conditions and reagents as detailed in Methods, Section 2.5.4.

5.3.3. Animal monitoring, sacrifice, necropsy and tissue preparation

Mice were monitored daily for signs of ill-health, and evidence of distress. The project licence for this work was limited to moderate severity. As such, mice were sacrificed earlier than planned time points if they displayed evidence of ill health; signs for which included, reduced activity, hunching and pale feet. The experimental times were chosen based on our own group's experience with constitutional *Mlh1* loss with *Apc^{min}*; these mice developed a phenotype at the 90-day timepoint. There was no previously reports of inducible intestinal *Mlh1* loss to guide this decision. Experimental and control mice who reached defined 90, 120 and 180 day timepoints were sacrificed by approved Schedule 1 method (cervical dislocation) and subject to necropsy. Gut rolls were prepared in accordance with best practice methodology [444]. The entire gastrointestinal tract from the stomach to the rectum was removed en-bloc, the stomach removed and the remaining intestine washed in PBS. This was then subdivided into 4 sections: proximal, medial and distal small bowel, and colon. The intestinal lumen was flushed with PBS to remove its contents and opened longitudinally. Each intestinal section was dissected onto filter paper and pinned to maintain a flat sheet of tissue. It was then fixed in 10% neutral buffered formalin (10% NBF) for 24 hours before transfer to a 70% ethanol solution before further processing.

5.3.4. Experimental mice

Heterozygous floxed *Apc* (Adenomatous polyposis coli) mice: Mice with heterozygous floxed *Apc* (*Apc^{fl/fl}*) allele bred with *Vil-CreERT²*. The animals have been backcrossed >10 generations

on a Black6 background. Cre recombinase induction by 4-hydroxytamoxifen results in heterozygous *Apc* loss. MGI ID: 88039. Original publication: [445].

Heterozygous/Homozygous floxed *Mlh1* (MutL homolog 1) mice: Mice with heterozygous/homozygous floxed *Mlh1* allele bred with *Vil*-CreERT². The animals have been backcrossed >10 generations on a Black6 background. Cre recombinase induction by 4-hydroxytamoxifen results in heterozygous/homozygous *Mlh1* loss.

Heterozygous *Villin* CreERT2 mice: Mice carrying transgene with a 4-hydroxytamoxifen inducible Cre recombinase under the control of the mouse *Vill* promoter. MGID: 2662274. Original publication: [446]

5.3.5. Quantification of intestinal polyps

Intestinal polyps were measured in their maximum dimension and the number quantified using methylene blue staining and manual assessment using standard light microscopy (Section 2.5.5) . Polyp numbers were determined for each intestinal section, and for all in combination.

5.3.6. Tissue storage

Tissues were stored as detailed in Methods, Section 2.3.1.

5.3.7. Histological analysis

5.3.7.1. General methods

General methods for fixation, tissue processing, and staining as described in Methods, Section 2.3

5.3.8. Immune cell quantification

5.3.8.1. Immunohistochemistry of immune cells

Immunohistochemistry (IHC) methods as described in Methods, Section 2.3.1

5.3.8.2. Immune cell quantification

Immune cell quantification was conducted using QuPath software version 0.3.0. QuPath is a Java based bioimage analysis software that enables cell marker annotation and quantification on whole slide images (<https://qupath.github.io/>) [447]. QuPath was selected for immune quantification in mouse tissue for two reasons: i) it is a free, open source, software and as such, use is not prohibited by the high costs associated with other platforms (for example HALO, Indica Labs) ii) our group has extensive experience in using it, and it is compatible with the current slide scanning technologies employed.

Stained slides were imaged under brightfield microscopy using the Leica Aperio slide scanner, and the scanned images loaded into QuPath projects for each immune cell quantification. Firstly, the images were manually inspected for baseline quality control, to ensure no areas of tissue loss or folding. For positive marker cell quantification regions of interest (ROIs) were identified, which included areas of normal epithelium and intestinal adenomas. Initially, quantification was refined using one sample. After a ROI is defined, automated analysis allows for counting of the total number of cells and the number of marker positive cells. Visual inspection is then used to determine the accuracy, and the threshold values are adjusted to achieve optimal counting results. Once optimised, the same thresholds are then used to count all ROIs in all samples. Results are presented as positive marker cells/tissue area (cells⁺/mm²).

5.4. Results

5.4.1. Gastrointestinal tumorigenesis in *Vil-CreERT2*, *Apc*^{+/*fl*} *Mlh1*^{*fl/fl*} mice

5.4.1.1. Homozgous *Mlh1* deletion increases intestinal polyp formation

Mlh1 loss lineage was not traced in this murine model but it would be of potentially considerable value. As anticipated from previous studies [99, 408, 410, 417], heterozygous deletion of *Apc* in the murine intestine (*Vil-CreERT2*, *Apc*^{+/*fl*} *Mlh1*^{+/*+*} mice; hereafter referred to *Apc* del mice) led to the development of gastrointestinal polyps (Table 5.5), with median of 16 polyps for all time points (IQR=10.5 – 22.0). Representative H&E stains of these polyps are shown in Figure 5.2 upper row. There was evidence of greater polyp frequency with increasing time (median of 12 polyps at 90 days (IQR= 8.5-14.5), 13.5 at 120 days (IQR=10.5-22.0) and 26 at 180 days (IQR=NA). (Table 5.5; Figures 5.3A).

Timepoint (days)		<i>Apc</i> ^{+/<i>fl</i>} <i>Mlh1</i> ^{+/<i>+</i>}	<i>Apc</i> ^{+/<i>fl</i>} <i>Mlh1</i> ^{+/<i>fl</i>}	<i>Apc</i> ^{+/<i>fl</i>} <i>Mlh1</i> ^{<i>fl/fl</i>}	<i>Apc</i> ^{+/<i>+</i>} <i>Mlh1</i> ^{<i>fl/fl</i>}
90	No. of mice	3	3	5	0
	Polyps (IQR)	12 (8.5-14.5)	9 (8.5-14.5)	19 (13.0-27.0)	-
120	No. of mice	7	6	12	0
	Polyps (IQR)	16.0 (10.5-22.0)	19.0 (15.25 – 22.75)	38.5 (27.75 – 58.75)	-
180	No. of mice	1	1	3	7
	Polyps (IQR)	26.0 (-)	11.0 (-)	34.5 (32.5-36.5)	8 (3.0-17.0)
All time points	No. of mice	11	10	20	7
	Polyps (IQR)	16 (10.5 – 22.0)	15.5 (11.75 – 21.5)	34 (20.5-39.75)	8 (3.0-17.0)

Table 5-5. Gastrointestinal polyps in experimental mice. Median number of polyps in the *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null. IQR: Interquartile Range

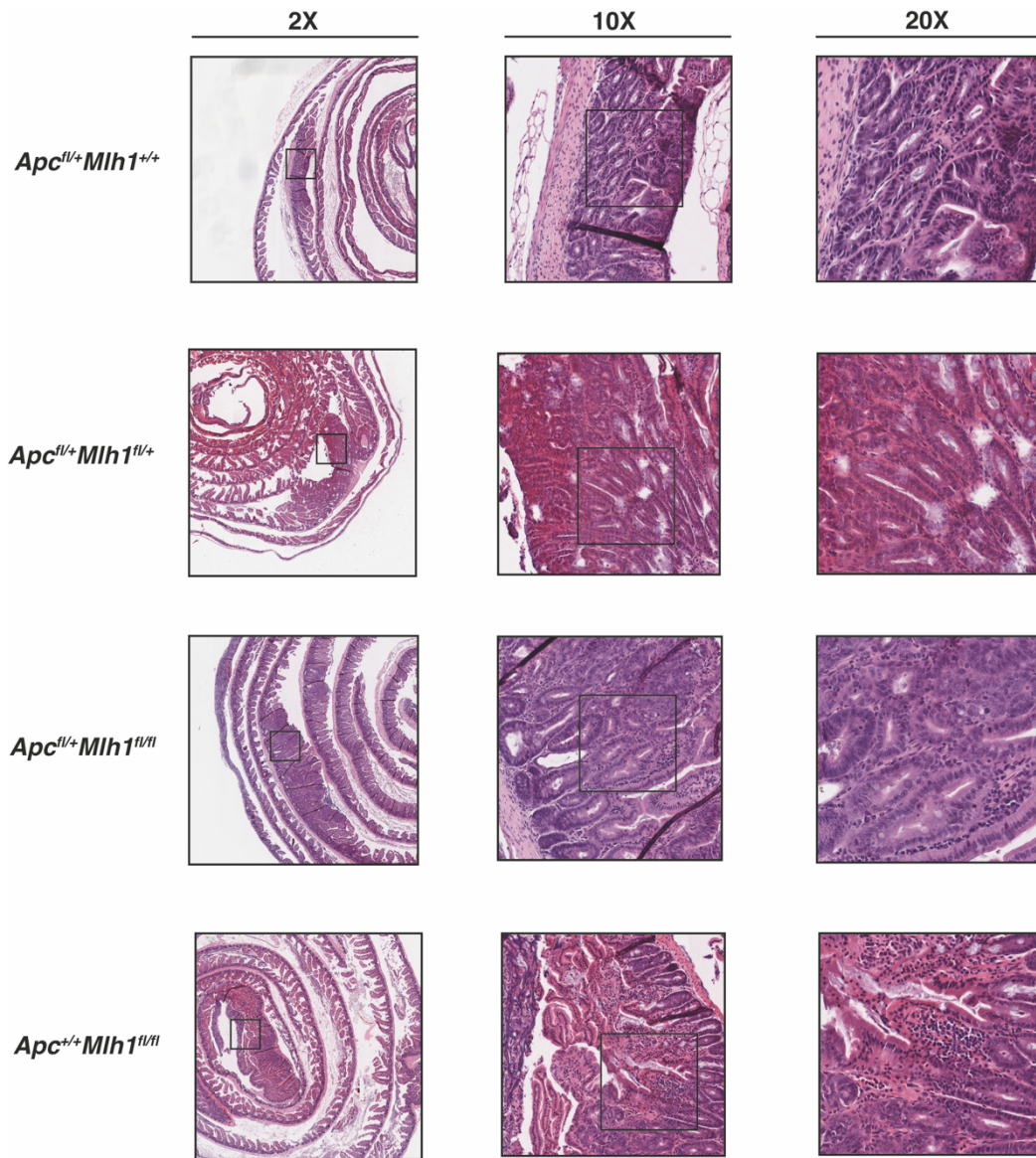


Figure 5-2. **H&E images of representative GI polyps seen in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null mice shown at 2x, 10x and 20x magnification.**

3227

3228 Heterozygous loss of *Mlh1* did not increase the frequency of intestinal polyp formation compared to
 3229 heterozygous *Apc* loss alone. Mice with heterozygous deletion of *Mlh1* and *Apc* (*Vil*-CreERT2, *Apc*<sup>+/^{fl}
 3230 *Mlh1*<sup>+/^{fl} mice; hereafter referred to as *Apc* del/*Mlh1* del mice) displayed no statistically significant
 3231 difference in the number of polyps than *Apc* del mice (median 16.0 IQR(10.5-22.0) versus 15.5
 3232 IQR(11.75-21.5) for all time points, *P*=0.12 by Mann-Whitney test) (Table 5.5; Figures 5.2 & 5.3A).
 3233 In contrast, homozygous loss of *Mlh1* substantially increased the frequency of polyp formation
 3234 compared to *Apc* loss alone. Mice with homozygous deletion of *Mlh1* and heterozygous deletion of
 3235 *Apc* (*Vil*-CreERT2, *Apc*^{+/^{fl} *Mlh1*^{fl/fl} mice; hereafter referred to as *Apc* del/*Mlh1* null mice) had}</sup></sup>

significantly greater number of polyps than *Apc* del mice (median 34 IQR (20.5-39.75) versus 16, $P=0.00016$ Mann-Whitney U test for all time points) (Table 5.5; Figure 5.2 & 5.3A). There was also a greater number of polyps in the *Apc* del/*Mlh1* null mice compared to the *Apc* del/*Mlh1* del mice (median 34 polyps versus 15.5 polyps, $P=2.2e-5$ Mann-Whitney U test) (Figures 5.3A).

Previous work in our group using a constitutive, conditional model for deletion of *Mlh1* and *Apc* min resulted in a statistically significant difference in polyp frequent at 90 days between genotypes of interest. However, in my work the polyp frequency and size at 90 days did not differ significantly between *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null ($P=0.32$ and $P=0.15$ respectively by Mann Whitney U Test) (Table 5.5, Figure 5.3A). Consequently, the mice were aged to a timepoint of 120 days. Prior to this timepoint, several *Apc* del/*Mlh1* null mice began to exhibit signs of illness (hunching, pale feet and reduced activity), and required early sacrifice for humane reasons. Necropsy of these animals revealed substantial burden of intestinal polyps, with several large lesions evident (Figure 5.4). Formal quantification of polyp number in these mice and matched controls confirmed combined homozygous *Apc* del/ *Mlh1* null mice had a substantially and significantly greater number of GI polyps (median 38.5 polyps IQR(27.75 – 58.75)) than *Apc* del (16.0 IQR(10.5-22.0)) and *Apc* del/ *Mlh1* del (19 IQR(15.25-22.75)) ($P=0.0037$ and $P=5.5e-6$ respectively, Mann-Whitney U test) (Figure 5.3A)

3254

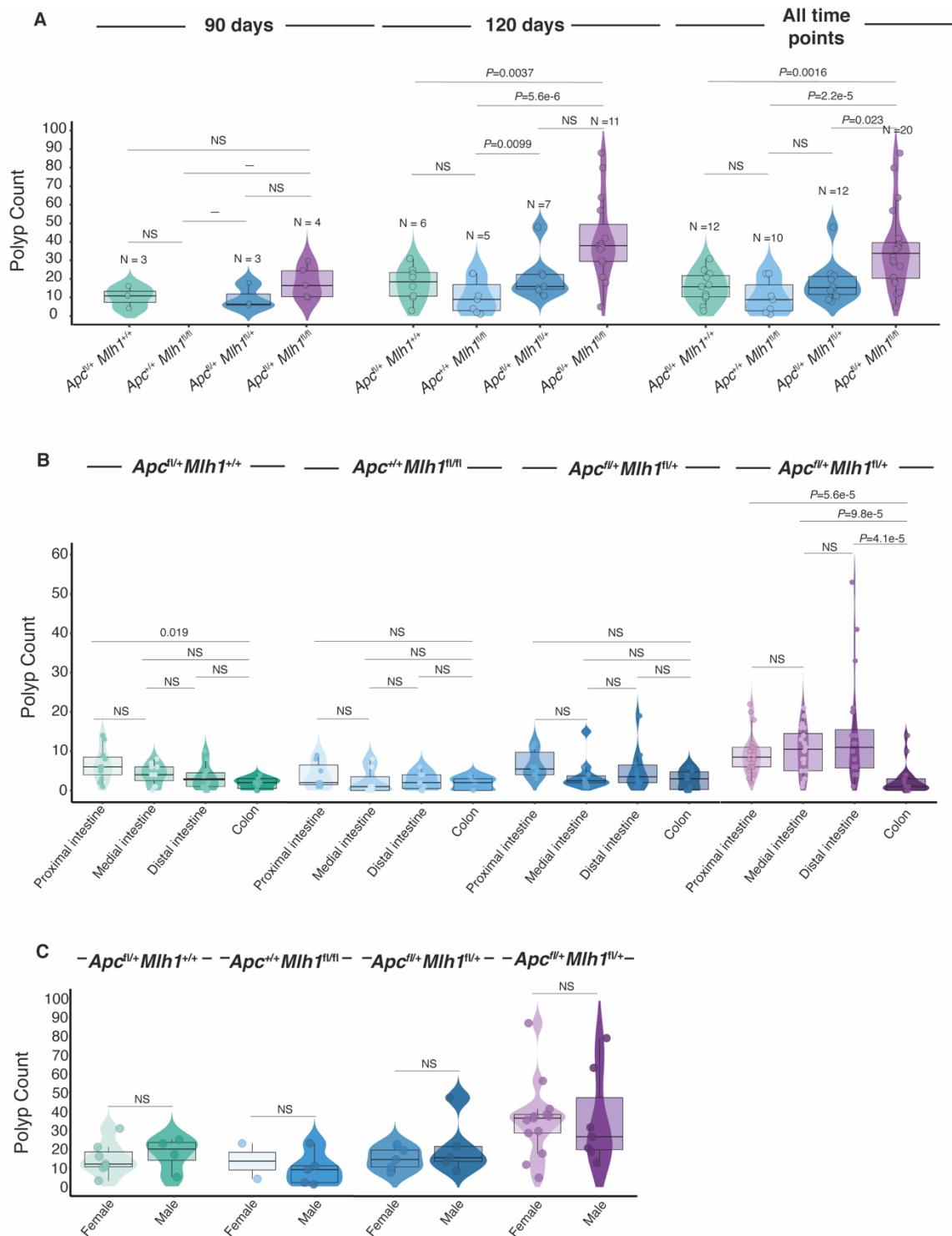


Figure .5-3. **Gastrointestinal (GI) tract polyps** in the *Apc* *del*, *Apc del/Mlh1 del* and *Apc del/Mlh1 null* and *Mlh1 null* **A**. total polyp counts for 90, 120 and all time points for the four genotypes included in this work. **B** GI Polyp formation in the proximal, medial and distal intestine, and colon for each of the four genotypes. **C**. GI polyps in the four genotypes by sex. NS: not significant

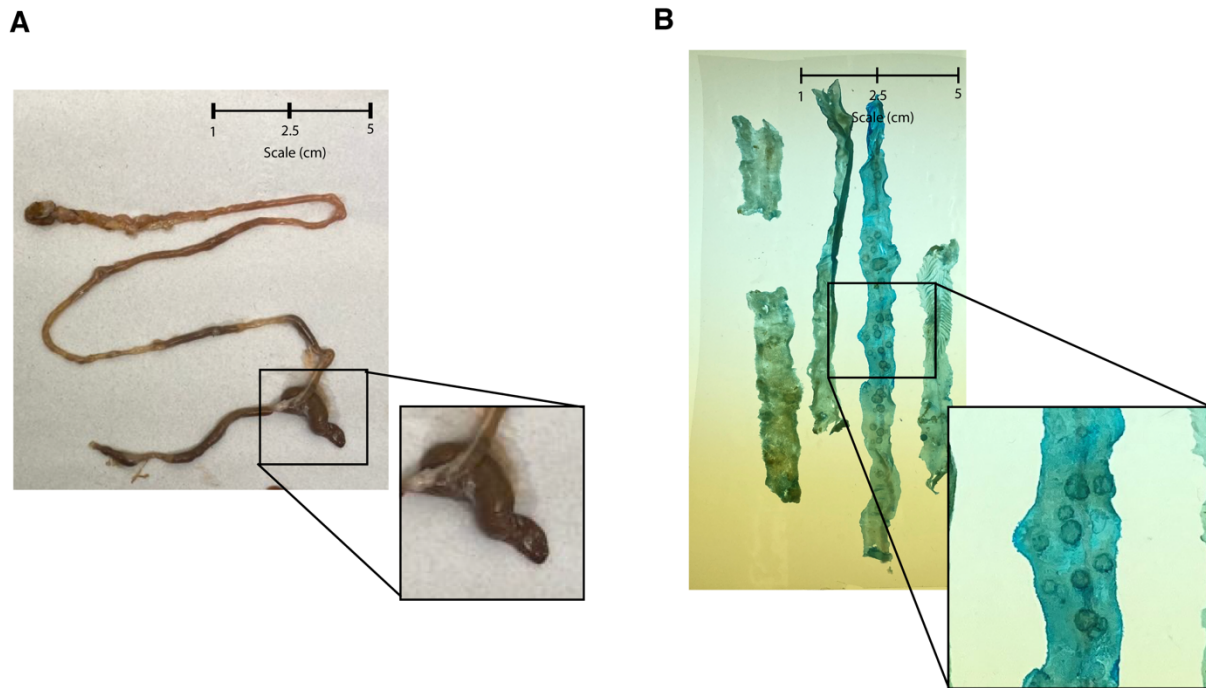


Figure 5-4: **Example of phenotype of *Apc del/Mlh1* null mice at 120 days.** The severity of the phenotype required culling of the mice for humane reasons. A. Macroscopically, the mice had distended and necrotic caeca and appendices B. Evidence of multiple polyps throughout the gastrointestinal tract.

5.4.1.2. Homozygous *Mlh1* loss alone is sufficient to cause intestinal polyps

The results above clearly confirmed the ability of *Mlh1* loss to promote intestinal tumorigenesis. In view of this, I next examined whether *Mlh1* loss alone was sufficient to induce intestinal polyps. Mice were bred and aged out to a timepoint of 180 days after completing 5 days of IP Tamoxifen injections. The mice were sacrificed at this defined timepoint, unless they had to be culled earlier for humane reasons. One mouse with homozygous *Mlh1* deletion alone (*Vil-CreERT2*, *Apc*^{+/+} *Mlh1*^{fl/fl} mice; hereafter referred to as *Mlh1* null mice) was sacrificed at 90 days to see if there was evidence of GI tumour formation. Because there was no evidence, the majority of the remaining mice were aged out to 180 days. At this point, the majority of these displayed few intestinal polyps with a median of 8 polyps (IQR: 3.0-17.0) at 180 days. These polyps occurred throughout the intestinal tract. They were small (no polyp >3mm) and there was no evidence of invasion beyond the mucosa. This analysis shows that *Mlh1* loss alone is sufficient to cause intestinal tumorigenesis in mice, but typically at a slower rate than *Apc del* alone, or *Apc* and *Mlh1* loss combined. Further analyses of these mice are ongoing.

Whilst most homozygous *Mlh1* loss mice had few intestinal polyps compared to *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null mice, one mouse did have to be culled early due to signs of ill health (substantial hunching, pale feet and reduced activity). Necropsy of this mouse revealed a large caecal tumour that had caused intestinal obstruction. H&E staining of FFPE sections showed features consistent with an invasive carcinoma; due to time constraints it was not possible to have this formally reviewed by a pathologist (Figure 5.5).

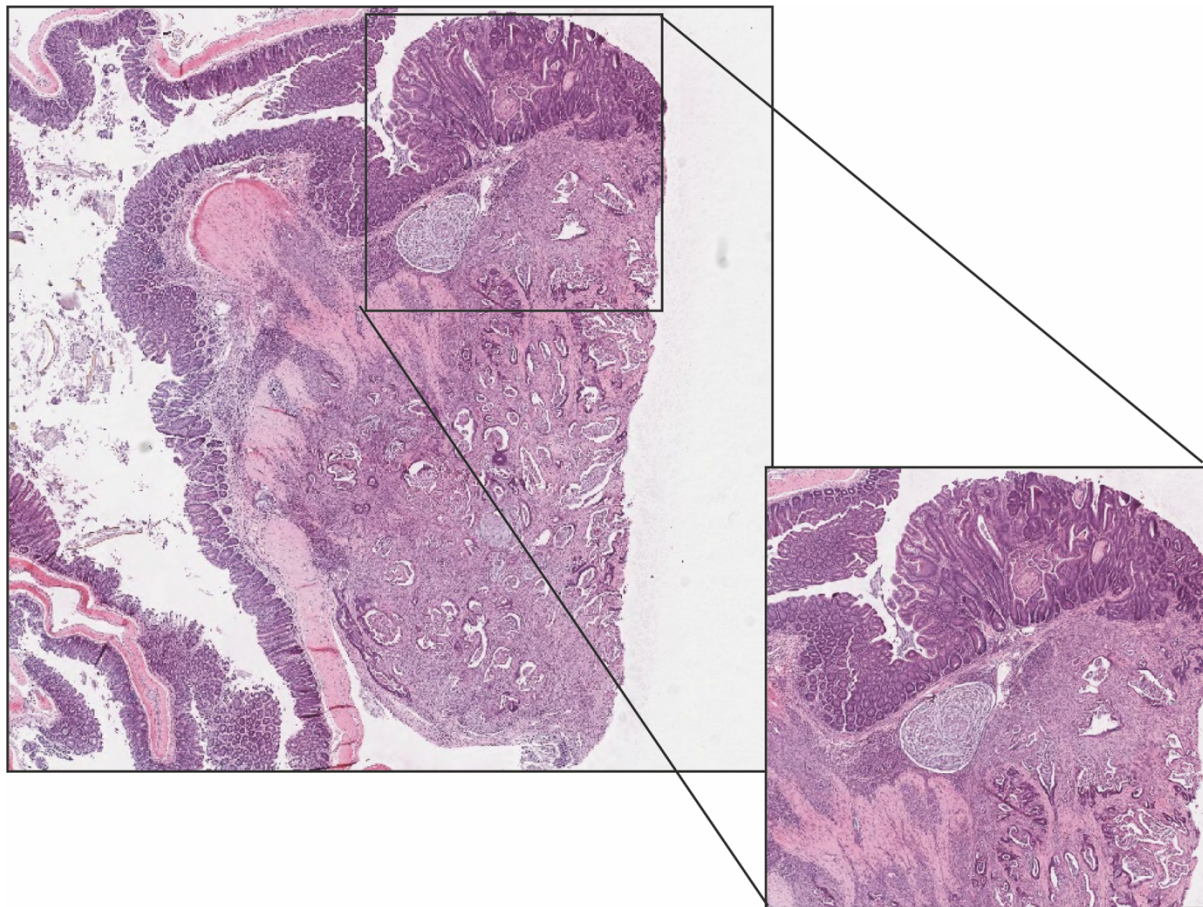


Figure 5-5. H&E staining of an invasive caecal carcinoma in a *Mlh1* null mouse (*Apc* wildtype). This mouse had to be culled early due to severe ill health, and necropsy revealed a large caecal tumour with intestinal obstruction.

5.4.1.3. Histological assessment

Lesion grade was not examined across the different genotypes due to time constraints. Final analyses of the mouse models are ongoing currently, and will include assessment by a pathologist. Assessment

of grade would be an important analysis as it would be expected from human cancers that lesions with *Mlh1* loss would be a higher grade than those without.

5.4.1.4. *Mlh1* and *Apc* loss predominantly results in small intestine tumorigenesis

In humans, somatic MMRd (most commonly secondary to *MLH1* promoter hypermethylation) is typically associated with cancers of the colon, rectum, endometrium and infrequently stomach [304]. However, Lynch syndrome patients carrying *Mlh1* or *Msh2* mutations are also at a marginally increased risk of small bowel malignancies [293]. Previous mice models of non-conditional *Msh2*, *Msh6* or *Pms2* loss have most commonly resulted in lymphoma formation, which has precluded analysis of intestinal tumorigenesis [430, 436]. Non-conditional *Mlh1* knockout or intestinal *Msh2* knockout has been shown to produce tumours primarily in the intestine. Unlike human disease, however, tumour formation has tended to occur in the small intestine [99, 429]. I wanted to investigate the site of polyp formation in the *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null mice. In keeping with these previous studies, analysis of polyp location in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null showed they were predominantly in the small intestine, with colonic polyps infrequently observed (Figure 5.3B, Table 5.6). Normalising the number of polyps for the length of GI tract showed that there was a significantly greater number of small intestine polyps than large intestine polyps in the *Apc* del (0.42 polyps/cm versus 0.26, $P=0.03$) and *Apc* del/*Mlh1* null (1.1 polyps/cm versus 0.38, $P=0.002$); there was no significant difference in the *Apc* del/*Mlh1* del mice (0.45 polyps/cm versus 0.38, $P=0.062$) (all P values calculated by Mann-Whitney U Test).

	Timepoint (days)	<i>Apc</i> ^{+/-} <i>Mlh1</i> ^{+/-}	<i>Apc</i> ^{+/-} <i>Mlh1</i> ^{+/-}	<i>Apc</i> ^{+/-} <i>Mlh1</i> ^{fl/fl}	<i>Apc</i> ^{+/-} <i>Mlh1</i> ^{fl/fl}
Proximal	90	4.0 (3.0-5.5)	6.0 (3.5-8.0)	5.0 (4.0-6.0)	-
	120	6.0 (4.0-8.5)	7.0 (5.3 – 9.8)	12.0 (8.0-13.5)	-
	180	13.0 (-)	4.0 (-)	10.0 (8.0-10.5)	2.0 (1.5-6.5)
	All time points	6.0 (4.0-8.5)	5.5 (4.0-9.75)	8.5 (6.0-11.0)	2.0 (1.5-6.5)
Medial	90	4.0 (2.5 – 5.0)	4.0 (2.5-4.5)	5.0 (3.0-8.0)	-
	120	4.0 (2.5 -6.5)	2.5 (1.3 -3.8)	12.0 (6.5-16.3)	-
	180	4 (-)	2.0 (-)	11.0 (12.0-15.0)	1.0 (0.0-3.5)
	All time points	4.0 (2.5-6.0)	2.5 (0.0-3.5)	10.5 (5.0-14.5)	1.0 (0.0-3.5)
Distal	90	2.0 (1.5 – 2.5)	2.0 (1.5-2.5)	10.0 (5.0-14.0)	-
	120	3.0 (0.5-4.5)	5.5 (2.5-8.5)	12.0 (5.5 -24.0)	-
	180	10.0 (-)	5.0 (-)	12 (10.0-12.0)	2.0 (0.5-4.0)
	All time points	3.0 (1.0-4.5)	3.5 (2.0-6.5)	11.0 (5.8-15.5)	2.0 (0.50-4.0)
Colon	90	2.0 (1.5-2.0)	1 (0.0-1.5)	0.5 (0.0-1.0)	-
	120	2.0 (0.0-3.0)	4.5 (3.3-5.8)	2.0 (0.8-4.0)	-
	180	4.0 (-)	0.0 (0)	1.0 (0.0-1.5)	2 (0.0-3.0)
	All time points	2.0 (0.5-3.0)	1.0 (0.0-3.0)	2.0 (0.0-3.0)	2.0 (0.0-3.0)

Table 5-6. **Site of gastrointestinal polyp formation.** Median number of polyps in the *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null according to gastrointestinal section

5.4.1.5. Sex does not influence GI tumorigenesis in *Mlh1* or *Apc* loss mice

In humans, male LS suffers are at increased risk of cancer than females afflicted with the disease [293]. Additionally, previous studies have shown that deletion of MMR genes has different consequences in terms of fertility and severity of phenotype in mice according to sex. It can cause sterility in males but not in females [430]. I wanted to examine whether the intestinal tumorigenesis resulting from *Mlh1* loss varied according to sex. There was no significant difference in intestinal

tumour formation in male compared with female mice in any genotype at any timepoint, or when all timepoints were combined (Figure 5.3C).

5.4.1.6. Heterozygous *Apc* loss and homozygous *Mlh1* deletion is significantly associated with increased polyp formation independent of other variables

I wanted to analyse the interaction between the variables described above, and the number of intestinal polyps observed. To do so, I performed multiple linear regression analysis using the covariables: timepoint (continuous variable), sex (categorical variable: female (referent) and male), *Apc* mutation status (ordinal variable: wildtype (referent) and heterozygous loss), and *Mlh1* mutation status (ordinal variable: wildtype (referent), heterozygous and homozygous loss). The results of the regression model showed that only variables that significantly predicted increased numbers of intestinal adenomas were *Apc* heterozygous loss (Effect 37.63, $P=0.0058$, Table 5.7) and *Mlh1* homozygous loss (Effect 37.63, $P=0.0058$, Table 5.7). These results suggest that genetic alterations, and in particular *Apc* mutations, are the main determinant of intestinal tumour frequency.

Table 5-7. Multiple linear regression analysis of the relationship between total polyp number and mouse sex, timepoint and genotype

	Effect	<i>P</i>
Timepoint (days)	0.19	0.13
Sex: female (referent)	0.00	-
Sex: male	4.46	0.43
<i>Apc</i> WT (referent)	0.0	-
<i>Apc</i> del	37.63	0.0058
<i>Mlh1</i> WT (referent)	0.0	-
<i>Mlh1</i> del	2.43	0.74
<i>Mlh1</i> null	21.63	0.0035

5.4.2. The immune microenvironment

As shown in Chapters 3 and 4, MMRd in CRC results in an enhanced immune response. Examination of the immune microenvironment in human MMRd CRC revealed increased numbers of tumour infiltrating immune populations, including CD8⁺ and CD4⁺ T cells and macrophages. The first point I wanted to investigate in my mouse model was whether *Mlh1* loss had a similar impact on immune cell infiltrate in intestinal polyps. In addition, I wanted to further investigate the role of immune escape mechanisms in determining the composition and intensity of the anti-tumour immune response. A further motivation for this work is that it would be compared to additional mouse models, developed in parallel, examining *Mlh1* loss in the endometrium. Given the findings presented in Chapter 4, it would be pertinent to compare the immunophenotypes between the two cancer types. Unfortunately, due to disruption caused by the COVID19 pandemic, this was not possible.

5.4.2.1. CD8⁺ T lymphocyte infiltration in normal mucosa and intestinal polyps

As shown in Chapter 4, human MMRd CRCs display increased density of CD8⁺ cytotoxic T-cells compared to MMRp CRCs (Figure 4.8A). However, focusing solely on the cancer specimen precludes comparison with the normal mucosa, which may be highly informative. The intestine is an immune-rich organ, and in which the intensity and type of immune infiltrate could plausibly vary as a consequence of genetic manipulation, even in the absence of neoplasia. I therefore began by quantifying the background level of lymphocyte infiltration in normal mucosa in my models. The aims of this were two-fold (i) to serve as a comparator for the intensity/type of infiltrate in polyps; (ii) to assess whether field cancerisation could lead to widespread lymphocyte infiltrate in histologically normally epithelium. Preliminary visualisation revealed that CD8⁺ cells densely infiltrated the normal mucosa of Black 6 wildtype (WT) mice and *Vil*-CreERT2⁺ mice (Figure 5.6). Similarly, examination of the normal mucosa of the four experimental genotypes (*Apc* del, *Apc* del/*Mlh1* del, *Apc* del/*Mlh1* null and *Mlh1* null) showed dense and widespread CD8⁺ cell infiltrate (Figure 5.9).

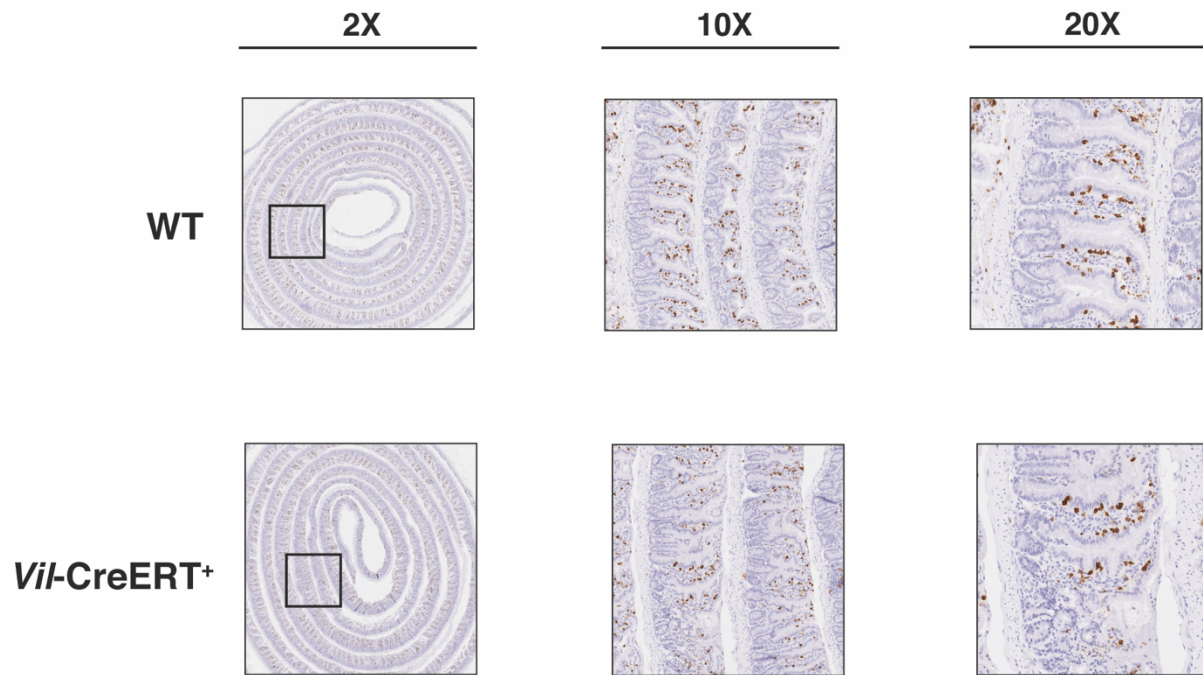


Figure 5-6. Examples of IHC from CD8⁺ cells in normal intestinal mucosa of Black-6 WT and Vil-CreERT² mice shown at 2x, 10x and 20x magnification. CD8⁺ cell infiltrate was widespread across normal mucosal tissue

3357
 3358 I wanted to formally quantify CD8⁺ cell infiltrate in both the normal epithelial tissue and in the polyp
 3359 mucosa. This was done using the image analysis software QuPath (Section 5.3.5, Figure 5.7). There
 3360 was no significant difference in the CD8⁺ cell density in the normal mucosa between *Apc* del (median
 3361 731.3 IQR (333.8 – 856.2) CD8⁺ cells/mm²), *Apc* del/*Mlh1* del (639.5 (591.7 – 737.5) CD8⁺ cells/mm²)
 3362 and *Apc* del/*Mlh1* null (743.91 (266.4 – 1950.9) CD8⁺ cells/mm²) (Table 5.8, Figure 5.8A). However,
 3363 *Mlh1* null mice did have significantly lower levels of CD8⁺ cell infiltrate in normal mucosa compared
 3364 to *Apc* del/*Mlh* del and *Apc* del/*Mlh* del *Mlh1* null mice (188.7 (154.1–412.2) CD8⁺ cells/mm²;
 3365 $P=0.015$ and $P=0.018$ respectively; calculated by Mann Whitney U test).

3366

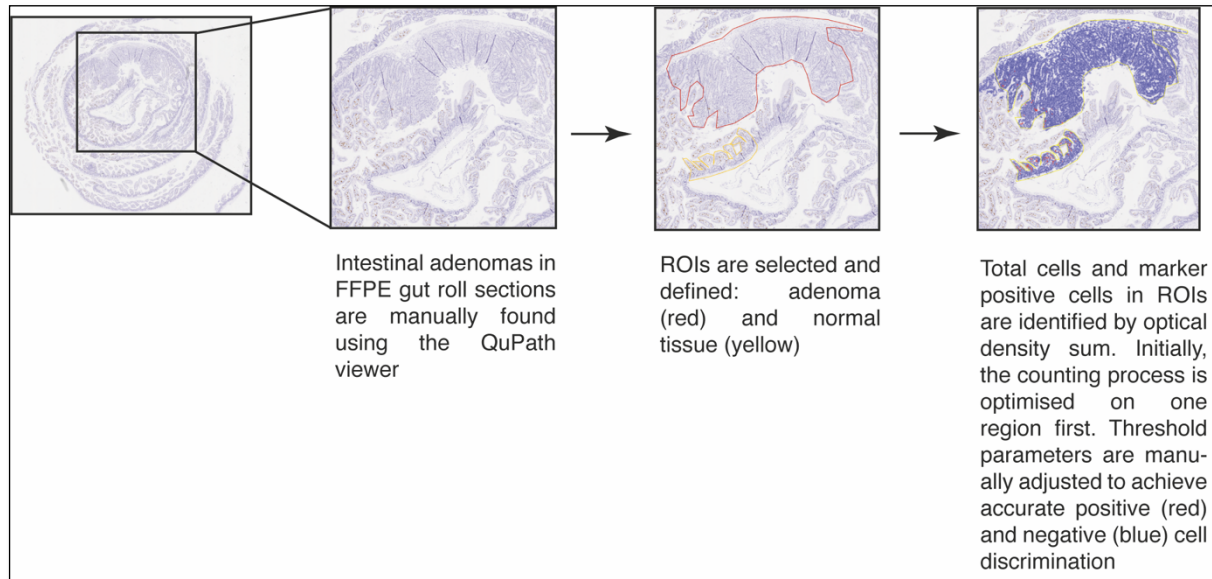
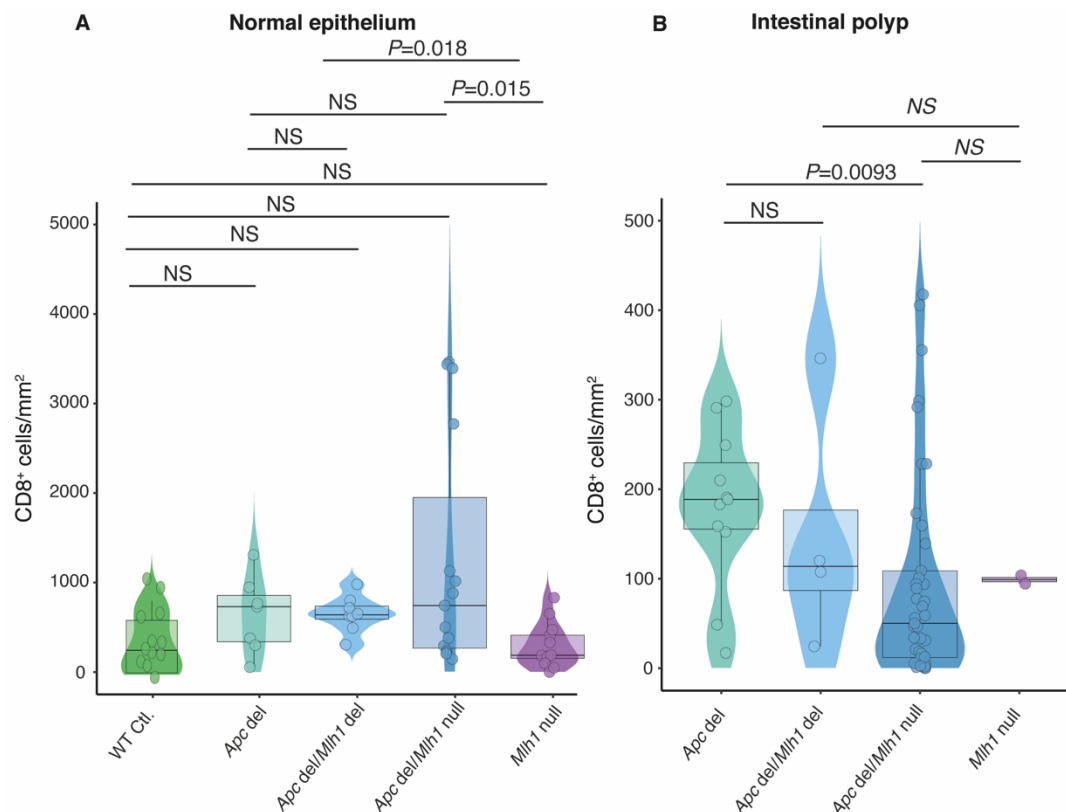


Figure 5-7. Quantification of IHC CD8⁺ positive cells using the QuPath analysis software platform. Adenomas are manually identified in FFPE gut roll sections. ROIs for quantification are annotated as appropriate (adenoma – red; normal – yellow). Quantification is then conducted: marker positive and negative cells are identified. Initially, the quantification is refined by adjusting thresholds before quantification of all samples. ROI: region of interest

3367

3368



3369

Figure 5-8. CD8⁺ quantified by IHC in A) normal intestinal epithelium and B) intestinal polyps in Apc del, Apc del/Mlh1 del, Apc del/Mlh1 null and Mlh1 null mice

Table 5-8 **CD8⁺ cell infiltrate** in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null normal intestine and polyps. Median values of CD8⁺ cells/mm² presented with IQR in parentheses

	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{+/+}	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{+/fl}	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{fl/fl}	<i>Apc</i> ^{+/+} <i>Mlh1</i> ^{fl/fl}
No. of mice (No. of polyps)	4 (18)	2 (12)	5 (56)	5 (20)
Normal	731.3 (338.8-856.2)	639.5 (591.7-737.5)	743.91 (266.4-1950.9)	188.7 (154.1-412.2)
Polyp	188.6 (114.4-252.1)	113.9 (49.2 – 204.7)	50.2 (0-146.3)	103.8 (90.2-110.2)

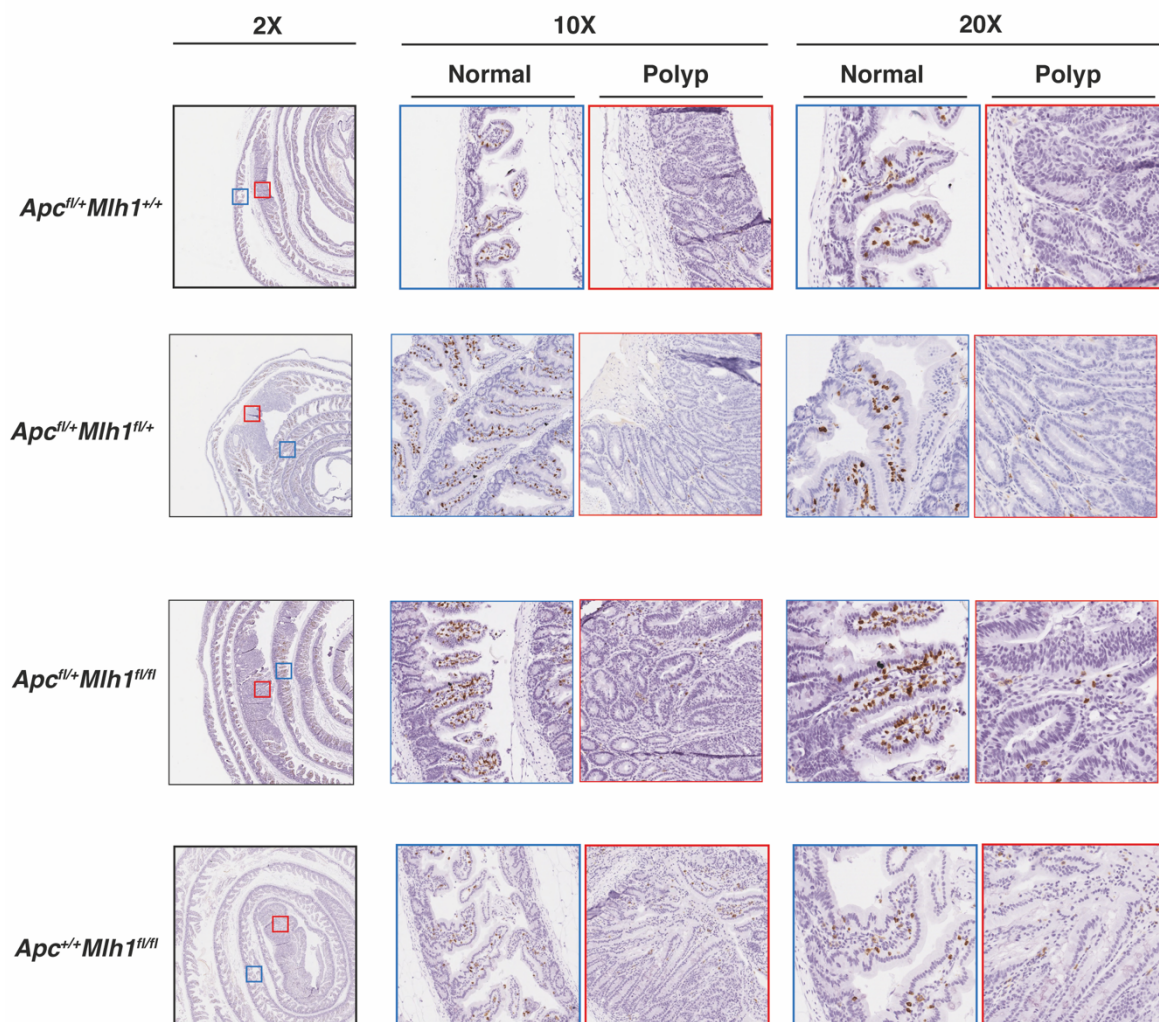


Figure 5-9. **Examples of IHC for CD8⁺ cells in normal intestinal mucosa and intestinal adenomas** in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null mice shown at 2x, 10x and 20x magnification. Visual inspection revealed marked CD8⁺ cell exclusion in polyps in each of the four genotypes.

In contrast to the normal mucosa, polyps arising in *Apc* del, *Apc* del/*Mlh1* del, *Apc* del/*Mlh1* null and *Mlh1* null mice were characterised by exclusion of CD8⁺ cells (Figure 5.9), with significantly lower levels of CD8⁺ cell in intestinal polyps compared to normal mucosa of the same genotype ($P=0.0021$, $P=0.0039$ and $P=0.003$ by Mann-Whitney U Test) (Table 5.8). Furthermore, there was significantly lower density of polyp CD8⁺ cells in *Apc* del/*Mlh1* null compared with *Apc* del (median CD8⁺ cell density 50.2 IQR (0-146.3) Vs 188.6 IQR (114.4-252.1) cells/mm², $P=0.0093$ Mann Whitney U test) (Figure 5.8B). There was no significant difference between *Apc* del/*Mlh1* del and *Apc* del mice (188.6 IQR (114.4-252.1) versus 113.9 IQR (49.2 – 204.7) cells/mm², $P=0.12$ Mann Whitney U test) (Figure 5.8B). This is in contrast to what would be expected based on MMRd CRCs in humans, which, as I have shown in Chapter 4, are characterised by increased CD8⁺ cell infiltrate. Interestingly, whilst the number of polyps available for analysis was modest, CD8⁺ cell exclusion was also observed in the *Mlh1* null mice (103.8 (90.2-110.2) cells/mm²). Although this was not significant in comparison with the normal mucosa from the same genotype ($P=0.23$, Mann Whitney U test). One previously hypothesised mechanism of immune cell exclusion is upregulation of Wnt signalling, which in CRC is most commonly associated with *APC* mutations [448]. The observation that CD8⁺ cell exclusion was present in polyps from mice with *Mlh1* loss in the background of wild-type *Apc* suggests either that these lesions have acquired (epi)mutations which activate the Wnt pathway or that alternative mechanisms are responsible. One method to detect Wnt pathway activation is IHC for nuclear localisation of its downstream target beta-catenin. Unfortunately, I did not have sufficient time to perform this in my inducible models. However, previous work our group did not show an association between CD8⁺ cell exclusion and B-catenin levels in *Apc*^{min} *Mlh1*^{del} mice.

5.4.2.2. CD4⁺ T lymphocyte infiltration

Whilst much effort has focused on the significance of CD8⁺ cytotoxic T-cells in the anti-tumour immune response, the important contribution of CD4⁺ T-helper cells is becoming increasingly well recognised [216]. T-helper cells are highly versatile, multi-functional cells, which are essential for coordinating several elements of the adaptive immune response [216]. Specifically, CD4⁺ cells are essential for the CD8⁺ cytotoxic T cells response against cancers, and for maintaining their

effectiveness [449, 450]. Furthermore, work presented in Chapter 4 and done by others have shown that MMRd/MSI CRCs exhibit enrichment for T_H1-like subset of CD4⁺ cells that express the effector cytokine *IFNG* [451] (Figure 4.9). Moreover, it seems that IFN- γ signalling is integral to the adaptive immune response in MMRd/MSI CRCs (Figures 4.11 and 4.15), in marked contrast to MMRd/MSI ECs. As such, I wanted to explore CD4⁺ infiltration in my intestinal *Mlh1* loss mouse models.

CD4⁺ cells were frequently observed in the normal intestinal mucosa of WT and *Vil-CreERT2* mice (Figure 5.10). Furthermore, CD4⁺ cells were also observed in the normal mucosa of all four experimental genotypes (*Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null) (Figure 5.11), and density did not differ significantly between *Apc* del (median 373.1 IQR(283.2-502.3) CD4⁺ cells/mm²), *Apc* del/*Mlh1* del (616.3 IQR(541.9-660.7) CD4⁺ cells/mm²), *Apc* del/*Mlh1* null (446.9 IQR(294.4 – 463.9) CD4⁺ cells/mm²) and *Mlh1* null mice (555.5 IQR(411.1-575.5) CD4⁺ cells/mm²) (no significant differences for all comparisons calculated by pairwise Wilcox with Dunn's test correction for multiple testing) (Table 5.9). Interestingly, it appeared that CD4⁺ cells resided predominantly within the stromal compartment of the normal intestine (Figure 5.11). Strikingly, and in marked contrast to CD8⁺ cells, there was no evidence of CD4⁺ exclusion from polyps (Figure 5.12).

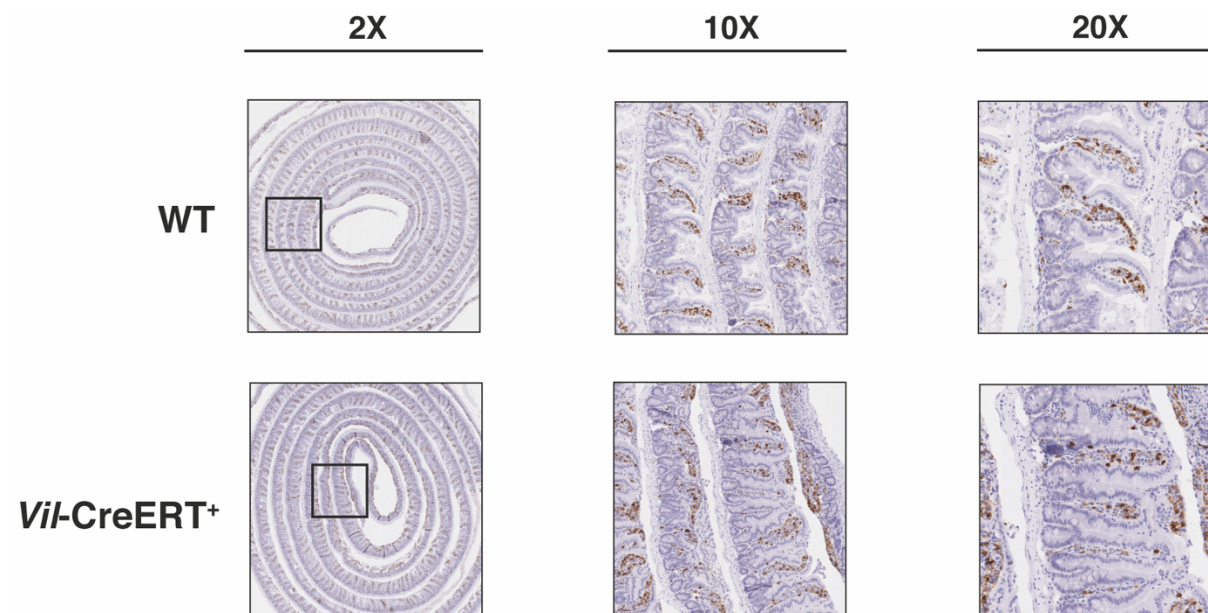
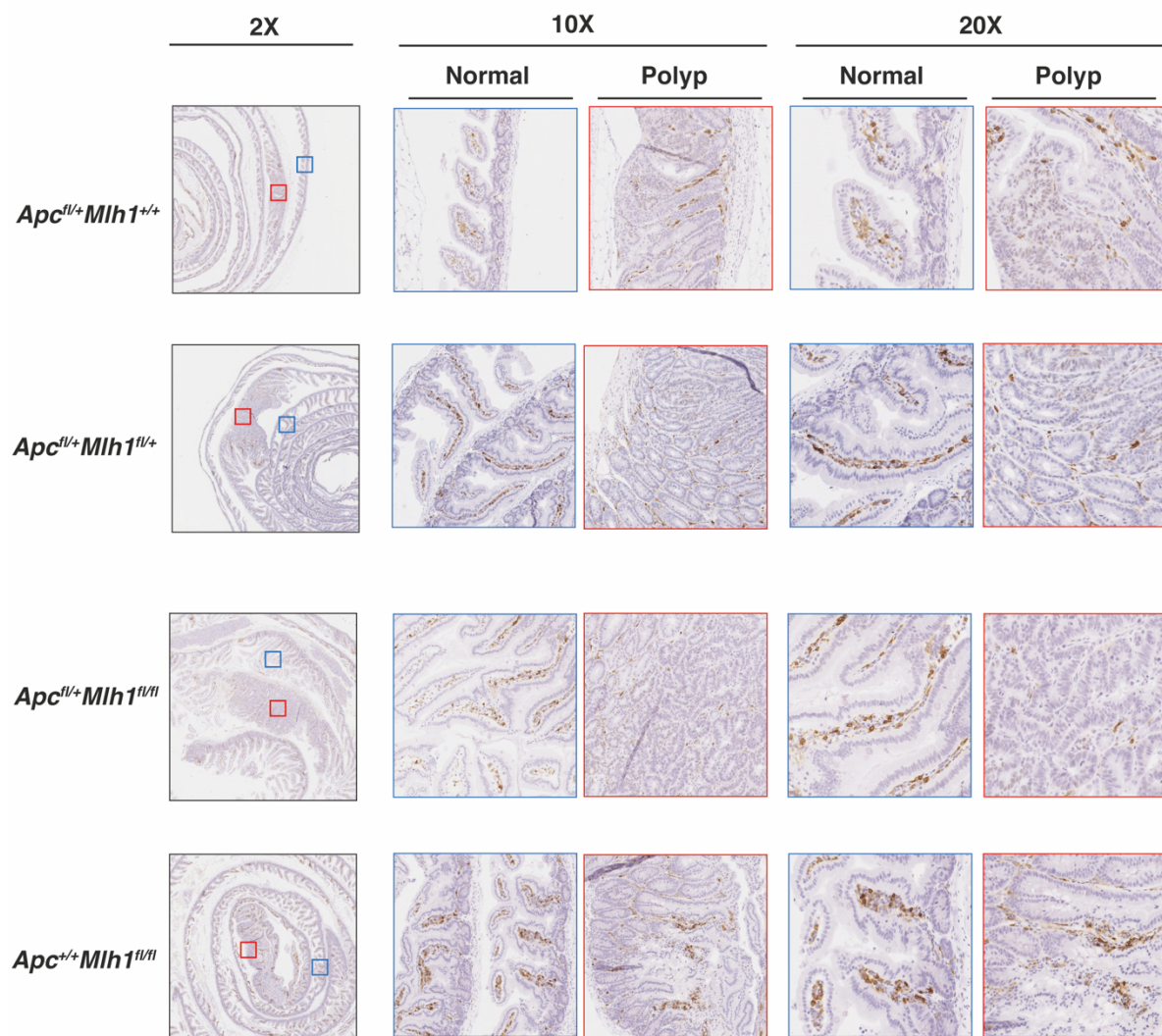


Figure 5-10. Examples of IHC from CD4⁺ cells in normal intestinal mucosa of Black-6 WT and *Vil-CreERT2* mice. Shown at 2x, 10x and 20x magnification. CD4⁺ cell infiltrate was widespread across normal mucosal tissue

3422 In contrast to CD8⁺ cell infiltrate, there was no evidence of CD4⁺ exclusion from intestinal polyps.
 3423 *Apc* del mice had significantly higher intra-polyp CD4⁺ infiltrate than normal epithelium (620.0
 3424 IQR(531.0-709.1) versus 373.1 IQR(283.2-502.3) $P=0.027$ by Mann Whitney). The difference
 3425 between normal mucosa and polyp CD4⁺ density was not significant for the other three genotypes.
 3426 Also interestingly, intra-polyp CD4⁺ cell density did not differ significantly between *Apc* del (median
 3427 620.0 IQR(531.0 – 709.1) CD4⁺cells/mm²), *Apc* del/*Mlh1* del (812.9 IQR(753.5-817.0)
 3428 CD4⁺cells/mm²), *Apc* del/*Mlh1* null (568.1 IQR(404.7-737.0) CD4⁺cells/mm²) and *Mlh1* null mice
 3429 (1024.4 IQR(823.3-1232.2) CD4⁺cells/mm²) (no significant differences for all comparisons calculated
 3430 by pairwise Wilcox with Dunn's test to correct for multiple testing) (Table 5.9, Figure 5.12).



3431 **Figure 5-11 Examples of IHC for CD4⁺ cells in normal intestinal mucosa and intestinal polyps in *Apc* del,**
 3432 ***Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null mice shown at 2x, 10x and 20x magnification.**

3433

Table 5-9 **CD4⁺ cell infiltrate** in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null normal intestine and polyps. Median values of CD4⁺ cells/mm² presented with IQR in parentheses

	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{+/+}	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{+/fl}	<i>Apc</i> ^{+/fl} <i>Mlh1</i> ^{fl/fl}	<i>Apc</i> ^{+/+} <i>Mlh1</i> ^{fl/fl}
Mice (No. of polyps)	8 (12)	3 (5)	9 (33)	2 (2)
Normal	373.1 (283.2 – 502.3)	616.3 (541.9 – 660.7)	446.9 (294.4 – 463.9)	555.5 (411.1 – 575.5)
Polyp	620.0 (531.0 – 709.1)	812.9 (753.5 – 817.0)	568.1 (404.7 – 737.0)	1024.4 (823.3–1223.2)

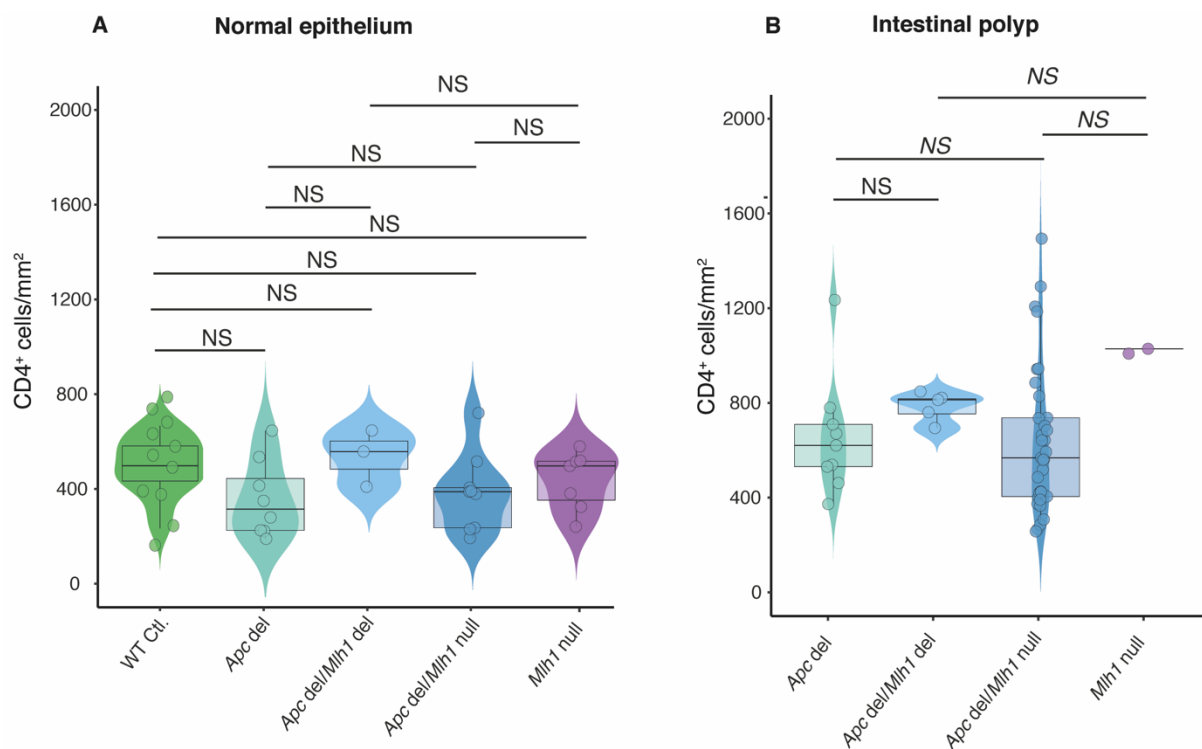


Figure 5-12. **CD4⁺ quantified by IHC** in A) normal intestinal epithelium and B) intestinal polyps in *Apc* del, *Apc* del/*Mlh1* del *Apc* del/*Mlh1* null and *Mlh1* null mice

5.4.3. CD19⁺ B lymphocyte infiltration

The adaptive immune response to cancer is also shaped by B-cells. They are able to do so through cytokine and antibody production, and through direct antigen presentation to CD4⁺ and CD8⁺ T cells [384]. In colorectal cancer, B-cell levels have been strongly correlated with CD8⁺ T-cell infiltrate and independently associated with improved survival in multivariable analysis [261]. However, analyses in Chapter 4 did not suggest that there was significantly increased B-cell infiltrate in MSI cancers compared with MSS. Given the absence of a dense CD8⁺ cell infiltrate in mouse polyps as described above, I wanted to analyse whether this was compensated for by a B-cell response.

Immunohistochemistry for the pan-B cell marker CD19 revealed that CD19⁺ cells were observed infrequently in the normal intestinal mucosa of WT and Vil-CreERT2 mice (Figure 5.13)

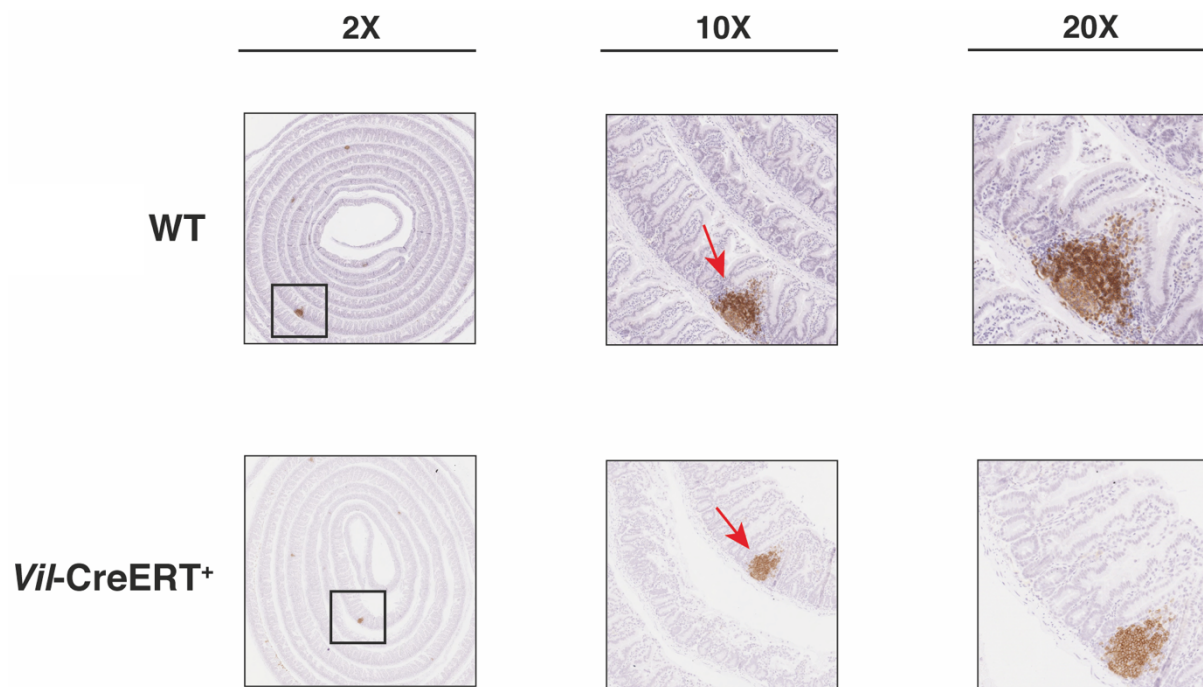


Figure 5-13. Examples of IHC from CD19⁺ cells in normal intestinal mucosa of WT and Vil-CreERT² mice shown at 2x, 10x and 20x magnification. CD19⁺ cell infiltrate was sparse in normal intestinal epithelium, except for Peyer's patch areas (red arrow)

Similarly, CD19⁺ cell infiltrate was sparse in all four experiment mice genotypes (*Apc* del, *Apc* del/*Mlh1* del, *Apc* del/*Mlh1* null and *Mlh1* null) (Figure 5.14). Additionally, very few CD19⁺ cells were seen in polyp tissue, irrespective of the underlying genotype. Given the very low numbers,

formal quantification of density was not undertaken. These results suggest that either B-cells do not play a significant part in the anti-tumour immune response in murine models of MMR loss, or that they localise outside the areas quantified in this study.

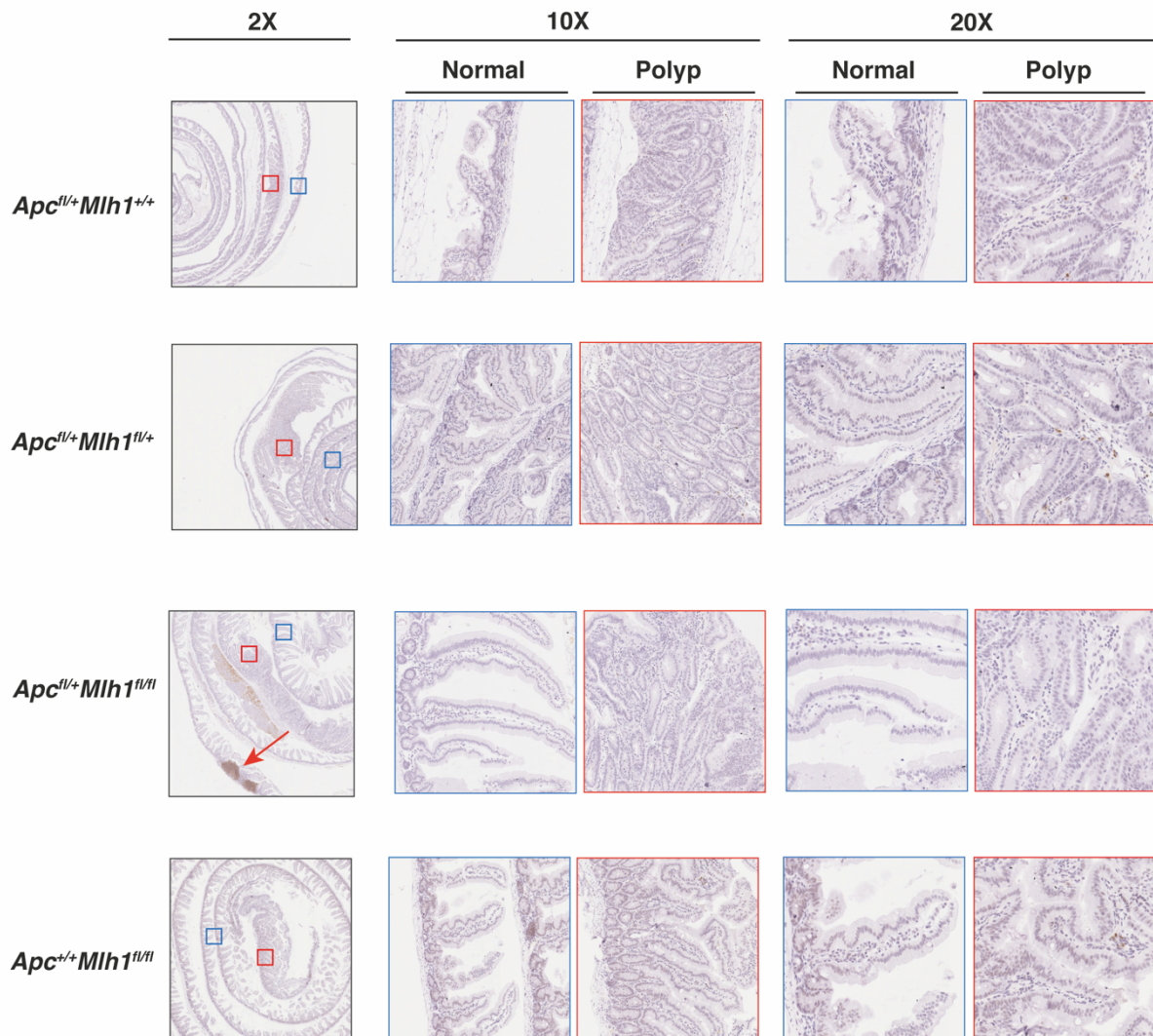


Figure 5-14 Examples of IHC for CD19⁺ cells in normal intestinal mucosa and intestinal polyps in Apc del, Apc del/Mlh1 del and Apc del/Mlh1 null and Mlh1 null mice shown at 2x, 10x and 20x magnification. Note the red arrow denotes the presence of a Peyer's patch, which acts as an internal positive control for CD19 IHC staining.

5.4.3.1. Immune escape is universal in murine intestinal polyps

In humans, the acquisition of immune escape mutations enables tumours to avoid immune-mediated cytotoxicity, and an active immune response selects for cancer cell clones that possess such genetic aberrations [171]. As discussed previously, the MHC class I complex presents intracellular peptides to CD8⁺ T cells, and is a critical component of the antitumour immune response. Beta-2 microglobulin (β_2m), the light chain of the MHC-I complex [182] is frequently targeted by loss of function mutations, which can retard this process [452]. In Chapter 4, I demonstrated an enrichment for β_2m mutations in MMRd CRCs with a mutation frequency of 22% – in agreement with previous publications [171, 183]. Whilst *B2M* loss of function mutations have been associated with MMRd in CRC, its influence on the tumour microenvironment have not been fully elucidated. I wanted to assess whether the observed exclusion of CD8⁺ cell from polyps in the mice models could be caused by loss of β_2m expression.

β_2m IHC was optimised using normal kidney, and was widely observed in the normal intestinal mucosa of all four mouse genotypes. It appeared to be strongest in the villi with limited expression in the intestinal crypts (Figure 5.15, 20 x magnification). This served as an appropriate internal control. Strikingly, examination of polyps revealed universal loss of β_2m protein in *Apc* del (10 of 14 polyps from 8 mice), *Apc* del/*Mlh1* del (11 of 15 polyps from 5 mice), *Apc* del/*Mlh1* null (31 of 36 polyps from 9 mice) and *Mlh1* null (2 of 2 polyps from 2 mice) mice polyps examined. This raises the interesting possibility that loss of β_2m expression may contribute to the absence of infiltrating CD8⁺ cell observed.

The mechanism of β_2m loss remains unknown. Further efforts are underway to delineate this and will include whole exome sequencing. It possible that the lack of labelling in tumours employs the same mechanisms as the lack of labelling in crypts and this will be an important focus of future investigation.

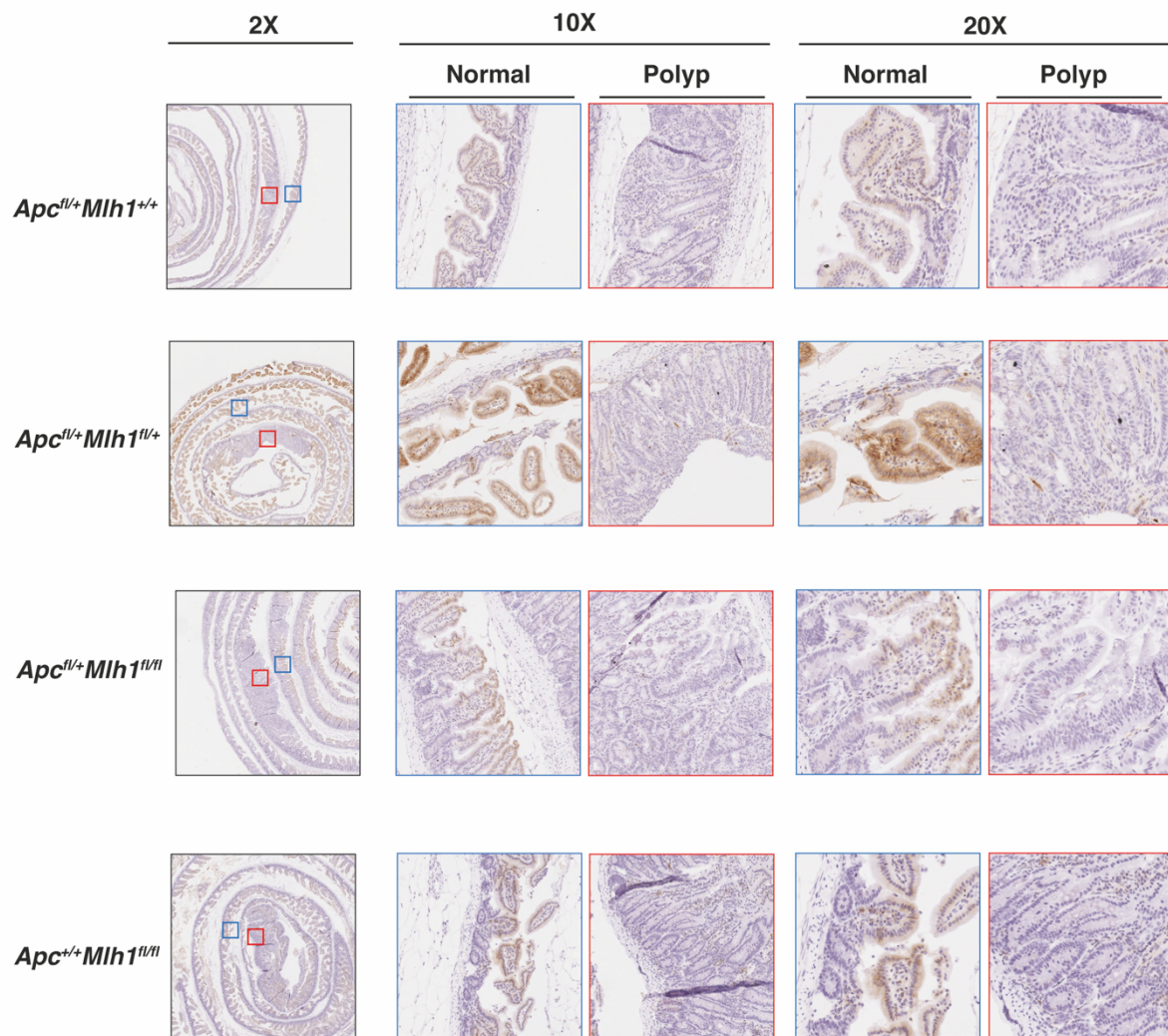


Figure 5-15. **Examples of IHC for β_{2m} cells in normal intestinal mucosa and GI polyps in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null mice shown at 2x, 10x and 20x magnification. B2m loss was universal to all polyps in each of the four genotypes examined.**

Figure 5-16. **Examples of IHC for β_{2m} cells in normal intestinal mucosa and GI polyps in *Apc* del, *Apc* del/*Mlh1* del and *Apc* del/*Mlh1* null and *Mlh1* null mice shown at 2x, 10x and 20x magnification. B2m loss was universal to all polyps in each of the four genotypes examined.**

3493

3494 5.4.4.Planned analyses not completed due to the COVID19 pandemic

3495 There were a number of further investigations I wished to perform, but unfortunately due to the
 3496 COVID-19 pandemic they were not possible. The first is to examine the genetic impact of *Mlh1* loss
 3497 in mice. I would wish to firstly to confirm that *Mlh1* loss did in fact result in MSI and increased
 3498 mutational burden, such is seen in human tumours with *Mlh1* loss. In addition to analysing the

overall mutational count, I would want to examine for the presence of specific mutations that could potentially explain the CD8⁺ cell exclusion that is seen in homozygous *Mlh1* deletion tumours shown in this work. It could be possible that *Apc* del/*Mlh1* null mice have acquired additional mutations in genes key to *MHC* class I pathway and T cell recruitment. To this end, WES on large mouse tumours from *Apc* del and *Apc* del/*Mlh1* null mice is planned.

Furthermore, I would want to expand my analysis of the immune microenvironment. Due to cost and time constraints, immune cell quantification was carried out using IHC on individual slides. To better investigate the immune cells, a multispectral IF approach would be extremely useful. Along with analysing lymphocyte populations, it would be beneficial to examine for other immune cells that could explain the lack for CD8⁺ cell infiltrate, such as myeloid derived suppressor cells. Multispectral IF quantification would also allow for determination of the spatial interaction between immune cell subsets. My work presented above has presented the surprising finding that β_2m loss appears to be ubiquitous to murine intestinal adenomas. An interesting line of analysis would be to see whether β_2m deletion alone is sufficient to cause murine intestinal tumorigenesis, and whether induced β_2m deletion exacerbates polyp formation in *Apc* loss or *Mlh1* null mice.

Finally, one further area that would greatly benefit from functional studies is to investigate why MMRd has discordant effects in CRC and EC. The development of endometrial specific *Mlh1* loss model is underway. Comparisons would include of analysis of immune cell infiltrate, tumour mutational burden and immune escape mutation enrichment. Furthermore, it will be interesting to analyse the effect of *Jak1* mutations in murine CRCs, and β_2m mutations in murine ECs.

5.5. Discussion

As discussed in detail in Chapter 4, MMRd plays a critical role in the development of CRC [304]. Moreover, it is particularly relevant in clinical management of the disease: MMRd is a biomarker for better prognosis [95] and more favourable response to immunotherapies [208]. These two characteristics are at least in part the result of a vigorous anti-tumour immune response that typifies MMRd CRCs. Genetically modified mouse models have been vital in advancing our understanding of cancer biology, and have led to a number of seminal discoveries [385]. Whilst previous studies have described MMRd mouse models, the degree to which they recapitulate human disease varies considerably [428]. Moreover, no studies have reported detailed immunophenotype analysis. Development of a mouse cancer model, which better portrays mismatch repair loss in humans, and description of the associated tumour immune microenvironment is needed. This will enable detailed functional study of the interplay between this particular form of genomic instability and the anti-tumour immune response.

To this end, a *Vil-CreERT2*, *Apc*^{+/^{fl}} *Mlh1*^{fl/fl} mouse model was generated to study mismatch repair loss in gastrointestinal carcinogenesis. Consistent with previous publications, both heterozygous *Apc* (*Apc*^{+/^{fl}} *Mlh1*^{+/+}) and homozygous *Mlh1* loss (*Apc*^{+/^{fl}} *Mlh1*^{fl/fl}) resulted in the development of gastrointestinal polyps. Combining heterozygous *Apc* loss and homozygous *Mlh1* loss (*Apc*^{+/^{fl}} *Mlh1*^{fl/fl}) resulted in a stronger phenotype with more polyps than either *Apc* loss or homozygous *Mlh1* loss alone. Of particular relevance, homozygous *Mlh1* loss alone (*Apc* wildtype) was sufficient to cause intestinal polyps, albeit with a substantially greater latency than those arising in the context of *Apc* loss. In mice, polyps were predominantly found in the small intestine and colonic polyps were infrequent. Surprisingly, across all mouse genotypes examined the intestinal polyps exhibited CD8⁺ T-cell exclusion. In *Apc* del and *Apc* del/*Mlh1* null, the polyp CD8⁺ cell density was significantly less than in the normal mucosa. In contrast, similar depletion was not observed for CD4⁺ cells: increased CD4⁺ cell infiltrate was observed in the polyps of *Apc* del mice compared to normal mucosa, whilst there was no significant difference for the same comparison in *Apc* del/*Mlh1* del, *Apc* del/*Mlh1* null,

Mlh1 null mice. There was little evidence of compensatory B-cell response these mice: few to no CD19⁺ cells were found in either normal mucosa or gastrointestinal polyps. The exclusion of CD8⁺ cells may be partly explained by the loss of β_2m expression in polyps, which was again observed in all four genotypes examined.

Taken together, these results demonstrate that *Vil*-CreERT2, *Apc*^{+/-} *Mlh1*^{fl/fl} loss results in an intestinal specific murine model of carcinogenesis. Whilst this model shows some similarities to human disease, it differs in several important aspects. CD8⁺ T-cell exclusion from gastrointestinal polyps appeared to be an almost universal phenomenon across all genotypes examined, irrespective of the *Apc* or *Mlh1* mutational status. Furthermore, loss of β_2m expression appeared to be ubiquitous to all intestinal polyps, again regardless of *Apc* or *Mlh1* mutational status. In human CRCs, loss of function β_2m mutations are found in approximately 30% of MSI CRCs [453] and far less frequently in MSS CRCs; there are no studies to my knowledge that have examined β_2m loss in mouse GI cancers. The finding that β_2m loss was common to all polyps examined suggests that it may be an early and fundamental step in murine GI tract carcinogenesis. The loss of β_2m inactivates MHC class I antigen presentation pathway [182], and this may explain the scarcity of intraneoplastic CD8⁺ T-cells. In contrast, CD4⁺ cells are activated via the MHC class II complex, which is independent of β_2m [384]. This may be particularly pertinent given that previous studies have suggested that β_2m loss can result in a compensatory CD4⁺ cell based anti-tumour reaction. In a murine model of β_2m deletion, it was shown that mice remained sensitive to ICI treatment (anti-PD1 and anti-CTLA4) and this was based on CD4⁺ cell-based immunity [454]. Furthermore, this relates to human disease; it has been shown patients with MMRd CRC tumours remain sensitive to ICI treatment irrespective of β_2m mutational status [455]. Results from these studies suggest that CD4⁺ cells may have a crucial role in determining ICI treatment response [455]. The absence of a meaningful B-cell anti-tumour response cannot be entirely explained by β_2m loss, except that co-localisation of CD20⁺ (human equivalent of CD19⁺) and CD8⁺ cells has been documented in breast, melanoma and ovarian cancers [166, 168, 456, 457]. It's possible that B-cell anti-cancer immunity is partly dependent on CD8⁺ cell activation.

5.5.1.Limitations

There are a number of limitations to the work presented above. Firstly, in human cancers mismatch repair loss is an initiating driving event in colorectal carcinogenesis, and APC mutations occur secondary to this. In this model, *Apc* and *Mlh1* loss are combined. As such, it is not possible to recreate the temporal relationship between key genetic events that are seen in human colorectal carcinogenesis. Secondly, whilst it has been previously documented that homozygous *Mlh1* loss in mice results MSI, this was not examined in this work due to the pandemic restriction. In humans, the strength of the antitumour response is recognised to be positively correlated with a tumour's mutational burden. And in colorectal cancer, this is most commonly a function of genomic instability. To fully validate this model of MMRd carcinogenesis, it will be important to show that *Mlh1* loss results in MSI and the consequent increased mutational rate. Additionally, this may offer insights into why increased CD8⁺ T-cell infiltrate is not seen in intestinal polyps with *Mlh1* loss. In a similar vein, this work did not confirm loss of Mlh1 protein expression. Thirdly, this work characterised a relatively limited subset of immune cell populations. A more in-depth immunophenotyping effort would be of considerable use. This work was limited by the use of consecutive FFPE sections and IHC based quantification. Whereas utilisation of multispectral immunofluorescence quantification would allow not only for analysis of more immune cell subtypes but also examination of the spatial interaction between different cell populations. Finally, the original study design had that *Apc* del mice would be the main control for experimental mice: *Apc* del/*Mlh1* del, *Apc* del/*Mlh1* null and *Mlh1* null. Improvement on the work presented above would include larger numbers of control mice (both WT and Vil-CreERT2⁺). This would allow for further characterisation of the immune cell infiltrate, particular CD8⁺ cell infiltrate, in normal mucosa tissue.

5.5.2.Future work

The work presented above lays the foundation for several avenues of future investigation. Initially, it will be important to explore the mechanisms behind the observed CD8⁺ exclusion from intestinal neoplasia. This could include analysis of further immune escape mutations beyond β_2m , such as *Hla*-

a, *Hla-b* or *TAP1/TAP2* mutations. Alternative routes of immune evasion should also be considered, which could include increased numbers of myeloid derived suppressor cells, upregulation of immune inhibitory molecules such as indoleamine 2,3-dioxygenase-1 (IDO-1) or immune checkpoints such as CTLA4, PD-1 or PD-L1. Additionally, it will be important to understand whether β_2m loss is caused by genetic mutations, or negative regulation at the transcription or post transcriptional level. Depending on these findings, it would be interesting to see whether polyp CD8⁺ infiltrate could be increased with reconstitution of β_2m expression. This work showed that polyps exhibited CD4⁺ cell infiltrate, irrespective of *Apc* and *Mlh1* mutational status. Whilst the findings of intraneoplastic T helper cells suggests that they may be contributing to the anti-tumour response, it is important to confirm that they are in fact active. In part this could be answered by knockout of CD4⁺ cells, and analysis to see whether this potentiates tumour formation. Murine models also offer the opportunity to contrast and compare the effects of mismatch deficiency in other organs beyond the intestine. Work presented in Chapter 4 highlighted that MMRd has substantially different consequences in the endometrium compared to the colon. In particular, an attenuated immune phenotype is observed in MMRd ECs, and this may be a consequence of an enrichment for *JAK1* loss of function mutations. Mice models may help to explain why MMRd results in selective enrichment for different immune escape mutations depending on anatomical site. It would be of value to understand whether β_2m loss impacts carcinogenesis in the endometrium, and whether *Jak1* loss impacts colorectal carcinogenesis; this work is underway but unfortunately delayed due to the COVID-19 pandemic. Finally, a small portion of endometrial cancers are characterised by two forms of genomic instability: MSI and *POLE* proofreading domain mutations [67]. The same has not been observed in colorectal cancer. Mouse models offer a potential method to explore why this is the case.

6. Conclusions

The work presented in this thesis demonstrates the critical importance of the antitumour immune response in determining tumour behaviour and its complex interaction with the tumour genome and epigenome[12]. Through correlative and functional analyses it provides evidence which supports the hypothesis that evasion of immune destruction is a fundamental characteristic of human cancers [9]. And it provides an initial attempt, albeit a limited one, to understand the mechanisms responsible for the considerable variation between the tumour immune microenvironment of different cancer types, as well as between tumours of the same type of cancer. Whilst several factors have been suggested to influence anti-tumour immunity, my work suggests that the underlying tumour mutational landscape is the most important determinant. Much remains to be learnt about how the interaction between tumour genetics and the immune system shapes the tumour microenvironment. This is an area of great importance; it is becoming increasingly evident that the tumour immune microenvironment strongly influences clinical outcome and response to treatment.

The following summarises and provides context to some pertinent findings from my work.

6.1. Tumour infiltrating CD8⁺ T lymphocytes are independently associated with improved recurrence free survival and overall survival in stage II/III colorectal cancer patients

Currently, treatment decisions for colorectal cancer patients are made on the basis of TNM staging system. Unfortunately, this is insufficient to capture the variation in clinical outcome observed between early-stage CRC patients. A wealth of previous studies has demonstrated the positive prognostic association of tumour infiltrating T-cells, in particular CD8⁺ T-lymphocytes, in CRC patients [4, 12]. However, translating this promise into a usable clinical biomarker has not yet been achieved. Using a large, clinical trial cohort of 1,804 patients, I showed that tumour CD8⁺ cell fraction was significantly associated with reduced recurrence risk and improved overall survival in stage II and III CRC patients [1]. Moreover, this relationship was independent of key confounding

variables, such as mismatch repair deficiency (MMRd), *POLE* proofreading domain mutations and chromosomal instability (CIN). This is particularly meaningful given that these forms of genomic instability have been proven to influence the composition of immune microenvironment. Furthermore, I demonstrated that treating CD8⁺ cell infiltrate as a continuous variable was a valid way of analysing the relationship between tumour infiltrating immune cells and clinical outcome. This is in contrast to previous studies that used forced dichotomisation analysis, which poses a number of issues. These results reinforce the prognostic utility of CD8⁺ cell quantification, and may guide future studies to avoid forced dichotomisation.

6.2. The prognostic association of CD8⁺ T lymphocyte varies according to tumour and nodal status

One notable advantage of using clinical trials is the detailed molecular and clinicopathological case annotations. It has not been previously shown whether the prognostic association of tumour infiltrating T-cells is constant across all early-stage CRCs, or whether it varies according to individual cancer characteristics. My results showed that the prognostic association of CD8⁺ cell infiltrate is not uniform across all early-stage CRCs, but rather varies according to tumour and nodal status. The association was effectively absent in CRCs at lowest risk of recurrence (pT3,N0), stronger in intermediate risk CRCs (pT4,N0 or pT1-3,N1/2) and strongest in highest risk CRCs (pT4,N1/2). This variation has not been reported previously, and if further validated, may hold important implications for the future utilisation of T-cell infiltrate as a prognostic biomarker. It may also give vital clues into the underlying immune/cancer interaction.

6.3. The consequences of mismatch repair deficiency vary between endometrial and colorectal cancer

MMRd is an important driver of carcinogenesis in endometrial and colorectal cancer. In CRC, MMRd is associated with improved clinical outcomes, which may in part be explained by a vigorous anti-tumour immune response [304]. In contrast, the same effect has not been well described in EC. By

systematic review and meta-analysis, I confirmed that MMRd is significantly associated with reduced recurrence risk and improved survival in CRC, but not in EC. This might be explained by the differences in the immune infiltrate seen in these cancers. I found that MMRd CRCs showed much greater increases in CD8⁺ T cells, and other lymphocytes of known prognostic import compared to MMRd ECs. These results were confirmed by multispectral IF quantification in FFPE tumour tissue and by analysis of TCGA RNA-seq data. Explanation of this discordance can in part be derived from the differing enrichment of immune escape mutations in MMRd CRCs versus EC. MMRd ECs showed a striking enrichment for *JAK1* loss of function mutations. I showed that these mutations are associated with abrogation of IFN- γ signalling, which is an important driver of an effective anti-tumour immune response [458]. These results have implications for the use of MMRd as a prognostic and a predictive biomarker. Immune checkpoint inhibition therapy effectiveness is dependent on tumour immune microenvironment rich in cytotoxic T-cells and other immune cells associated with IFN- γ signalling [109]. *JAK1* loss of function mutations have also been associated with immunotherapy resistance [381, 382].

6.4. A murine model of *Mlh1* loss is associated with intestinal tumorigenesis

Whilst previous studies have sought to investigate the effects of MMR loss on intestinal tumour formation in mice, there is a lack of studies that have investigate the immune microenvironment in these tumours. I showed that conditional, inducible *Mlh1* deletion combined with *Apc* loss, which was driven by Cre recombinase, resulted in intestinal tumorigenesis. In contrast to human disease, these polyps were located in the small intestine, and colonic polyps were infrequent. Strikingly, homozygous loss of *Mlh1* alone was sufficient to cause polyps, though numbers were smaller than in animals with concomitant *Apc* loss.

6.5. CD8⁺ T cell exclusion and *B2m* loss are common features of mouse intestinal adenomas

Given my findings regarding the significant association of CD8⁺ T-cells and clinical outcome in human CRC, I wanted to analyse CD8⁺ cell infiltrate in mouse tumours. Strikingly, my results showed that CD8⁺ cells were almost completely excluded from murine polyps resulting from *Mlh1* loss. Furthermore, and in contrast to expected findings based on human disease, CD8⁺ cell infiltrate was not increased in *Mlh1*^{fl/fl} *Apc*^{+/fl}/*Mlh1*^{+/fl} *Apc*^{+/fl} mice compared to *Mlh1*^{+/+} *Apc*^{+/fl} mice. Next, given that I had postulated an important role for immune escape mutations in determining the tumour immune microenvironment, I wanted to investigate whether β_2m mutations could explain the immune cell exclusion. Surprisingly, β_2m loss was common to all murine adenomas I analysed, regardless of the underlying genetic alterations. Taken together, these results show that gastrointestinal mouse *Mlh1* loss, unlike human disease, does not result in an enhanced T-cell immune infiltrate. Moreover, β_2m loss may be an early and fundamental step in murine gastrointestinal carcinogenesis.

6.6. Future Perspectives

My results above have highlighted the importance of CD8⁺ CTLs in the prognosis of early-stage CRC. Furthering our understanding of the prognostic association of more immune cell populations, and the interactions between these populations, will be of substantial benefit for efforts to introduce immune cell-based biomarkers into routine clinical use. Additionally, my results have shown that the same type of genomic instability can have substantially different effects depending the tumour's anatomical location. Elucidating why there is specific enrichment for different immune escape mutations may help us understand how cancers are able to overcome the of immunosurveillance hurdle in the carcinogenesis pathway.

The introduction of immune checkpoint inhibitors represents the paradigm of translational research: basic science understanding of tumour immunology leading to a substantial therapeutic breakthrough. These therapies have delivered remarkable results for a select group of patients with previously

Conclusions

3732 untreatable cancers, but their benefits are far from widespread. Future work must concentrate on
3733 continuing to build our understanding of the anti-tumour immunity, particularly how it is shaped
3734 through interaction with tumour genomics, and translating this knowledge into clinical advances for
3735 patients who have recurrent or currently untreatable disease.

3736

7. Appendices

7.1. Gene set enrichment analysis results

7.1.1. Top 30 upregulated gene sets for MSI CRC vs MSS CRC

ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust
GO:0045087	innate immune response	717	0.5762935	2.74612018	4.91E-60	7.07E-56
GO:0002250	adaptive immune response	373	0.60246508	2.74575731	1.76E-35	8.46E-32
GO:0034341	response to interferon-gamma	186	0.64116488	2.70051914	3.28E-22	3.94E-19
GO:0071346	cellular response to interferon-gamma	168	0.63707755	2.64567661	3.94E-20	2.47E-17
GO:0060333	interferon-gamma-mediated signaling pathway	85	0.69384899	2.63775457	1.00E-14	3.70E-12
GO:0032609	interferon-gamma production	100	0.66184277	2.61368401	3.50E-14	1.17E-11
GO:0001906	cell killing	128	0.62511577	2.55779071	7.35E-14	2.12E-11
GO:0032729	positive regulation of interferon-gamma production	59	0.70996884	2.55539056	3.62E-10	5.42E-08
GO:0051607	defense response to virus	207	0.59286871	2.55440042	6.04E-19	3.27E-16
GO:0002228	natural killer cell mediated immunity	55	0.71427579	2.54684124	1.20E-09	1.57E-07
GO:0032649	regulation of interferon-gamma production	88	0.65511298	2.52462083	3.17E-12	6.82E-10
GO:0009615	response to virus	284	0.56198772	2.50498086	8.74E-21	6.99E-18
GO:0031341	regulation of cell killing	84	0.6547814	2.48417811	1.88E-11	3.61E-09
GO:0140375	immune receptor activity	119	0.61134194	2.46230737	9.31E-13	2.27E-10
GO:0050778	positive regulation of immune response	585	0.52208857	2.45840796	6.18E-35	2.22E-31
GO:0002831	regulation of response to biotic stimulus	364	0.53745183	2.43732721	6.02E-24	7.89E-21
GO:0045088	regulation of innate immune response	274	0.54678968	2.43504789	6.13E-19	3.27E-16
GO:0042613	MHC class II protein complex	14	0.93447046	2.43484829	1.10E-09	1.46E-07
GO:0042611	MHC protein complex	22	0.84250721	2.43231492	2.23E-08	2.12E-06
GO:0050776	regulation of immune response	800	0.50716732	2.4259466	8.25E-43	5.94E-39
GO:0002833	positive regulation of response to biotic stimulus	225	0.5548974	2.42568439	1.79E-15	7.38E-13
GO:0042267	natural killer cell mediated cytotoxicity	51	0.68797079	2.42348428	7.65E-08	6.59E-06
GO:0002449	lymphocyte mediated immunity	225	0.55114681	2.409289	3.54E-15	1.38E-12
GO:0001774	microglial cell activation	41	0.72979908	2.40903881	3.98E-08	3.54E-06
GO:0002269	leukocyte activation involved in inflammatory response	41	0.72979908	2.40903881	3.98E-08	3.54E-06
GO:0002715	regulation of natural killer cell mediated immunity	40	0.73050616	2.40680268	2.22E-08	2.12E-06
GO:0150076	neuroinflammatory response	59	0.66460883	2.39212626	6.04E-08	5.27E-06
GO:0023023	MHC protein complex binding	25	0.79570973	2.383244	2.57E-07	1.93E-05
GO:0002717	positive regulation of natural killer cell mediated immunity	31	0.76163346	2.38178742	2.46E-07	1.87E-05
GO:0030101	natural killer cell activation	60	0.66143138	2.37770143	2.18E-08	2.11E-06

3741 **7.1.2. Top 30 downregulated gene sets for MSI CRC vs MSS CRC**

ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust
GO:0002002	regulation of angiotensin levels in blood	11	-0.8491573	-1.9063311	1.76E-04	0.0060584
GO:0002003	angiotensin maturation	11	-0.8491573	-1.9063311	1.76E-04	0.0060584
GO:0032281	AMPA glutamate receptor complex	17	-0.7440771	-1.8564499	7.59E-04	0.01884386
GO:0035461	vitamin transmembrane transport	13	-0.7962026	-1.85603	0.00117555	0.0266348
GO:0070696	transmembrane receptor protein serine/threonine kinase binding	17	-0.7371452	-1.8391551	0.0012063	0.02702454
GO:0016328	lateral plasma membrane	56	-0.5929213	-1.8384568	1.58E-04	0.00558331
GO:0099061	integral component of postsynaptic density membrane	36	-0.634942	-1.8309625	3.60E-04	0.01061365
GO:0006067	ethanol metabolic process	19	-0.7131015	-1.8274773	0.00231792	0.04493892
GO:1901338	catecholamine binding	11	-0.8138636	-1.827098	9.51E-04	0.02245976
GO:0099095	ligand-gated anion channel activity	12	-0.7881844	-1.8256023	0.0026369	0.04894909
GO:0098962	regulation of postsynaptic neurotransmitter receptor activity	10	-0.8288291	-1.8106784	0.00130445	0.02847063
GO:0034310	primary alcohol catabolic process	13	-0.7729961	-1.8019333	0.00239485	0.04612
GO:0060393	regulation of pathway-restricted SMAD protein phosphorylation	52	-0.5881803	-1.8001524	4.98E-04	0.01377621
GO:0033612	receptor serine/threonine kinase binding	21	-0.6910863	-1.8000842	0.0019559	0.03907725
GO:0099146	intrinsic component of postsynaptic density membrane	39	-0.6106504	-1.7951543	8.83E-04	0.02133533
GO:0060389	pathway-restricted SMAD protein phosphorylation	55	-0.5799141	-1.7935115	3.06E-04	0.00925715
GO:0004497	monooxygenase activity	78	-0.5419804	-1.7837051	1.28E-04	0.00472713
GO:0006068	ethanol catabolic process	11	-0.7939508	-1.7823943	0.00197806	0.03941075
GO:0007631	feeding behavior	68	-0.5531395	-1.7821808	3.32E-04	0.00996806
GO:0005237	inhibitory extracellular ligand-gated ion channel activity	8	-0.8701946	-1.7755669	0.00129698	0.02843682
GO:0035240	dopamine binding	5	-0.9668131	-1.7653564	6.58E-05	0.00266375
GO:0099060	integral component of postsynaptic specialization membrane	50	-0.5813076	-1.7625311	6.24E-04	0.01633643
GO:0017147	Wnt-protein binding	37	-0.6085601	-1.758955	8.45E-04	0.02056145
GO:0030917	midbrain-hindbrain boundary development	5	-0.9626848	-1.7578184	1.03E-04	0.00392387
GO:0008083	growth factor activity	128	-0.4917645	-1.7476104	1.84E-04	0.0062604
GO:0042756	drinking behavior	8	-0.8551556	-1.7448809	0.00243414	0.04668941
GO:0098948	intrinsic component of postsynaptic specialization membrane	53	-0.5653859	-1.744605	0.00108027	0.02489799
GO:0002016	regulation of blood volume by renin-angiotensin	6	-0.9243597	-1.7355375	2.67E-04	0.00831828
GO:0005179	hormone activity	66	-0.5353191	-1.7224678	0.00102518	0.02370411
GO:0098953	receptor diffusion trapping	5	-0.9399918	-1.716382	7.45E-04	0.01858981

3742

3743 **7.1.3.Top 30 upregulated gene sets for MSI EC vs MSS EC**

ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust
GO:0009226	nucleotide-sugar biosynthetic process	21	0.67714664	2.22438841	9.77E-05	0.02331239
GO:0006778	porphyrin-containing compound metabolic process	30	0.61104727	2.17747564	1.22E-04	0.02752574
GO:0042168	heme metabolic process	26	0.62355356	2.16441398	2.56E-04	0.04269825
GO:0016667	oxidoreductase activity, acting on a sulfur group of donors	52	0.50210744	2.00818291	5.61E-05	0.01588506
GO:0097237	cellular response to toxic substance	152	0.39581838	1.92268324	6.70E-07	5.83E-04
GO:0097646	calcitonin family receptor signaling pathway	4	0.9702124	1.86690917	8.47E-05	0.02103222
GO:1990410	adrenomedullin receptor signaling pathway	4	0.9702124	1.86690917	8.47E-05	0.02103222
GO:0009108	coenzyme biosynthetic process	115	0.40518845	1.86569308	1.00E-05	0.00549699
GO:0050662	coenzyme binding	210	0.37474243	1.85292177	6.81E-08	9.21E-05
GO:1901568	fatty acid derivative metabolic process	74	0.43014996	1.82708443	1.84E-04	0.03450958
GO:0051188	cofactor biosynthetic process	184	0.37514264	1.82608027	5.34E-07	5.79E-04
GO:0006732	coenzyme metabolic process	184	0.3743395	1.82217083	6.07E-07	5.79E-04
GO:0098754	detoxification	72	0.43091839	1.81712106	1.90E-04	0.03501015
GO:0016491	oxidoreductase activity	468	0.34815348	1.78010715	1.05E-13	1.27E-09
GO:0016054	organic acid catabolic process	172	0.35989775	1.77603493	6.90E-06	0.00419803
GO:0046395	carboxylic acid catabolic process	172	0.35989775	1.77603493	6.90E-06	0.00419803
GO:0097647	amylin receptor signaling pathway	3	0.99506041	1.7628042	1.03E-05	0.00549699
GO:0055114	oxidation-reduction process	667	0.30822579	1.73192093	5.14E-13	3.13E-09
GO:0005759	mitochondrial matrix	406	0.33813969	1.72036635	2.03E-10	8.25E-07
GO:0006090	pyruvate metabolic process	109	0.36699396	1.71517891	1.80E-04	0.03430643
GO:0044282	small molecule catabolic process	281	0.3404947	1.68027251	2.97E-08	4.51E-05
GO:0051186	cofactor metabolic process	302	0.34075094	1.67351456	2.70E-08	4.51E-05
GO:0042542	response to hydrogen peroxide	103	0.36548029	1.66879907	2.91E-04	0.04658269
GO:0048037	cofactor binding	325	0.31365887	1.59467595	4.60E-07	5.60E-04
GO:0006631	fatty acid metabolic process	218	0.32150459	1.56498013	5.89E-05	0.01591797
GO:0019752	carboxylic acid metabolic process	664	0.27418535	1.54582543	6.63E-09	1.34E-05
GO:0006790	sulfur compound metabolic process	246	0.31296417	1.54060154	4.43E-05	0.01380968
GO:0034599	cellular response to oxidative stress	215	0.30574248	1.51449245	2.21E-04	0.03894467
GO:0009636	response to toxic substance	295	0.29843818	1.4741614	1.27E-05	0.00556615
GO:0006875	cellular metal ion homeostasis	254	0.29208377	1.45474318	7.12E-05	0.01844023

3744

3745

3746

3747

3748 **7.1.4.Top 30 downregulated gene sets for MSI EC vs MSS EC**

ID	Description	setSize	enrichmentScore	NES	pvalue	p.adjust
GO:0005201	extracellular matrix structural constituent	76	-0.6699209	-2.1893705	3.01E-09	9.14E-06
GO:0099055	integral component of postsynaptic membrane	20	-0.7495716	-1.9307752	1.78E-04	0.03430643
GO:0035455	response to interferon-alpha	17	-0.766587	-1.9101572	1.58E-04	0.03154439
GO:0060337	type I interferon signaling pathway	68	-0.5852244	-1.8929476	5.47E-05	0.01585901
GO:0071357	cellular response to type I interferon	68	-0.5852244	-1.8929476	5.47E-05	0.01585901
GO:0031012	extracellular matrix	215	-0.5194645	-1.8747182	5.86E-09	1.34E-05
GO:0034340	response to type I interferon	72	-0.5758138	-1.8725912	3.06E-05	0.01053427
GO:0062023	collagen-containing extracellular matrix	186	-0.5048691	-1.7972397	6.18E-07	5.79E-04
GO:0007411	axon guidance	134	-0.5140074	-1.7912636	1.28E-05	0.00556615
GO:0097485	neuron projection guidance	134	-0.5140074	-1.7912636	1.28E-05	0.00556615
GO:0042734	presynaptic membrane	54	-0.5721218	-1.7660485	3.06E-04	0.04779806
GO:0006617	SRP-dependent cotranslational protein targeting to membrane, signal sequence recognition	4	-0.9757961	-1.6865562	1.04E-04	0.02393695
GO:0001654	eye development	178	-0.4599092	-1.6320788	1.25E-04	0.02764906
GO:0150063	visual system development	180	-0.4568097	-1.6226181	9.53E-05	0.02320259
GO:0048667	cell morphogenesis involved in neuron differentiation	321	-0.4415983	-1.6224371	5.37E-06	0.00363409
GO:0048880	sensory system development	181	-0.4561886	-1.6217708	6.16E-05	0.0162922
GO:0043010	camera-type eye development	155	-0.4576585	-1.6095642	3.05E-04	0.04779806
GO:2001199	negative regulation of dendritic cell differentiation	3	-0.9931967	-1.5952128	3.14E-05	0.01053427
GO:0048812	neuron projection morphogenesis	347	-0.4284907	-1.5803002	4.69E-06	0.00352361
GO:0120039	plasma membrane bounded cell projection morphogenesis	359	-0.4223092	-1.5599921	1.05E-05	0.00549699
GO:0048858	cell projection morphogenesis	362	-0.4215149	-1.557184	1.19E-05	0.00556615
GO:0007409	axonogenesis	248	-0.4267887	-1.5529123	2.35E-04	0.04080547
GO:0010975	regulation of neuron projection development	290	-0.4222451	-1.5468766	1.53E-04	0.03135071
GO:0045664	regulation of neuron differentiation	366	-0.4180963	-1.5439182	2.37E-05	0.00902327
GO:0032990	cell part morphogenesis	379	-0.4148801	-1.5330217	1.08E-05	0.00549699
GO:0005887	integral component of plasma membrane	523	-0.3952332	-1.4753324	2.10E-05	0.00824861
GO:0050767	regulation of neurogenesis	450	-0.3917751	-1.4533853	1.04E-04	0.02393695
GO:0031175	neuron projection development	562	-0.3881804	-1.4527834	3.20E-05	0.01053427
GO:0009887	animal organ morphogenesis	555	-0.3872437	-1.4484325	3.09E-05	0.01053427
GO:0051960	regulation of nervous system development	493	-0.3893171	-1.4475535	5.36E-05	0.01585901

3749

3750

8. References

1. Glaire, M.A., et al., *Tumour-infiltrating CD8(+) lymphocytes and colorectal cancer recurrence by tumour and nodal stage*. Br J Cancer, 2019. **121**(6): p. 474-482.
2. Ijsselstein, M.E., et al., *Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification*. J Pathol Clin Res, 2019. **5**(1): p. 3-11.
3. Kuipers, E.J., et al., *Colorectal cancer*. Nat Rev Dis Primers, 2015. **1**: p. 15065.
4. Bruni, D., H.K. Angell, and J. Galon, *The immune contexture and Immunoscore in cancer prognosis and therapeutic efficacy*. Nat Rev Cancer, 2020. **20**(11): p. 662-680.
5. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 2021. **71**(3): p. 209-249.
6. Jiricny, J., *The multifaceted mismatch-repair system*. Nat Rev Mol Cell Biol, 2006. **7**(5): p. 335-46.
7. Omran, A.R., *The epidemiologic transition. A theory of the epidemiology of population change*. Milbank Mem Fund Q, 1971. **49**(4): p. 509-38.
8. Lee, E.Y.H.P. and W.J. Muller, *Oncogenes and tumor suppressor genes*. Cold Spring Harbor perspectives in biology, 2010. **2**(10): p. a003236-a003236.
9. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: The next generation*. 2011. p. 646-674.
10. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
11. Virchow, R., *Cellular pathology. As based upon physiological and pathological histology. Lecture XVI--Atheromatous affection of arteries*. 1858. Nutr Rev, 1989. **47**(1): p. 23-5.
12. Galon, J. and D. Bruni, *Tumor Immunology and Tumor Evolution: Intertwined Histories*. Immunity, 2020. **52**(1): p. 55-81.
13. Robert, C., *A decade of immune-checkpoint inhibitors in cancer therapy*. Nat Commun, 2020. **11**(1): p. 3801.
14. Sharma, P. and James P. Allison, *Immune Checkpoint Targeting in Cancer Therapy: Toward Combination Strategies with Curative Potential*. Cell, 2015. **161**(2): p. 205-214.
15. Rizvi, N.A., et al., *Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer*. Science, 2015. **348**(6230): p. 124-8.
16. Van Allen, E.M., et al., *Genomic correlates of response to CTLA4 blockade in metastatic melanoma*. Science, 2015. **350**(6257): p. 207-211.
17. Cunningham, D., et al., *Colorectal cancer*. Lancet, 2010. **375**(9719): p. 1030-47.
18. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012*. European journal of cancer (Oxford, England : 1990), 2013. **49**(6): p. 1374-403.
19. Bray, F., et al., *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. CA Cancer J Clin, 2018. **68**(6): p. 394-424.
20. UK, C.R. *Bowel Cancer Statistics*. 2021 [cited 2021 21st September 2021]; Available from: <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/bowel-cancer#heading-One>.
21. Hardcastle, J.D., et al., *Randomised controlled trial of faecal-occult-blood screening for colorectal cancer*. Lancet, 1996. **348**(9040): p. 1472-7.
22. Henrikson, N.B., et al., *Family history and the natural history of colorectal cancer: systematic review*. Genet Med, 2015. **17**(9): p. 702-12.
23. Vasen, H.F., I. Tomlinson, and A. Castells, *Clinical management of hereditary colorectal cancer syndromes*. Nat Rev Gastroenterol Hepatol, 2015. **12**(2): p. 88-97.
24. Tomlinson, I., *The Mendelian colorectal cancer syndromes*. Annals of clinical biochemistry, 2015. **52**(Pt 6): p. 690-692.
25. Lynch, H.T., et al., *Hereditary colorectal cancer syndromes: molecular genetics, genetic counseling, diagnosis and management*. Fam Cancer, 2007. **7**(1): p. 27-39.
26. Nieuwenhuis, M.H., et al., *Genotype predicting phenotype in familial adenomatous polyposis: a practical application to the choice of surgery*. Dis Colon Rectum, 2009. **52**(7): p. 1259-63.
27. Monahan, K.J., et al., *Guidelines for the management of hereditary colorectal cancer from the British Society of Gastroenterology (BSG)/Association of Coloproctology of Great Britain and Ireland (ACPGBI)/United Kingdom Cancer Genetics Group (UKCGG)*. Gut, 2020. **69**(3): p. 411-444.
28. Cheadle, J.P. and J.R. Sampson, *MUTYH-associated polyposis--from defect in base excision repair to clinical genetic testing*. DNA Repair (Amst), 2007. **6**(3): p. 274-9.
29. Nielsen, M., et al., *MUTYH-associated polyposis (MAP)*. Crit Rev Oncol Hematol, 2011. **79**(1): p. 1-16.
30. Sampson, J.R. and N. Jones, *MUTYH-associated polyposis*. Best Pract Res Clin Gastroenterol, 2009. **23**(2): p. 209-18.
31. Song, M., W.S. Garrett, and A.T. Chan, *Nutrients, foods, and colorectal cancer prevention*. Gastroenterology, 2015. **148**(6): p. 1244-60.e16.
32. Botteri, E., et al., *Smoking and colorectal cancer: a meta-analysis*. Jama, 2008. **300**(23): p. 2765-78.
33. Liang, P.S., T.Y. Chen, and E. Giovannucci, *Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis*. Int J Cancer, 2009. **124**(10): p. 2406-15.
34. Cai, S., et al., *Alcohol drinking and the risk of colorectal cancer death: a meta-analysis*. Eur J Cancer Prev, 2014. **23**(6): p. 532-9.
35. Kyrgiou, M., et al., *Adiposity and cancer at major anatomical sites: umbrella review of the literature*. Bmj, 2017. **356**: p. j477.
36. Guraya, S.Y., *Association of type 2 diabetes mellitus and the risk of colorectal cancer: A meta-analysis and systematic review*. World J Gastroenterol, 2015. **21**(19): p. 6026-31.
37. Algra, A.M. and P.M. Rothwell, *Effects of regular aspirin on long-term cancer incidence and metastasis: a systematic comparison of evidence from observational studies versus randomised trials*. Lancet Oncol, 2012. **13**(5): p. 518-27.
38. Grady, W.M. and J.M. Carethers, *Genomic and epigenetic instability in colorectal cancer pathogenesis*. Gastroenterology, 2008. **135**(4): p. 1079-99.
39. Colussi, D., et al., *Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention*. Int J Mol Sci, 2013. **14**(8): p. 16365-85.
40. Negrini, S., V.G. Gorgoulis, and T.D. Halazonetis, *Genomic instability--an evolving hallmark of cancer*. Nat Rev Mol Cell Biol, 2010. **11**(3): p. 220-8.
41. Zeki, S.S., T.A. Graham, and N.A. Wright, *Stem cells and their implications for colorectal cancer*. Nature Reviews Gastroenterology & Hepatology, 2011. **8**(2): p. 90-100.

References

42. Barker, N., et al., *Identification of stem cells in small intestine and colon by marker gene Lgr5*. Nature, 2007. **449**(7165): p. 1003-1007.
43. Buczaccki, S.J.A., et al., *Intestinal label-retaining cells are secretory precursors expressing Lgr5*. Nature, 2013. **495**(7439): p. 65-69.
44. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis*. Cell, 1990. **61**(5): p. 759-67.
45. Dekker, E., et al., *Colorectal cancer*. Lancet, 2019. **394**(10207): p. 1467-1480.
46. Rowan, A.J., et al., *APC and mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits"*. Proceedings of the National Academy of Sciences, 2000. **97**(7): p. 3352.
47. Loree, J.M., et al., *Classifying Colorectal Cancer by Tumor Location Rather than Sidedness Highlights a Continuum in Mutation Profiles and Consensus Molecular Subtypes*. Clin Cancer Res, 2018. **24**(5): p. 1062-1072.
48. Guinney, J., et al., *The consensus molecular subtypes of colorectal cancer*. Nature Medicine, 2015. **21**(11): p. 1350-6.
49. Argilés, G., et al., *Localised colon cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up*. Ann Oncol, 2020. **31**(10): p. 1291-1305.
50. Brierley, J.D., M.K. Gospodarowicz, and C. Wittekind, *TNM Classification of Malignant Tumours*. 2017, Chichester, UNITED KINGDOM: John Wiley & Sons, Incorporated.
51. Biller, L.H. and D. Schrag, *Diagnosis and Treatment of Metastatic Colorectal Cancer: A Review*. Jama, 2021. **325**(7): p. 669-685.
52. André, T., et al., *Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer*. N Engl J Med, 2020. **383**(23): p. 2207-2218.
53. NICE. *Colorectal cancer: Management of metastatic disease*. 2021 [cited 2021 21st September 2021]; Available from: <https://www.nice.org.uk/guidance/ng151/chapter/Recommendations#management-of-metastatic-disease>.
54. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018*. Eur J Cancer, 2018. **103**: p. 356-387.
55. UK, C.R. *Uterine Cancer statistics*. 2021 [cited 2021 21st September 2021].
56. Moller, P., et al., *Cancer incidence and survival in Lynch syndrome patients receiving colonoscopic and gynaecological surveillance: first report from the prospective Lynch syndrome database*. Gut, 2015.
57. Ryan, N.A.J., et al., *Association of Mismatch Repair Mutation With Age at Cancer Onset in Lynch Syndrome: Implications for Stratified Surveillance Strategies*. JAMA Oncol, 2017. **3**(12): p. 1702-1706.
58. Amant, F., et al., *Cancer of the corpus uteri*. Int J Gynaecol Obstet, 2018. **143 Suppl 2**: p. 37-50.
59. Grady, D., et al., *Hormone replacement therapy and endometrial cancer risk: a meta-analysis*. Obstet Gynecol, 1995. **85**(2): p. 304-13.
60. Calle, E.E. and R. Kaaks, *Overweight, obesity and cancer: epidemiological evidence and proposed mechanisms*.
61. Renehan, A.G., et al., *Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies*. Lancet, 2008. **371**(9612): p. 569-78.
62. Bergman, L., et al., *Risk and prognosis of endometrial cancer after tamoxifen for breast cancer. Comprehensive Cancer Centres' ALERT Group. Assessment of Liver and Endometrial cancer Risk following Tamoxifen*. Lancet, 2000. **356**(9233): p. 881-7.
63. Swerdlow, A.J. and M.E. Jones, *Tamoxifen treatment for breast cancer and risk of endometrial cancer: a case-control study*. J Natl Cancer Inst, 2005. **97**(5): p. 375-84.
64. Brinton, L.A., et al., *Reproductive, menstrual, and medical risk factors for endometrial cancer: results from a case-control study*. Am J Obstet Gynecol, 1992. **167**(5): p. 1317-25.
65. Setiawan, V.W., et al., *Type I and II endometrial cancers: have they different risk factors?* J Clin Oncol, 2013. **31**(20): p. 2607-18.
66. Russo, M., et al., *Clonal evolution in paired endometrial intraepithelial neoplasia/atypical hyperplasia and endometrioid adenocarcinoma*. Hum Pathol, 2017. **67**: p. 69-77.
67. Cancer Genome Atlas Research, N., et al., *Integrated genomic characterization of endometrial carcinoma*. Nature, 2013. **497**(7447): p. 67-73.
68. Colombo, N., et al., *Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up*. Ann Oncol, 2013. **24 Suppl 6**: p. vi33-8.
69. Timmermans, A., et al., *Endometrial thickness measurement for detecting endometrial cancer in women with postmenopausal bleeding: a systematic review and meta-analysis*. Obstet Gynecol, 2010. **116**(1): p. 160-167.
70. Pecorelli, S., *Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium*. Int J Gynaecol Obstet, 2009. **105**(2): p. 103-4.
71. Saleh, M., et al., *Endometrial Cancer, the Current International Federation of Gynecology and Obstetrics Staging System, and the Role of Imaging*. J Comput Assist Tomogr, 2020. **44**(5): p. 714-729.
72. Lu, K.H. and R.R. Broaddus, *Endometrial Cancer*. N Engl J Med, 2020. **383**(21): p. 2053-2064.
73. de Boer, S.M., et al., *Adjuvant chemoradiotherapy versus radiotherapy alone in women with high-risk endometrial cancer (PORTEC-3): patterns of recurrence and post-hoc survival analysis of a randomised phase 3 trial*. Lancet Oncol, 2019. **20**(9): p. 1273-1285.
74. Randall, M.E., et al., *Phase III Trial: Adjuvant Pelvic Radiation Therapy Versus Vaginal Brachytherapy Plus Paclitaxel/Carboplatin in High-Intermediate and High-Risk Early Stage Endometrial Cancer*. J Clin Oncol, 2019. **37**(21): p. 1810-1818.
75. Fong, W., Q. Li, and J. Yu, *Gut microbiota modulation: a novel strategy for prevention and treatment of colorectal cancer*. Oncogene, 2020. **39**(26): p. 4925-4943.
76. Aguilera, A. and T. García-Muse, *Causes of genome instability*. Annu Rev Genet, 2013. **47**: p. 1-32.
77. Helleday, T., S. Eshtad, and S. Nik-Zainal, *Mechanisms underlying mutational signatures in human cancers*. Nature reviews. Genetics, 2014. **15**(9): p. 585-598.
78. Alexandrov, L.B., et al., *Signatures of mutational processes in human cancer*. Nature, 2013. **500**(7463): p. 415-21.
79. Alexandrov, L.B., et al., *The repertoire of mutational signatures in human cancer*. Nature, 2020. **578**(7793): p. 94-101.
80. Drake, J.W., et al., *Rates of spontaneous mutation*. Genetics, 1998. **148**(4): p. 1667-86.
81. Kunkel, T.A., *DNA replication fidelity*. J Biol Chem, 2004. **279**(17): p. 16895-8.
82. Shevelev, I.V. and U. Hubscher, *The 3' 5' exonucleases*. Nat Rev Mol Cell Biol, 2002. **3**(5): p. 364-76.
83. Albertson, T.M., et al., *DNA polymerase epsilon and delta proofreading suppress discrete mutator and cancer phenotypes in mice*. Proc Natl Acad Sci U S A, 2009. **106**(40): p. 17101-4.
84. Murphy, K., et al., *A method to select for mutator DNA polymerase deltas in Saccharomyces cerevisiae*. Genome, 2006. **49**(4): p. 403-10.

References

85. Simon, M., L. Giot, and G. Faye, *The 3' to 5' exonuclease activity located in the DNA polymerase delta subunit of Saccharomyces cerevisiae is required for accurate replication*. EMBO J, 1991. **10**(8): p. 2165-70.
86. Goldsby, R.E., et al., *High incidence of epithelial cancers in mice deficient for DNA polymerase delta proofreading*. Proc Natl Acad Sci U S A, 2002. **99**(24): p. 15560-5.
87. TCGA, *Comprehensive molecular characterization of human colon and rectal cancer*. Nature, 2012. **487**(7407): p. 330-7.
88. Rayner, E., et al., *A panoply of errors: polymerase proofreading domain mutations in cancer*. Nat Rev Cancer, 2016. **16**(2): p. 71-81.
89. Palles, C., et al., *Germline mutations affecting the proofreading domains of POLE and POLD1 predispose to colorectal adenomas and carcinomas*. Nature genetics, 2013. **45**(2): p. 136-44.
90. Domingo, E., et al., *Somatic POLE proofreading domain mutation, immune response, and prognosis in colorectal cancer: a retrospective, pooled biomarker study*. Lancet Gastroenterol Hepatol, 2016. **1**(3): p. 207-216.
91. van Gool, I.C., et al., *POLE Proofreading Mutations Elicit an Antitumor Immune Response in Endometrial Cancer*. Clinical cancer research : an official journal of the American Association for Cancer Research, 2015. **21**(14): p. 3347-55.
92. Schumacher, T.N., W. Scheper, and P. Kvistborg, *Cancer Neoantigens*. Annu Rev Immunol, 2019. **37**: p. 173-200.
93. Kunkel, T.A. and D.A. Erie, *Eukaryotic Mismatch Repair in Relation to DNA Replication*. Annu Rev Genet, 2015. **49**: p. 291-313.
94. Hause, R.J., et al., *Classification and characterization of microsatellite instability across 18 cancer types*. Nat Med, 2016. **22**(11): p. 1342-1350.
95. Popat, S., *Systematic Review of Microsatellite Instability and Colorectal Cancer Prognosis*. Journal of Clinical Oncology, 2004. **23**(3): p. 609-618.
96. Balkwill, F. and A. Mantovani, *Inflammation and cancer: back to Virchow?* Lancet, 2001. **357**(9255): p. 539-45.
97. Grivennikov, S.I., F.R. Greten, and M. Karin, *Immunity, Inflammation, and Cancer*. Cell, 2010. **140**(6): p. 883-99.
98. Kanda, Y., M. Osaki, and F. Okada, *Chemopreventive Strategies for Inflammation-Related Carcinogenesis: Current Status and Future Direction*. Int J Mol Sci, 2017. **18**(4).
99. Edelmann, W., et al., *Tumorigenesis in Mlh1 and Mlh1/Apc1638N mutant mice*. Cancer Res, 1999. **59**(6): p. 1301-7.
100. Colotta, F., et al., *Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability*. Carcinogenesis, 2009. **30**(7): p. 1073-81.
101. Lin, R., et al., *Chronic inflammation-associated genomic instability paves the way for human esophageal carcinogenesis*. Oncotarget, 2016. **7**(17): p. 24564-71.
102. Templeton, A.J., et al., *Prognostic role of neutrophil-to-lymphocyte ratio in solid tumors: a systematic review and meta-analysis*. J Natl Cancer Inst, 2014. **106**(6): p. djv124.
103. Kowalczyk, O., et al., *CXCL5 as a potential novel prognostic factor in early stage non-small cell lung cancer: results of a study of expression levels of 23 genes*. Tumour Biol, 2014. **35**(5): p. 4619-28.
104. Li, A., et al., *Overexpression of CXCL5 is associated with poor survival in patients with pancreatic cancer*. Am J Pathol, 2011. **178**(3): p. 1340-9.
105. Zhou, S.L., et al., *CXCL5 contributes to tumor metastasis and recurrence of intrahepatic cholangiocarcinoma by recruiting infiltrative intratumoral neutrophils*. Carcinogenesis, 2014. **35**(3): p. 597-605.
106. Zhu, X., et al., *CXCL5 is a potential diagnostic and prognostic marker for bladder cancer patients*. Tumour Biol, 2016. **37**(4): p. 4569-77.
107. Vesely, M.D., et al., *Natural innate and adaptive immunity to cancer*. Annual review of immunology, 2011. **29**: p. 235-71.
108. Dunn, G.P., et al., *Cancer immunoediting: from immunosurveillance to tumor escape*. Nature Immunology, 2002. **3**(11): p. 991-998.
109. Fridman, W.H., et al., *The immune contexture in cancer prognosis and treatment*. Nat Rev Clin Oncol, 2017. **14**(12): p. 717-734.
110. Galon, J., et al., *Towards the introduction of the 'Immunoscore' in the classification of malignant tumours*. J Pathol, 2014. **232**(2): p. 199-209.
111. Llosa, N.J., et al., *The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints*. Cancer discovery, 2015. **5**(1): p. 43-51.
112. Glasner, A., et al., *NKp46 Receptor-Mediated Interferon-γ Production by Natural Killer Cells Increases Fibronectin 1 to Alter Tumor Architecture and Control Metastasis*. Immunity, 2018. **48**(1): p. 107-119.e4.
113. Ivashkiv, L.B., *IFNγ: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy*. Nat Rev Immunol, 2018. **18**(9): p. 545-558.
114. Parker, B.S., J. Rautela, and P.J. Hertzog, *Antitumour actions of interferons: implications for cancer therapy*. Nat Rev Cancer, 2016. **16**(3): p. 131-44.
115. Schroeder, H.W., Jr. and L. Cavacini, *Structure and function of immunoglobulins*. J Allergy Clin Immunol, 2010. **125**(2 Suppl 2): p. S41-52.
116. Martinez, F.O. and S. Gordon, *The M1 and M2 paradigm of macrophage activation: time for reassessment*. F1000Prime Rep, 2014. **6**: p. 13.
117. Forssell, J., et al., *High macrophage infiltration along the tumor front correlates with improved survival in colon cancer*. Clin Cancer Res, 2007. **13**(5): p. 1472-9.
118. Lackner, C., et al., *Prognostic relevance of tumour-associated macrophages and von Willebrand factor-positive microvessels in colorectal cancer*. Virchows Arch, 2004. **445**(2): p. 160-7.
119. Zhou, Q., et al., *The density of macrophages in the invasive front is inversely correlated to liver metastasis in colon cancer*. J Transl Med, 2010. **8**: p. 13.
120. Mahmoud, S.M., et al., *Tumour-infiltrating macrophages and clinical outcome in breast cancer*. J Clin Pathol, 2012. **65**(2): p. 159-63.
121. Yin, S., et al., *The Prognostic and Clinicopathological Significance of Tumor-Associated Macrophages in Patients with Gastric Cancer: A Meta-Analysis*. PLoS One, 2017. **12**(1): p. e0170042.
122. Zhang, Q.W., et al., *Prognostic significance of tumor-associated macrophages in solid tumor: a meta-analysis of the literature*. PLoS One, 2012. **7**(12): p. e50946.
123. Zhao, X., et al., *Prognostic significance of tumor-associated macrophages in breast cancer: a meta-analysis of the literature*. Oncotarget, 2017. **8**(18): p. 30576-30586.
124. Böttcher, J.P., et al., *NK Cells Stimulate Recruitment of cDC1 into the Tumor Microenvironment Promoting Cancer Immune Control*. Cell, 2018. **172**(5): p. 1022-1037.e14.
125. Malietzis, G., et al., *Prognostic Value of the Tumour-Infiltrating Dendritic Cells in Colorectal Cancer: A Systematic Review*. Cell Commun Adhes, 2015. **22**(1): p. 9-14.

References

126. Sandel, M.H., et al., *Prognostic value of tumor-infiltrating dendritic cells in colorectal cancer: role of maturation status and intratumoral localization*. Clin Cancer Res, 2005. **11**(7): p. 2576-82.
127. Coca, S., et al., *The prognostic significance of intratumoral natural killer cells in patients with colorectal carcinoma*. Cancer, 1997. **79**(12): p. 2320-8.
128. Eckl, J., et al., *Transcript signature predicts tissue NK cell content and defines renal cell carcinoma subgroups independent of TNM staging*. J Mol Med (Berl), 2012. **90**(1): p. 55-66.
129. Melero, I., et al., *T-cell and NK-cell infiltration into solid tumors: a key limiting factor for efficacious cancer immunotherapy*. Cancer Discov, 2014. **4**(5): p. 522-6.
130. Menon, A.G., et al., *Down-regulation of HLA-A expression correlates with a better prognosis in colorectal cancer patients*. Lab Invest, 2002. **82**(12): p. 1725-33.
131. Zhu, L.Y., et al., *[Prognostic significance of natural killer cell infiltration in hepatocellular carcinoma]*. Ai Zheng, 2009. **28**(11): p. 1198-202.
132. Angell, T.E., et al., *Circulating Myeloid-Derived Suppressor Cells Predict Differentiated Thyroid Cancer Diagnosis and Extent*. Thyroid, 2016. **26**(3): p. 381-9.
133. Huang, A., et al., *Increased CD14(+)HLA-DR (-/low) myeloid-derived suppressor cells correlate with extrathoracic metastasis and poor response to chemotherapy in non-small cell lung cancer patients*. Cancer Immunol Immunother, 2013. **62**(9): p. 1439-51.
134. Jordan, K.R., et al., *Myeloid-derived suppressor cells are associated with disease progression and decreased overall survival in advanced-stage melanoma patients*. Cancer Immunol Immunother, 2013. **62**(11): p. 1711-22.
135. Li, X., et al., *Neutrophil count is associated with myeloid derived suppressor cell level and presents prognostic value of for hepatocellular carcinoma patients*. Oncotarget, 2017. **8**(15): p. 24380-24388.
136. Yang, G., et al., *Accumulation of myeloid-derived suppressor cells (MDSCs) induced by low levels of IL-6 correlates with poor prognosis in bladder cancer*. Oncotarget, 2017. **8**(24): p. 38378-38388.
137. Fleischmann, A., et al., *Immunological microenvironment in prostate cancer: high mast cell densities are associated with favorable tumor characteristics and good prognosis*. Prostate, 2009. **69**(9): p. 976-81.
138. Mehdawi, L., et al., *High tumor mast cell density is associated with longer survival of colon cancer patients*. Acta Oncol, 2016. **55**(12): p. 1434-1442.
139. Rajput, A.B., et al., *Stromal mast cells in invasive breast cancer are a marker of favourable prognosis: a study of 4,444 cases*. Breast Cancer Res Treat, 2008. **107**(2): p. 249-57.
140. Tan, S.Y., et al., *Prognostic significance of cell infiltrations of immunosurveillance in colorectal cancer*. World J Gastroenterol, 2005. **11**(8): p. 1210-4.
141. Al-Shibli, K.I., et al., *Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer*. Clin Cancer Res, 2008. **14**(16): p. 5220-7.
142. Lugli, A., et al., *CD8+ lymphocytes/ tumour-budding index: an independent prognostic factor representing a 'pro-/anti-tumour' approach to tumour host interaction in colorectal cancer*. Br J Cancer, 2009. **101**(8): p. 1382-92.
143. Naito, Y., et al., *CD8+ T cells infiltrated within cancer cell nests as a prognostic factor in human colorectal cancer*. Cancer Res, 1998. **58**(16): p. 3491-4.
144. Sinicrope, F.A., et al., *Association of immune markers and Immunoscore with survival of stage III colon carcinoma (CC) patients (pts) treated with adjuvant FOLFOX: NCCTG N0147 (Alliance)*. Journal of Clinical Oncology, 2017. **35**(15_suppl): p. 3579-3579.
145. Kondratiev, S., et al., *Intratumoral CD8+ T lymphocytes as a prognostic factor of survival in endometrial carcinoma*. Clin Cancer Res, 2004. **10**(13): p. 4450-6.
146. Pagès, F., et al., *Effector memory T cells, early metastasis, and survival in colorectal cancer*. The New England journal of medicine, 2005. **353**(25): p. 2654-66.
147. Pagès, F., et al., *In Situ Cytotoxic and Memory T Cells Predict Outcome in Patients With Early-Stage Colorectal Cancer*. <http://dx.doi.org/10.1200/JCO.2008.19.6147>, 2016.
148. Pages, F., et al., *International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study*. Lancet, 2018. **391**(10135): p. 2128-2139.
149. Prall, F., et al., *Prognostic role of CD8+ tumor-infiltrating lymphocytes in stage III colorectal cancer with and without microsatellite instability*. Hum Pathol, 2004. **35**(7): p. 808-16.
150. Tosolini, M., et al., *Clinical impact of different classes of infiltrating T cytotoxic and helper cells (Th1, th2, treg, th17) in patients with colorectal cancer*. Cancer Res, 2011. **71**(4): p. 1263-71.
151. Oldford, S.A., et al., *Tumor cell expression of HLA-DM associates with a Th1 profile and predicts improved survival in breast carcinoma patients*. Int Immunol, 2006. **18**(11): p. 1591-602.
152. Kondo, T., et al., *Favorable prognosis of renal cell carcinoma with increased expression of chemokines associated with a Th1-type immune response*. Cancer Sci, 2006. **97**(8): p. 780-6.
153. Galon, J., et al., *Type, Density, and Location of Immune Cells Within Human Colorectal Tumors Predict Clinical Outcome*. Science, 2006. **313**(September): p. 1960-1964.
154. Camus, M., et al., *Coordination of intratumoral immune reaction and human colorectal cancer recurrence*. Cancer Res, 2009. **69**(6): p. 2685-93.
155. De Monte, L., et al., *Intratumor T helper type 2 cell infiltrate correlates with cancer-associated fibroblast thymic stromal lymphopoietin production and reduced survival in pancreatic cancer*. J Exp Med, 2011. **208**(3): p. 469-78.
156. Chen, X., et al., *Increased IL-17-producing cells correlate with poor survival and lymphangiogenesis in NSCLC patients*. Lung Cancer, 2010. **69**(3): p. 348-54.
157. Zhang, J.P., et al., *Increased intratumoral IL-17-producing cells correlate with poor survival in hepatocellular carcinoma patients*. J Hepatol, 2009. **50**(5): p. 980-9.
158. Bindea, G., et al., *Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer*. Immunity, 2013. **39**(4): p. 782-95.
159. Gu-Trantien, C., et al., *CD4+ follicular helper T cell infiltration predicts breast cancer survival*. J Clin Invest, 2013. **123**(7): p. 2873-92.
160. Badoual, C., et al., *Prognostic value of tumor-infiltrating CD4+ T-cell subpopulations in head and neck cancers*. Clin Cancer Res, 2006. **12**(2): p. 465-72.
161. Frey, D.M., et al., *High frequency of tumor-infiltrating FOXP3(+) regulatory T cells predicts improved survival in mismatch repair-proficient colorectal cancer patients*. Int J Cancer, 2010. **126**(11): p. 2635-43.

References

162. Li, J.F., et al., *The prognostic value of peritumoral regulatory T cells and its correlation with intratumoral cyclooxygenase-2 expression in clear cell renal cell carcinoma*. BJU Int, 2009. **103**(3): p. 399-405.
163. Salama, P., et al., *Tumor-Infiltrating FOXP3+ T Regulatory Cells Show Strong Prognostic Significance in Colorectal Cancer*. <http://dx.doi.org/10.1200/JCO.2008.18.7229>, 2016.
164. Shah, W., et al., *A reversed CD4/CD8 ratio of tumor-infiltrating lymphocytes and a high percentage of CD4(+)FOXP3(+) regulatory T cells are significantly associated with clinical outcome in squamous cell carcinoma of the cervix*. Cell Mol Immunol, 2011. **8**(1): p. 59-66.
165. Winerdal, M.E., et al., *FOXP3 and survival in urinary bladder cancer*. BJU Int, 2011. **108**(10): p. 1672-8.
166. Ladányi, A., et al., *Prognostic impact of B-cell density in cutaneous melanoma*. Cancer Immunol Immunother, 2011. **60**(12): p. 1729-38.
167. Milne, K., et al., *Systematic analysis of immune infiltrates in high-grade serous ovarian cancer reveals CD20, FoxP3 and TIA-1 as positive prognostic factors*. PLoS One, 2009. **4**(7): p. e6412.
168. Nielsen, J.S., et al., *CD20+ tumor-infiltrating lymphocytes have an atypical CD27- memory phenotype and together with CD8+ T cells promote favorable prognosis in ovarian cancer*. Clin Cancer Res, 2012. **18**(12): p. 3281-92.
169. Xu, Y., S. Lan, and Q. Zheng, *Prognostic significance of infiltrating immune cell subtypes in invasive ductal carcinoma of the breast*. Tumori, 2018. **104**(3): p. 196-201.
170. Yeong, J., et al., *High Densities of Tumor-Associated Plasma Cells Predict Improved Prognosis in Triple Negative Breast Cancer*. Front Immunol, 2018. **9**: p. 1209.
171. Rooney, Michael S., et al., *Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity*. Cell, 2015. **160**(1-2): p. 48-61.
172. Postow, M.A., M.K. Callahan, and J.D. Wolchok, *Immune Checkpoint Blockade in Cancer Therapy*. Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2015. **33**(17): p. 1974-U161.
173. Chen, Q., et al., *Immunodominant CD4+ responses identified in a patient vaccinated with full-length NY-ESO-1 formulated with ISCOMATRIX adjuvant*. Proc Natl Acad Sci U S A, 2004. **101**(25): p. 9363-8.
174. D'Angelo, S.P., et al., *Antitumor Activity Associated with Prolonged Persistence of Adoptively Transferred NY-ESO-1 (c259)T Cells in Synovial Sarcoma*. Cancer Discov, 2018. **8**(8): p. 944-957.
175. Davis, I.D., et al., *Recombinant NY-ESO-1 protein with ISCOMATRIX adjuvant induces broad integrated antibody and CD4(+) and CD8(+) T cell responses in humans*. Proc Natl Acad Sci U S A, 2004. **101**(29): p. 10697-702.
176. Jager, E., et al., *Induction of primary NY-ESO-1 immunity: CD8+ T lymphocyte and antibody responses in peptide-vaccinated patients with NY-ESO-1+ cancers*. Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12198-203.
177. Boon, T., et al., *Tumor antigens recognized by T lymphocytes*. Annual review of immunology, 1994. **12**: p. 337-65.
178. Chen, Y.T., et al., *A testicular antigen aberrantly expressed in human cancers detected by autologous antibody screening*. Proc Natl Acad Sci U S A, 1997. **94**(5): p. 1914-8.
179. Marchand, M., et al., *Tumor regressions observed in patients with metastatic melanoma treated with an antigenic peptide encoded by gene MAGE-3 and presented by HLA-A1*. Int J Cancer, 1999. **80**(2): p. 219-30.
180. Schreiber, R.D., L.J. Old, and M.J. Smyth, *Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion*. Science (New York, N.Y.), 2011. **331**(6024): p. 1565-70.
181. Medema, J.P., et al., *Immune escape of tumors in vivo by expression of cellular FLICE-inhibitory protein*. J Exp Med, 1999. **190**(7): p. 1033-8.
182. Neeffjes, J., et al., *Towards a systems understanding of MHC class I and MHC class II antigen presentation*. Nat Rev Immunol, 2011. **11**(12): p. 823-36.
183. Grasso, C.S., et al., *Genetic Mechanisms of Immune Evasion in Colorectal Cancer*. Cancer Discov, 2018. **8**(6): p. 730-749.
184. Brunet, J.F., et al., *A new member of the immunoglobulin superfamily--CTLA-4*. Nature, 1987. **328**(6127): p. 267-70.
185. Ishida, Y., et al., *Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death*. Embo j, 1992. **11**(11): p. 3887-95.
186. Freeman, G.J., et al., *Engagement of the PD-1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation*. J Exp Med, 2000. **192**(7): p. 1027-34.
187. Keller, M.F., et al., *Trans-ethnic meta-analysis of white blood cell phenotypes*. Hum Mol Genet, 2014. **23**(25): p. 6944-60.
188. Mangino, M., et al., *Innate and adaptive immune traits are differentially affected by genetic and environmental factors*. Nat Commun, 2017. **8**: p. 13850.
189. Roederer, M., et al., *The genetic architecture of the human immune system: a bioresource for autoimmunity and disease pathogenesis*. Cell, 2015. **161**(2): p. 387-403.
190. Vogelsang, M., et al., *The Expression Quantitative Trait Loci in Immune Pathways and their Effect on Cutaneous Melanoma Prognosis*. Clin Cancer Res, 2016. **22**(13): p. 3268-80.
191. Sayaman, R.W., et al., *Germline genetic contribution to the immune landscape of cancer*. Immunity, 2021. **54**(2): p. 367-386.e8.
192. Chowell, D., et al., *Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy*. Science, 2018. **359**(6375): p. 582-587.
193. Queirolo, P., et al., *Association of CTLA-4 Gene Variants with Response to Therapy and Long-term Survival in Metastatic Melanoma Patients Treated with Ipilimumab: An Italian Melanoma Intergroup Study*. Front Immunol, 2017. **8**: p. 386.
194. Chat, V., et al., *Autoimmune genetic risk variants as germline biomarkers of response to melanoma immune-checkpoint inhibition*. Cancer Immunol Immunother, 2019. **68**(6): p. 897-905.
195. Cani, P.D., *Human gut microbiome: hopes, threats and promises*. Gut, 2018. **67**(9): p. 1716-1725.
196. Gonzalez, C.A. and E. Riboli, *Diet and cancer prevention: Contributions from the European Prospective Investigation into Cancer and Nutrition (EPIC) study*. Eur J Cancer, 2010. **46**(14): p. 2555-62.
197. Bullman, S., et al., *Analysis of Fusobacterium persistence and antibiotic response in colorectal cancer*. Science, 2017. **358**(6369): p. 1443-1448.
198. Rubinstein, M.R., et al., *Fusobacterium nucleatum promotes colorectal cancer by inducing Wnt/ β -catenin modulator Annexin A1*. EMBO Rep, 2019. **20**(4).
199. Iida, N., et al., *Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment*. Science, 2013. **342**(6161): p. 967-70.
200. Vetizou, M., et al., *Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota*. Science, 2015. **350**(6264): p. 1079-84.
201. Sivan, A., et al., *Commensal Bifidobacterium promotes antitumor immunity and facilitates anti-PD-L1 efficacy*. Science, 2015. **350**(6264): p. 1084-9.

References

202. Matson, V., et al., *The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients*. Science, 2018. **359**(6371): p. 104-108.
203. Iwai, Y., et al., *Involvement of PD-L1 on tumor cells in the escape from host immune system and tumor immunotherapy by PD-L1 blockade*. Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12293-7.
204. Hodi, F.S., et al., *Improved survival with ipilimumab in patients with metastatic melanoma*. N Engl J Med, 2010. **363**(8): p. 711-23.
205. Borghaei, H., et al., *Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer*. N Engl J Med, 2015. **373**(17): p. 1627-39.
206. Jenkins, R.W., D.A. Barbie, and K.T. Flaherty, *Mechanisms of resistance to immune checkpoint inhibitors*. British Journal of Cancer, 2018. **118**(1): p. 9-16.
207. Makker, V., et al., *Lenvatinib plus pembrolizumab in patients with advanced endometrial cancer: an interim analysis of a multicentre, open-label, single-arm, phase 2 trial*. Lancet Oncol, 2019. **20**(5): p. 711-718.
208. Le, D.T., et al., *PD-1 Blockade in Tumors with Mismatch-Repair Deficiency*. New England Journal of Medicine, 2015. **372**(26): p. 2509-2520.
209. Tran, E., et al., *Cancer immunotherapy based on mutation-specific CD4+ T cells in a patient with epithelial cancer*. Science (New York, N.Y.), 2014. **344**(6184): p. 641-5.
210. Wang, Q., et al., *Unifying cancer and normal RNA sequencing data from different sources*. Sci Data, 2018. **5**: p. 180061.
211. Patro, R., et al., *Salmon provides fast and bias-aware quantification of transcript expression*. Nat Methods, 2017. **14**(4): p. 417-419.
212. Ellrott, K., et al., *Scalable Open Science Approach for Mutation Calling of Tumor Exomes Using Multiple Genomic Pipelines*. Cell Syst, 2018. **6**(3): p. 271-281.e7.
213. Weber, J., et al., *Adjuvant Nivolumab versus Ipilimumab in Resected Stage III or IV Melanoma*. N Engl J Med, 2017. **377**(19): p. 1824-1835.
214. Eggermont, A.M.M., et al., *Adjuvant Pembrolizumab versus Placebo in Resected Stage III Melanoma*. N Engl J Med, 2018. **378**(19): p. 1789-1801.
215. Diederichsen, A.C., et al., *Prognostic value of the CD4+/CD8+ ratio of tumour infiltrating lymphocytes in colorectal cancer and HLA-DR expression on tumour cells*. Cancer Immunol Immunother, 2003. **52**(7): p. 423-8.
216. Tay, R.E., E.K. Richardson, and H.C. Toh, *Revisiting the role of CD4(+) T cells in cancer immunotherapy-new insights into old paradigms*. Cancer Gene Ther, 2021. **28**(1-2): p. 5-17.
217. Frey, D.M., et al., *High frequency of tumor-infiltrating FOXP3+ regulatory T cells predicts improved survival in mismatch repair-proficient colorectal cancer patients*. International Journal of Cancer, 2017. **126**(11): p. 2635-2643.
218. Marisa, L., et al., *The Balance Between Cytotoxic T-cell Lymphocytes and Immune Checkpoint Expression in the Prognosis of Colon Tumors*. J Natl Cancer Inst, 2018. **110**(1).
219. Eriksen, A.C., et al., *The Prognostic Value of Tumor-Infiltrating lymphocytes in Stage II Colon Cancer. A Nationwide Population-Based Study*. Transl Oncol, 2018. **11**(4): p. 979-987.
220. Bemtsson, J., et al., *The clinical impact of tumour-infiltrating lymphocytes in colorectal cancer differs by anatomical subsite: A cohort study*. Int J Cancer, 2017. **141**(8): p. 1654-1666.
221. Daster, S., et al., *Absence of myeloperoxidase and CD8 positive cells in colorectal cancer infiltrates identifies patients with severe prognosis*. Oncoimmunology, 2015. **4**(12): p. e1050574.
222. Deschoolmeester, V., et al., *Tumor infiltrating lymphocytes: an intriguing player in the survival of colorectal cancer patients*. BMC Immunol, 2010. **11**: p. 19.
223. Mlecnik, B., et al., *Integrative Analyses of Colorectal Cancer Show Immunoscore Is a Stronger Predictor of Patient Survival Than Microsatellite Instability*. Immunity, 2016. **44**(3): p. 698-711.
224. Wirta, E.V., et al., *Immunoscore in mismatch repair-proficient and -deficient colon cancer*. J Pathol Clin Res, 2017. **3**(3): p. 203-213.
225. Pagès, F., et al., *Prognostic and predictive value of the Immunoscore in stage III colon cancer patients treated with oxaliplatin in the prospective IDEA France PRODIGE-GERCOR cohort study*. Ann Oncol, 2020. **31**(7): p. 921-929.
226. Simon, R.M., S. Paik, and D.F. Hayes, *Use of archived specimens in evaluation of prognostic and predictive biomarkers*. J Natl Cancer Inst, 2009. **101**(21): p. 1446-52.
227. Davoli, T., et al., *Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy*. Science, 2017. **355**(6322).
228. Walther, A., R. Houlston, and I. Tomlinson, *Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis*. Gut, 2008. **57**(7): p. 941-50.
229. Roth, A.D., et al., *Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial*. J Clin Oncol, 2010. **28**(3): p. 466-74.
230. Taieb, J., et al., *Prognostic Value of BRAF and KRAS Mutations in MSI and MSS Stage III Colon Cancer*. J Natl Cancer Inst, 2017. **109**(5).
231. Hamarshesh, S.a., et al., *Immune modulatory effects of oncogenic KRAS in cancer*. Nature Communications, 2020. **11**(1): p. 5439.
232. Ischenko, I., et al., *KRAS drives immune evasion in a genetic model of pancreatic cancer*. Nature Communications, 2021. **12**(1): p. 1482.
233. Albitar, M., et al., *Expression of PD-L1 in colorectal cancer that lack mutations in RAS or TP53 genes*. Journal of Clinical Oncology, 2017. **35**(15_suppl): p. e14500-e14500.
234. Marco, M., et al., *A KRAS mutation is associated with an immunosuppressive tumor microenvironment in mismatch-repair proficient colorectal cancer*. Journal of Clinical Oncology, 2019. **37**(4_suppl): p. 609-609.
235. DG, A. and R. P., *The cost of dichotomising continuous variables*. BMJ (Clinical research ed.), 2006. **332**(7549).
236. Mlecnik, B., et al., *Comprehensive Intrametastatic Immune Quantification and Major Impact of Immunoscore on Survival*. J Natl Cancer Inst, 2018. **110**(1).
237. Royston, P., D.G. Altman, and W. Sauerbrei, *Dichotomizing continuous predictors in multiple regression: a bad idea*. Stat Med, 2006. **25**(1): p. 127-41.
238. Kerr, R.S., et al., *Adjuvant capecitabine plus bevacizumab versus capecitabine alone in patients with colorectal cancer (QUASAR 2): an open-label, randomised phase 3 trial*. Lancet Oncol, 2016. **17**(11): p. 1543-1557.
239. Midgley, R.S., et al., *Phase III randomized trial assessing rofecoxib in the adjuvant setting of colorectal cancer: final results of the VICTOR trial*. J Clin Oncol, 2010. **28**(30): p. 4575-80.
240. Hurwitz, H., et al., *Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer*. N Engl J Med, 2004. **350**(23): p. 2335-42.

References

241. Blaker, Y.N., et al., *Computerized image analysis of the Ki-67 proliferation index in mantle cell lymphoma*. *Histopathology*, 2015. **67**(1): p. 62-9.
242. Domingo, E., et al., *Evaluation of PIK3CA mutation as a predictor of benefit from nonsteroidal anti-inflammatory drug therapy in colorectal cancer*. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 2013. **31**(34): p. 4297-305.
243. McShane, L.M., et al., *REporting recommendations for tumour MARKer prognostic studies (REMARK)*, in *Br J Cancer*. 2005. p. 387-91.
244. Little, R.J., et al., *The prevention and treatment of missing data in clinical trials*. *N Engl J Med*, 2012. **367**(14): p. 1355-60.
245. van Buuren, S., H.C. Boshuizen, and D.L. Knook, *Multiple imputation of missing blood pressure covariates in survival analysis*. *Stat Med*, 1999. **18**(6): p. 681-94.
246. Graham, J.W., *Missing data analysis: making it work in the real world*. *Annu Rev Psychol*, 2009. **60**: p. 549-76.
247. Sterne, J.A., et al., *Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls*. *Bmj*, 2009. **338**: p. b2393.
248. Marshall, A., D.G. Altman, and R.L. Holder, *Comparison of imputation methods for handling missing covariate data when fitting a Cox proportional hazards model: a resampling study*. *BMC Med Res Methodol*, 2010. **10**: p. 112.
249. White, I.R., P. Royston, and A.M. Wood, *Multiple imputation using chained equations: Issues and guidance for practice*. *Stat Med*, 2011. **30**(4): p. 377-99.
250. Kioussis, D. and W. Ellmeier, *Chromatin and CD4, CD8A and CD8B gene expression during thymic differentiation*. *Nat Rev Immunol*, 2002. **2**(12): p. 909-19.
251. WL, A., et al., *Transcriptional subtyping and CD8 immunohistochemistry identifies poor prognosis stage II/III colorectal cancer patients who benefit from adjuvant chemotherapy*. *JCO precision oncology*, 2018. **2018**.
252. E, S., et al., *An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types*. *Journal of molecular medicine (Berlin, Germany)*, 2009. **87**(6).
253. F, d.S.E.M., et al., *Methylation of cancer-stem-cell-associated Wnt target genes predicts poor prognosis in colorectal cancer patients*. *Cell stem cell*, 2011. **9**(5).
254. L, M., et al., *Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value*. *PLoS medicine*, 2013. **10**(5).
255. M, S., et al., *Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer*. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. **106**(17).
256. AW, T., et al., *Extent and Location of Tumor-Infiltrating Lymphocytes in Microsatellite-Stable Colon Cancer Predict Outcome to Adjuvant Active Specific Immunotherapy*. *Clinical cancer research : an official journal of the American Association for Cancer Research*, 2016. **22**(2).
257. Iveson, T.J., et al., *3 versus 6 months of adjuvant oxaliplatin-fluoropyrimidine combination therapy for colorectal cancer (SCOT): an international, randomised, phase 3, non-inferiority trial*. *The Lancet Oncology*, 2018. **19**(4): p. 562-578.
258. Kim, K.J., et al., *Prognostic implications of tumor-infiltrating FoxP3+ regulatory T cells and CD8+ cytotoxic T cells in microsatellite-unstable gastric cancers*. *Hum Pathol*, 2014. **45**(2): p. 285-93.
259. Sinicrope, F.A., et al., *Intraepithelial effector (CD3+)/regulatory (FoxP3+) T-cell ratio predicts a clinical outcome of human colon carcinoma*. *Gastroenterology*, 2009. **137**(4): p. 1270-9.
260. Suzuki, H., et al., *Intratumoral CD8(+) T/FOXP3 (+) cell ratio is a predictive marker for survival in patients with colorectal cancer*. *Cancer Immunol Immunother*, 2010. **59**(5): p. 653-61.
261. Edin, S., et al., *The Prognostic Importance of CD20(+) B lymphocytes in Colorectal Cancer and the Relation to Other Immune Cell subsets*. *Sci Rep*, 2019. **9**(1): p. 19997.
262. Nosh, K., et al., *Tumour-infiltrating T-cell subsets, molecular changes in colorectal cancer and prognosis: cohort study and literature review*. *J Pathol*, 2010. **222**(4): p. 350-66.
263. Turksma, A.W., et al., *Extent and Location of Tumor-Infiltrating Lymphocytes in Microsatellite-Stable Colon Cancer Predict Outcome to Adjuvant Active Specific Immunotherapy*. *Clin Cancer Res*, 2016. **22**(2): p. 346-56.
264. Horeweg, N., et al., *Prognostic Integrated Image-Based Immune and Molecular Profiling in Early-Stage Endometrial Cancer*. *Cancer Immunol Res*, 2020. **8**(12): p. 1508-1519.
265. Garrido-Ramos, M.A., *Satellite DNA: An Evolving Topic*. *Genes (Basel)*, 2017. **8**(9).
266. Marti, T.M., C. Kunz, and O. Fleck, *DNA mismatch repair and mutation avoidance pathways*. *J Cell Physiol*, 2002. **191**(1): p. 28-41.
267. O'Donnell, M., L. Langston, and B. Stillman, *Principles and concepts of DNA replication in bacteria, archaea, and eukarya*. *Cold Spring Harb Perspect Biol*, 2013. **5**(7).
268. Palombo, F., et al., *GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells*. *Science (New York, N.Y.)*, 1995. **268**(5219): p. 1912-4.
269. Fishel, R., et al., *The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer*. *Cell*, 1993. **75**(5): p. 1027-38.
270. Kolodner, R.D. and G.T. Marsischky, *Eukaryotic DNA mismatch repair*. *Curr Opin Genet Dev*, 1999. **9**(1): p. 89-96.
271. Kunkel, T.A. and D.A. Erie, *DNA mismatch repair*. *Annual review of biochemistry*, 2005. **74**: p. 681-710.
272. Alani, E., et al., *Crystal structure and biochemical analysis of the MutS.ADP.beryllium fluoride complex suggests a conserved mechanism for ATP interactions in mismatch repair*. *J Biol Chem*, 2003. **278**(18): p. 16088-94.
273. Junop, M.S., et al., *Composite active site of an ABC ATPase: MutS uses ATP to verify mismatch recognition and authorize DNA repair*. *Mol Cell*, 2001. **7**(1): p. 1-12.
274. Lamers, M.H., H.H. Winterwerp, and T.K. Sixma, *The alternating ATPase domains of MutS control DNA mismatch repair*. *Embo j*, 2003. **22**(3): p. 746-56.
275. Obmolova, G., et al., *Crystal structures of mismatch repair protein MutS and its complex with a substrate DNA*. *Nature*, 2000. **407**(6805): p. 703-10.
276. Shell, S.S., C.D. Putnam, and R.D. Kolodner, *The N terminus of Saccharomyces cerevisiae Msh6 is an unstructured tether to PCNA*. *Mol Cell*, 2007. **26**(4): p. 565-78.
277. Johnson, R.E., et al., *Evidence for involvement of yeast proliferating cell nuclear antigen in DNA mismatch repair*. *J Biol Chem*, 1996. **271**(45): p. 27987-90.
278. Kadyrov, F.A., et al., *Endonucleolytic function of MutLalpha in human mismatch repair*. *Cell*, 2006. **126**(2): p. 297-308.
279. Zhang, Y., et al., *Reconstitution of 5'-directed human mismatch repair in a purified system*. *Cell*, 2005. **122**(5): p. 693-705.
280. McCulloch, S.D., L. Gu, and G.M. Li, *Bi-directional processing of DNA loops by mismatch repair-dependent and -independent pathways in human cells*. *J Biol Chem*, 2003. **278**(6): p. 3891-6.

References

281. Parsons, R., et al., *Hypermutability and mismatch repair deficiency in RER+ tumor cells*. Cell, 1993. **75**(6): p. 1227-1236.
282. Bhattacharyya, N.P., et al., *Mutator phenotypes in human colorectal carcinoma cell lines*. Proc Natl Acad Sci U S A, 1994. **91**(14): p. 6319-23.
283. Cortes-Ciriano, I., et al., *A molecular portrait of microsatellite instability across multiple cancers*. Nat Commun, 2017. **8**: p. 15180.
284. Lynch, H.T., et al., *Milestones of Lynch syndrome: 1895-2015*. Nat Rev Cancer, 2015. **15**(3): p. 181-94.
285. Leach, F.S., et al., *Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer*. Cell, 1993. **75**(6): p. 1215-25.
286. Nicolaides, N.C., et al., *Mutations of two PMS homologues in hereditary nonpolyposis colon cancer*. Nature, 1994. **371**(6492): p. 75-80.
287. Papadopoulos, N., et al., *Mutation of a mutL homolog in hereditary colon cancer*. Science, 1994. **263**(5153): p. 1625-9.
288. Bronner, C.E., et al., *Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer*. Nature, 1994. **368**(6468): p. 258-61.
289. Ligtenberg, M.J., et al., *Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1*. Nat Genet, 2009. **41**(1): p. 112-7.
290. Cunningham, J.M., et al., *The frequency of hereditary defective mismatch repair in a prospective series of unselected colorectal carcinomas*. Am J Hum Genet, 2001. **69**(4): p. 780-90.
291. Palomaki, G.E., et al., *EGAPP supplementary evidence review: DNA testing strategies aimed at reducing morbidity and mortality from Lynch syndrome*. Genet Med, 2009. **11**(1): p. 42-65.
292. Hampel, H., et al., *Feasibility of screening for Lynch syndrome among patients with colorectal cancer*. J Clin Oncol, 2008. **26**(35): p. 5783-8.
293. Bonadona, V., et al., *Cancer risks associated with germline mutations in MLH1, MSH2, and MSH6 genes in Lynch syndrome*. Jama, 2011. **305**(22): p. 2304-10.
294. Lin, K.M., et al., *Cumulative incidence of colorectal and extracolonic cancers in MLH1 and MSH2 mutation carriers of hereditary nonpolyposis colorectal cancer*. J Gastrointest Surg, 1998. **2**(1): p. 67-71.
295. Espenschied, C.R., et al., *Multigene Panel Testing Provides a New Perspective on Lynch Syndrome*. J Clin Oncol, 2017. **35**(22): p. 2568-2575.
296. Senter, L., et al., *The clinical phenotype of Lynch syndrome due to germ-line PMS2 mutations*. Gastroenterology, 2008. **135**(2): p. 419-28.
297. Hitchins, M.P., et al., *Dominantly inherited constitutional epigenetic silencing of MLH1 in a cancer-affected family is linked to a single nucleotide variant within the 5'UTR*. Cancer Cell, 2011. **20**(2): p. 200-13.
298. Veigl, M.L., et al., *Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers*. Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8698-702.
299. Gallinger, S., et al., *Gastrointestinal cancers and neurofibromatosis type 1 features in children with a germline homozygous MLH1 mutation*. Gastroenterology, 2004. **126**(2): p. 576-85.
300. Vasen, H.F.A., et al., *Guidelines for the clinical management of Lynch syndrome (hereditary non-polyposis cancer)*. Journal of Medical Genetics, 2007. **44**(6): p. 353-362.
301. Boland, C.R. and A. Goel, *Microsatellite Instability in Colorectal Cancer*. Gastroenterology, 2010. **138**(6): p. 2073-2087.e3.
302. Kane, M.F., et al., *Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines*. Cancer Res, 1997. **57**(5): p. 808-11.
303. Herman, J.G., et al., *Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma*. Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6870-5.
304. Baretti, M. and D.T. Le, *DNA mismatch repair in cancer*. Pharmacol Ther, 2018. **189**: p. 45-62.
305. Wölfel, T., et al., *A p16INK4a-insensitive CDK4 mutant targeted by cytolytic T lymphocytes in a human melanoma*. Science, 1995. **269**(5228): p. 1281-4.
306. Coulie, P.G., et al., *A mutated intron sequence codes for an antigenic peptide recognized by cytolytic T lymphocytes on a human melanoma*. Proc Natl Acad Sci U S A, 1995. **92**(17): p. 7976-80.
307. Lennerz, V., et al., *The response of autologous T cells to a human melanoma is dominated by mutated neoantigens*. Proc Natl Acad Sci U S A, 2005. **102**(44): p. 16013-8.
308. Rapoport, A.P., et al., *NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma*. Nat Med, 2015. **21**(8): p. 914-921.
309. Dieu-Nosjean, M.C., et al., *Long-term survival for patients with non-small-cell lung cancer with intratumoral lymphoid structures*. J Clin Oncol, 2008. **26**(27): p. 4410-7.
310. Dieu-Nosjean, M.C., et al., *Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers*. Immunol Rev, 2016. **271**(1): p. 260-75.
311. Schlitzer, A., N. McGovern, and F. Ginhoux, *Dendritic cells and monocyte-derived cells: Two complementary and integrated functional systems*. Semin Cell Dev Biol, 2015. **41**: p. 9-22.
312. Cabillie, F., et al., *Interleukin-6 and vascular endothelial growth factor release by renal cell carcinoma cells impedes lymphocyte-dendritic cell cross-talk*. Clin Exp Immunol, 2006. **146**(3): p. 518-23.
313. Carmi, Y., et al., *Allogeneic IgG combined with dendritic cell stimuli induce antitumour T-cell immunity*. Nature, 2015. **521**(7550): p. 99-104.
314. Moulin, V., et al., *B lymphocytes regulate dendritic cell (DC) function in vivo: increased interleukin 12 production by DCs from B cell-deficient mice results in T helper cell type 1 deviation*. J Exp Med, 2000. **192**(4): p. 475-82.
315. Nzula, S., J.J. Goings, and D.I. Stott, *Antigen-driven clonal proliferation, somatic hypermutation, and selection of B lymphocytes infiltrating human ductal breast carcinomas*. Cancer Res, 2003. **63**(12): p. 3275-80.
316. Rossetti, R.A.M., et al., *B lymphocytes can be activated to act as antigen presenting cells to promote anti-tumor responses*. PLoS One, 2018. **13**(7): p. e0199034.
317. Schumacher, T.N. and R.D. Schreiber, *Neoantigens in cancer immunotherapy*. Science, 2015. **348**(6230): p. 69-74.
318. Kim, S.T., et al., *Tumor-infiltrating lymphocytes, tumor characteristics, and recurrence in patients with early breast cancer*. Am J Clin Oncol, 2013. **36**(3): p. 224-31.
319. Park, M.H., J.S. Lee, and J.H. Yoon, *High expression of CX3CL1 by tumor cells correlates with a good prognosis and increased tumor-infiltrating CD8+ T cells, natural killer cells, and dendritic cells in breast carcinoma*. J Surg Oncol, 2012. **106**(4): p. 386-92.
320. Hamanishi, J., et al., *Programmed cell death 1 ligand 1 and tumor-infiltrating CD8+ T lymphocytes are prognostic factors of human ovarian cancer*. Proc Natl Acad Sci U S A, 2007. **104**(9): p. 3360-5.

References

321. Feichtenbeiner, A., et al., *Critical role of spatial interaction between CD8⁺ and Foxp3⁺ cells in human gastric cancer: the distance matters*. Cancer Immunol Immunother, 2014. **63**(2): p. 111-9.
322. Hamid, O., et al., *A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma*. J Transl Med, 2011. **9**: p. 204.
323. Tumeh, P.C., et al., *PD-1 blockade induces responses by inhibiting adaptive immune resistance*. Nature, 2014. **515**(7528): p. 568-71.
324. Gryfe, R., et al., *Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer*. N Engl J Med, 2000. **342**(2): p. 69-77.
325. Mouradov, D., et al., *Survival in stage II/III colorectal cancer is independently predicted by chromosomal and microsatellite instability, but not by specific driver mutations*. Am J Gastroenterol, 2013. **108**(11): p. 1785-93.
326. Guastadisegni, C., et al., *Microsatellite instability as a marker of prognosis and response to therapy: a meta-analysis of colorectal cancer survival data*. Eur J Cancer, 2010. **46**(15): p. 2788-98.
327. Lochhead, P., et al., *Microsatellite instability and BRAF mutation testing in colorectal cancer prognostication*. J Natl Cancer Inst, 2013. **105**(15): p. 1151-6.
328. Diaz-Padilla, I., et al., *Mismatch repair status and clinical outcome in endometrial cancer: a systematic review and meta-analysis*. Crit Rev Oncol Hematol, 2013. **88**(1): p. 154-67.
329. Stelloo, E., et al., *Refining prognosis and identifying targetable pathways for high-risk endometrial cancer; a TransPORTEC initiative*. Mod Pathol, 2015. **28**(6): p. 836-44.
330. Longley, D.B., D.P. Harkin, and P.G. Johnston, *5-fluorouracil: mechanisms of action and clinical strategies*. Nat Rev Cancer, 2003. **3**(5): p. 330-8.
331. Ribic, C.M., et al., *Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer*. N Engl J Med, 2003. **349**(3): p. 247-57.
332. Tejpar, S., et al., *Microsatellite instability, prognosis and drug sensitivity of stage II and III colorectal cancer: more complexity to the puzzle*, in J Natl Cancer Inst. 2011: United States. p. 841-4.
333. Hutchins, G., et al., *Value of mismatch repair, KRAS, and BRAF mutations in predicting recurrence and benefits from chemotherapy in colorectal cancer*. J Clin Oncol, 2011. **29**(10): p. 1261-70.
334. Resnick, K.E., et al., *Mismatch repair status and outcomes after adjuvant therapy in patients with surgically staged endometrial cancer*. Gynecol Oncol, 2010. **117**(2): p. 234-8.
335. Hellmann, M.D., et al., *Nivolumab plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer*. N Engl J Med, 2019. **381**(21): p. 2020-2031.
336. Larkin, J., et al., *Five-Year Survival with Combined Nivolumab and Ipilimumab in Advanced Melanoma*. N Engl J Med, 2019. **381**(16): p. 1535-1546.
337. Brahmer, J.R., et al., *Phase I study of single-agent anti-programmed death-1 (MDX-1106) in refractory solid tumors: safety, clinical activity, pharmacodynamics, and immunologic correlates*. Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2010. **28**(19): p. 3167-75.
338. Topalian, S.L., et al., *Safety, activity, and immune correlates of anti-PD-1 antibody in cancer*. The New England journal of medicine, 2012. **366**(26): p. 2443-54.
339. Le, D.T., et al., *KEYNOTE-164: Phase 2 study of pembrolizumab for patients with previously treated, microsatellite instability-high advanced colorectal carcinoma*. Journal of Clinical Oncology, 2016. **34**(15_suppl): p. TPS3631-TPS3631.
340. Schellens, J.H.M., et al., *Pembrolizumab for previously treated advanced cervical squamous cell cancer: Preliminary results from the phase 2 KEYNOTE-158 study*. Journal of Clinical Oncology, 2017. **35**(15_suppl): p. 5514-5514.
341. Seiwert, T.Y., et al., *Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial*. The Lancet Oncology, 2016. **17**(7): p. 956-965.
342. US Food and Drug Administration, F. *FDA grants accelerated approval to pembrolizumab for first tissue/site agnostic indication*. 2017 [cited 2021 4th September 2021]; Available from: <https://www.fda.gov/drugs/resources-information-approved-drugs/fda-grants-accelerated-approval-pembrolizumab-first-tissuesite-agnostic-indication>.
343. Network, T.C.G.A., *Integrated genomic characterization of endometrial carcinoma*. Nature, 2013. **497**(7447): p. 67-73.
344. Network, T.C.G.A., *Comprehensive molecular characterization of human colon and rectal cancer*. Nature, 2012. **487**(7407): p. 330-337.
345. Thompson, B.A., et al., *Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database*. Nat Genet, 2014. **46**(2): p. 107-115.
346. Thorsson, V., et al., *The Immune Landscape of Cancer*. Immunity, 2018. **48**(4): p. 812-830.e14.
347. Colaprico, A., et al., *TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data*. Nucleic Acids Res, 2016. **44**(8): p. e71.
348. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2*. Genome Biol, 2014. **15**(12): p. 550.
349. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
350. Schwartz, G., *Meta-Analysis with R*. Use R! 2015: Springer International Publishing.
351. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis*. 2016: Springer-Verlag New York.
352. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters*. Omics, 2012. **16**(5): p. 284-7.
353. Mayakonda, A., et al., *Maftools: efficient and comprehensive analysis of somatic variants in cancer*. Genome Res, 2018. **28**(11): p. 1747-1756.
354. Sinicrope, F.A., et al., *DNA mismatch repair status and colon cancer recurrence and survival in clinical trials of 5-fluorouracil-based adjuvant therapy*. J Natl Cancer Inst, 2011. **103**(11): p. 863-75.
355. Bertagnolli, M.M., et al., *Microsatellite instability and loss of heterozygosity at chromosomal location 18q: prospective evaluation of biomarkers for stages II and III colon cancer--a study of CALGB 9581 and 89803*. J Clin Oncol, 2011. **29**(23): p. 3153-62.
356. Taieb, J., et al., *Prognostic Effect of BRAF and KRAS Mutations in Patients With Stage III Colon Cancer Treated With Leucovorin, Fluorouracil, and Oxaliplatin With or Without Cetuximab: A Post Hoc Analysis of the PETACC-8 Trial*. JAMA Oncol, 2016. **2**(5): p. 643-653.
357. de Weger, V.A., et al., *Clinical effects of adjuvant active specific immunotherapy differ between patients with microsatellite-stable and microsatellite-unstable colon cancer*. Clin Cancer Res, 2012. **18**(3): p. 882-9.

References

358. Niedzwiecki, D., et al., *Association Between Results of a Gene Expression Signature Assay and Recurrence-Free Interval in Patients With Stage II Colon Cancer in Cancer and Leukemia Group B 9581 (Alliance)*. J Clin Oncol, 2016. **34**(25): p. 3047-53.
359. Westra, J.L., et al., *Determination of TP53 mutation is more relevant than microsatellite instability status for the prediction of disease-free survival in adjuvant-treated stage III colon cancer patients*. J Clin Oncol, 2005. **23**(24): p. 5635-43.
360. Sasaki, Y., et al., *Prognostic and predictive value of extended RAS mutation and mismatch repair status in stage III colorectal cancer*. Cancer Sci, 2016. **107**(7): p. 1006-12.
361. Stelloo, E., et al., *Improved Risk Assessment by Integrating Molecular and Clinicopathological Factors in Early-stage Endometrial Cancer-Combined Analysis of the PORTEC Cohorts*. Clin Cancer Res, 2016. **22**(16): p. 4215-24.
362. Cosgrove, C.M., et al., *An NRG Oncology/GOG study of molecular classification for risk prediction in endometrioid endometrial cancer*. Gynecol Oncol, 2018. **148**(1): p. 174-180.
363. Helen, J.M., et al., *Prognostic value of microsatellite instability (MSI) and PTEN expression in women with endometrial cancer: Results from studies of the NCIC Clinical Trials Group (NCIC CTG)*. European Journal of Cancer, 2010. **46**(8): p. 1365-1373.
364. Suemori, T., et al., *Intratumoral CD8+ Lymphocyte Infiltration as a Prognostic Factor and Its Relationship With Cyclooxygenase 2 Expression and Microsatellite Instability in Endometrial Cancer*. Int J Gynecol Cancer, 2015. **25**(7): p. 1165-72.
365. Pakish, J.B., et al., *Immune microenvironment in microsatellite instable endometrial cancers: Hereditary or sporadic origin matters*. Clin Cancer Res, 2017.
366. Newman, A.M., et al., *Robust enumeration of cell subsets from tissue expression profiles*. Nat Methods, 2015. **12**(5): p. 453-7.
367. Novetsky, A.P., et al., *Frequent mutations in the RPL22 gene and its clinical and functional implications*. Gynecol Oncol, 2013. **128**(3): p. 470-4.
368. Wei, J., et al., *Ribosomal Proteins Regulate MHC Class I Peptide Generation for Immunosurveillance*. Mol Cell, 2019. **73**(6): p. 1162-1173.e5.
369. Thomas, S.J., et al., *The role of JAK/STAT signalling in the pathogenesis, prognosis and treatment of solid tumours*. Br J Cancer, 2015. **113**(3): p. 365-71.
370. Dowty, M.E., et al., *Preclinical to Clinical Translation of Tofacitinib, a Janus Kinase Inhibitor, in Rheumatoid Arthritis*. Journal of Pharmacology and Experimental Therapeutics, 2014. **348**(1): p. 165-173.
371. Wolf, D.M., et al., *Gene co-expression modules as clinically relevant hallmarks of breast cancer diversity*. PLoS One, 2014. **9**(2): p. e88309.
372. Curtsinger, J.M., et al., *Autocrine IFN- γ promotes naive CD8 T cell differentiation and synergizes with IFN- α to stimulate strong function*. J Immunol, 2012. **189**(2): p. 659-68.
373. Maraskovsky, E., W.F. Chen, and K. Shortman, *IL-2 and IFN-gamma are two necessary lymphokines in the development of cytolytic T cells*. J Immunol, 1989. **143**(4): p. 1210-4.
374. Akbar, S.M., K. Inaba, and M. Onji, *Upregulation of MHC class II antigen on dendritic cells from hepatitis B virus transgenic mice by interferon-gamma: abrogation of immune response defect to a T-cell-dependent antigen*. Immunology, 1996. **87**(4): p. 519-27.
375. Haabeth, O.A., et al., *Inflammation driven by tumour-specific Th1 cells protects against B-cell cancer*. Nat Commun, 2011. **2**: p. 240.
376. Walter, W., et al., *MHC class II antigen presentation pathway in murine tumours: tumour evasion from immunosurveillance?* Br J Cancer, 2000. **83**(9): p. 1192-201.
377. Simpson, J.A.D., et al., *Intratumoral T cell infiltration, MHC class I and STAT1 as biomarkers of good prognosis in colorectal cancer*. Gut, 2010. **59**(7): p. 926-33.
378. Ayers, M., et al., *IFN- γ -related mRNA profile predicts clinical response to PD-1 blockade*. J Clin Invest, 2017. **127**(8): p. 2930-2940.
379. Karachaliou, N., et al., *Interferon gamma, an important marker of response to immune checkpoint blockade in non-small cell lung cancer and melanoma patients*. Ther Adv Med Oncol, 2018. **10**: p. 1758834017749748.
380. Mo, X., et al., *Interferon- γ Signaling in Melanocytes and Melanoma Cells Regulates Expression of CTLA-4*. Cancer Res, 2018. **78**(2): p. 436-450.
381. Shin, D.S., et al., *Primary Resistance to PD-1 Blockade Mediated by JAK1/2 Mutations*. Cancer Discov, 2017. **7**(2): p. 188-201.
382. Zaretsky, J.M., et al., *Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma*. N Engl J Med, 2016. **375**(9): p. 819-29.
383. Cao, Y., et al., *Immune checkpoint molecules in natural killer cells as potential targets for cancer immunotherapy*. Signal Transduct Target Ther, 2020. **5**(1): p. 250.
384. Sharonov, G.V., et al., *B cells, plasma cells and antibody repertoires in the tumour microenvironment*. Nat Rev Immunol, 2020. **20**(5): p. 294-307.
385. Cheon, D.J. and S. Orsulic, *Mouse models of cancer*. Annu Rev Pathol, 2011. **6**: p. 95-119.
386. Frese, K.K. and D.A. Tuveson, *Maximizing mouse cancer models*. Nat Rev Cancer, 2007. **7**(9): p. 645-58.
387. Becher, O.J. and E.C. Holland, *Genetically engineered models have advantages over xenografts for preclinical studies*. Cancer Res, 2006. **66**(7): p. 3355-8, discussion 3358-9.
388. Sikder, H., et al., *Disruption of Id1 reveals major differences in angiogenesis between transplanted and autochthonous tumors*. Cancer Cell, 2003. **4**(4): p. 291-9.
389. Stewart, T.A., P.K. Pattengale, and P. Leder, *Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes*. Cell, 1984. **38**(3): p. 627-37.
390. Brinster, R.L., et al., *Transgenic mice harboring SV40 T-antigen genes develop characteristic brain tumors*. Cell, 1984. **37**(2): p. 367-79.
391. Jacks, T., et al., *Effects of an Rb mutation in the mouse*. Nature, 1992. **359**(6393): p. 295-300.
392. Donehower, L.A., et al., *Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours*. Nature, 1992. **356**(6366): p. 215-21.
393. Gowen, L.C., et al., *Brcal deficiency results in early embryonic lethality characterized by neuroepithelial abnormalities*. Nat Genet, 1996. **12**(2): p. 191-4.
394. Hakem, R., et al., *The tumor suppressor gene Brca1 is required for embryonic cellular proliferation in the mouse*. Cell, 1996. **85**(7): p. 1009-23.
395. Bouabe, H. and K. Okkenhaug, *Gene targeting in mice: a review*. Methods Mol Biol, 2013. **1064**: p. 315-36.
396. Hoess, R.H., A. Wierzbicki, and K. Abremski, *The role of the loxP spacer region in P1 site-specific recombination*. Nucleic Acids Res, 1986. **14**(5): p. 2287-300.

References

397. Gierut, J.J., T.E. Jacks, and K.M. Haigis, *Strategies to achieve conditional gene mutation in mice*. Cold Spring Harb Protoc, 2014. **2014**(4): p. 339-49.
398. Feil, S., N. Valtcheva, and R. Feil, *Inducible Cre mice*. Methods Mol Biol, 2009. **530**: p. 343-63.
399. Indra, A.K., et al., *Temporally-controlled site-specific mutagenesis in the basal layer of the epidermis: comparison of the recombinase activity of the tamoxifen-inducible Cre-ER(T) and Cre-ER(T2) recombinases*. Nucleic Acids Res, 1999. **27**(22): p. 4324-7.
400. Schwenk, F., et al., *Temporally and spatially regulated somatic mutagenesis in mice*. Nucleic Acids Res, 1998. **26**(6): p. 1427-32.
401. Feldmann, G., et al., *Hedgehog inhibition prolongs survival in a genetically engineered mouse model of pancreatic cancer*. Gut, 2008. **57**(10): p. 1420-30.
402. Romer, J.T., et al., *Suppression of the Shh pathway using a small molecule inhibitor eliminates medulloblastoma in Ptc1(+/-) p53(-/-) mice*. Cancer Cell, 2004. **6**(3): p. 229-40.
403. Zeng, L., et al., *The mouse Fused locus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation*. Cell, 1997. **90**(1): p. 181-92.
404. Orsulic, S., et al., *Induction of ovarian cancer by defined multiple genetic changes in a mouse model system*. Cancer Cell, 2002. **1**(1): p. 53-62.
405. Xu, X., et al., *Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis*. Nat Genet, 2001. **28**(3): p. 266-71.
406. Ramus, S.J., et al., *Increased frequency of TP53 mutations in BRCA1 and BRCA2 ovarian tumours*. Genes Chromosomes Cancer, 1999. **25**(2): p. 91-6.
407. Moser, A.R., H.C. Pitot, and W.F. Dove, *A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse*. Science, 1990. **247**(4940): p. 322-4.
408. Oshima, M., et al., *Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene*. Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4482-6.
409. Fodde, R., et al., *A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors*. Proc Natl Acad Sci U S A, 1994. **91**(19): p. 8969-73.
410. Hinoi, T., et al., *Mouse model of colonic adenoma-carcinoma progression based on somatic Apc inactivation*. Cancer Res, 2007. **67**(20): p. 9721-30.
411. Nambiar, P.R., et al., *Preliminary analysis of azoxymethane induced colon tumors in inbred mice commonly used as transgenic/knockout progenitors*. Int J Oncol, 2003. **22**(1): p. 145-50.
412. Paulsen, J.E., et al., *Age-dependent susceptibility to azoxymethane-induced and spontaneous tumorigenesis in the Min/+ mouse*. Anticancer Res, 2003. **23**(1a): p. 259-65.
413. Suzui, M., et al., *Enhanced colon carcinogenesis induced by azoxymethane in min mice occurs via a mechanism independent of beta-catenin mutation*. Cancer Lett, 2002. **183**(1): p. 31-41.
414. Sussman, D.A., et al., *Cancer in inflammatory bowel disease: lessons from animal models*. Curr Opin Gastroenterol, 2012. **28**(4): p. 327-33.
415. Neufert, C., et al., *Inducible mouse models of colon cancer for the analysis of sporadic and inflammation-driven tumor progression and lymph node metastasis*. Nat Protoc, 2021. **16**(1): p. 61-85.
416. Colnot, S., et al., *Colorectal cancers in a new mouse model of familial adenomatous polyposis: influence of genetic and environmental modifiers*. Lab Invest, 2004. **84**(12): p. 1619-30.
417. Crist, R.C., et al., *The armadillo repeat domain of Apc suppresses intestinal tumorigenesis*. Mamm Genome, 2010. **21**(9-10): p. 450-7.
418. Clarke, A.R., M.C. Cummings, and D.J. Harrison, *Interaction between murine germline mutations in p53 and APC predisposes to pancreatic neoplasia but not to increased intestinal malignancy*. Oncogene, 1995. **11**(9): p. 1913-20.
419. Shankaran, V., et al., *IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity*. Nature, 2001. **410**(6832): p. 1107-11.
420. Smyth, M.J., et al., *Perforin-mediated cytotoxicity is critical for surveillance of spontaneous lymphoma*. J Exp Med, 2000. **192**(5): p. 755-60.
421. Street, S.E., et al., *Suppression of lymphoma and epithelial malignancies effected by interferon gamma*. J Exp Med, 2002. **196**(1): p. 129-34.
422. Zerafa, N., et al., *Cutting edge: TRAIL deficiency accelerates hematological malignancies*. J Immunol, 2005. **175**(9): p. 5586-90.
423. Street, S.E., et al., *Innate immune surveillance of spontaneous B cell lymphomas by natural killer cells and gammadelta T cells*. J Exp Med, 2004. **199**(6): p. 879-84.
424. Kaplan, D.H., et al., *Demonstration of an interferon gamma-dependent tumor surveillance system in immunocompetent mice*. Proc Natl Acad Sci U S A, 1998. **95**(13): p. 7556-61.
425. Bolitho, P., et al., *Perforin-mediated suppression of B-cell lymphoma*. Proc Natl Acad Sci U S A, 2009. **106**(8): p. 2723-8.
426. Leach, D.R., M.F. Krummel, and J.P. Allison, *Enhancement of antitumor immunity by CTLA-4 blockade*. Science, 1996. **271**(5256): p. 1734-6.
427. Iwai, Y., S. Terawaki, and T. Honjo, *PD-1 blockade inhibits hematogenous spread of poorly immunogenic tumor cells by enhanced recruitment of effector T cells*. Int Immunol, 2005. **17**(2): p. 133-44.
428. Lee, K., E. Tosti, and W. Edelmann, *Mouse models of DNA mismatch repair in cancer research*. DNA Repair (Amst), 2016. **38**: p. 140-146.
429. Reitmaier, A.H., et al., *MSH2 deficient mice are viable and susceptible to lymphoid tumours*. Nat Genet, 1995. **11**(1): p. 64-70.
430. Prolla, T.A., et al., *Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair*. Nat Genet, 1998. **18**(3): p. 276-9.
431. Kucherlapati, M.H., et al., *An Msh2 conditional knockout mouse for studying intestinal cancer and testing anticancer agents*. Gastroenterology, 2010. **138**(3): p. 993-1002.e1.
432. Wojciechowicz, K., et al., *Temozolomide increases the number of mismatch repair-deficient intestinal crypts and accelerates tumorigenesis in a mouse model of Lynch syndrome*. Gastroenterology, 2014. **147**(5): p. 1064-72.e5.
433. Edelmann, W., et al., *The DNA Mismatch Repair Genes Msh3 and Msh6 Cooperate in Intestinal Tumor Suppression*. Cancer Res., 2000. **60**(4): p. 803-807.
434. de Wind, N., et al., *Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer*. Cell, 1995. **82**(2): p. 321-330.
435. Smits, R., et al., *Somatic Apc mutations are selected upon their capacity to inactivate the beta-catenin downregulating activity*. Genes Chromosomes Cancer, 2000. **29**(3): p. 229-39.

References

436. Edelmann, W., et al., *The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression*. Cancer research, 2000. **60**(4): p. 803-7.
437. Baker, S.M., et al., *Male mice defective in the DNA mismatch repair gene PMS2 exhibit abnormal chromosome synapsis in meiosis*. Cell, 1995. **82**(2): p. 309-19.
438. Avdievich, E., et al., *Distinct effects of the recurrent Mlh1G67R mutation on MMR functions, cancer, and meiosis*. Proc Natl Acad Sci U S A, 2008. **105**(11): p. 4247-52.
439. van Oers, J.M., et al., *PMS2 endonuclease activity has distinct biological functions and is essential for genome maintenance*. Proc Natl Acad Sci U S A, 2010. **107**(30): p. 13384-9.
440. Yang, G., et al., *Dominant effects of an Msh6 missense mutation on DNA repair and cancer susceptibility*. Cancer Cell, 2004. **6**(2): p. 139-50.
441. Lin, D.P., et al., *An Msh2 point mutation uncouples DNA mismatch repair and apoptosis*. Cancer Res, 2004. **64**(2): p. 517-22.
442. Yao, X., et al., *Different mutator phenotypes in Mlh1- versus Pms2-deficient mice*. Proc Natl Acad Sci U S A, 1999. **96**(12): p. 6850-5.
443. Jahn, H.M., et al., *Refined protocols of tamoxifen injection for inducible DNA recombination in mouse astroglia*. Sci Rep, 2018. **8**(1): p. 5913.
444. Williams, J.M., et al., *Intestinal Preparation Techniques for Histological Analysis in the Mouse*. Curr Protoc Mouse Biol, 2016. **6**(2): p. 148-168.
445. Maitra, R., et al., *Development and Characterization of a Genetic Mouse Model of KRAS Mutated Colorectal Cancer*. Int J Mol Sci, 2019. **20**(22).
446. el Marjou, F., et al., *Tissue-specific and inducible Cre-mediated recombination in the gut epithelium*. Genesis, 2004. **39**(3): p. 186-93.
447. Bankhead, P., et al., *QuPath: Open source software for digital pathology image analysis*. Sci Rep, 2017. **7**(1): p. 16878.
448. Luke, J.J., et al., *WNT/ β -catenin Pathway Activation Correlates with Immune Exclusion across Human Cancers*. Clin Cancer Res, 2019. **25**(10): p. 3074-3083.
449. Zander, R., et al., *CD4(+) T Cell Help Is Required for the Formation of a Cytolytic CD8(+) T Cell Subset that Protects against Chronic Infection and Cancer*. Immunity, 2019. **51**(6): p. 1028-1042.e4.
450. Alspach, E., et al., *MHC-II neoantigens shape tumour immunity and response to immunotherapy*. Nature, 2019. **574**(7780): p. 696-701.
451. Zhang, L., et al., *Lineage tracking reveals dynamic relationships of T cells in colorectal cancer*. Nature, 2018. **564**(7735): p. 268-272.
452. Ambagala, A.P., J.C. Solheim, and S. Srikumaran, *Viral interference with MHC class I antigen presentation pathway: the battle continues*. Vet Immunol Immunopathol, 2005. **107**(1-2): p. 1-15.
453. Kloor, M., et al., *Beta2-microglobulin mutations in microsatellite unstable colorectal tumors*. Int J Cancer, 2007. **121**(2): p. 454-8.
454. Germano, G., et al., *CD4 T Cell-Dependent Rejection of Beta-2 Microglobulin Null Mismatch Repair-Deficient Tumors*. Cancer Discov, 2021. **11**(7): p. 1844-1859.
455. Middha, S., et al., *Majority of B2M-Mutant and -Deficient Colorectal Carcinomas Achieve Clinical Benefit From Immune Checkpoint Inhibitor Therapy and Are Microsatellite Instability-High*. JCO Precis Oncol, 2019. **3**.
456. Coronella, J.A., et al., *Antigen-driven oligoclonal expansion of tumor-infiltrating B cells in infiltrating ductal carcinoma of the breast*. J Immunol, 2002. **169**(4): p. 1829-36.
457. Iglesia, M.D., et al., *Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer*. Clin Cancer Res, 2014. **20**(14): p. 3818-29.
458. Dunn, G.P., C.M. Koebel, and R.D. Schreiber, *Interferons, immunity and cancer immunoediting*. Nature Reviews Immunology, 2006. **6**(11): p. 836.