Characterising peripheral responses to immune checkpoint inhibitors



Rosalin Cooper

Department of Oncology

University of Oxford

Wolfson College

Trinity Term 2021

Thesis submitted for the degree of Doctor of Philosophy

Abstract

The introduction of immune checkpoint blockade (ICB) therapy in metastatic MM has dramatically improved survival in this patient group. However, only some patients experience durable clinical response, and many will develop immune-related adverse events (irAEs). The identification of predictive biomarkers which can be assessed in a biopsy-free manner may facilitate early detection of non-responders and optimal targeting of treatment.

There is emerging evidence to show that the circulating monocyte population has a prognostic and predictive role in the context of ICB, but this has yet to be fully characterised. Circulating-cell free DNA (cfDNA) is derived predominantly from leukocytes, but in patients with cancer may comprise circulating tumour DNA (ctDNA). Assessment of epigenetic cfDNA profiles may provide tumour profiling and reflect immune responses as a 'liquid biopsy'. Although epigenetic tumour profiles have diagnostic and prognostic utility across multiple cancer types, epigenetic cfDNA profiles in MM and following ICB are not yet characterised.

In this thesis I dissect both circulating myeloid and epigenetic cfDNA profiles in MM patients. Firstly, I explore MM-associated monocyte transcriptomic profiles with bulk RNA sequencing (RNA-seq), characterising on-treatment modulation and association with clinical response. I then use scRNA-seq to explore the transcriptional heterogeneity of the peripheral monocyte compartment, and describe monocyte subset-wise responses to ICB. Lastly, I characterise epigenetic cfDNA profiles in this cohort, revealing distinct 5-hydroxymethylcytosine (5hmC) and 5-methylcytosine (5mC) cfDNA profiles in MM and following ICB.

In this thesis I present novel findings regarding peripheral myeloid and epigenetic cfDNA responses to ICB and demonstrate that these profiles may have potential diagnostic and predictive clinical utility.

Declaration

I declare that the work presented in this thesis is my own, except where explicitly stated and acknowledged (see below). I confirm that no part of this thesis has been submitted for another degree within this University or externally.

Sample collection and processing was carried out by myself, Ms Chelsea A Taylor, Dr Piyush Kumar Sharma, Mr Orion Tong, Dr Robert Watson, Dr Elise Mahé and Ms Sara Danielli. The scRNA-seq library preparation was carried out by myself, Dr Robert Watson (patient samples) and Ms Chelsea Taylor (healthy control samples) at the MRC Weatherall Institute of Molecular Medicine (WIMM), University of Oxford. I carried out 5hmC library preparation and sequencing both in collaboration with the Song group at the Target Discovery Institute, University of Oxford, under the guidance of Dr Paulina Siejka-Zielinska, and in-house at the WIMM. I carried out the 5mC library preparation with Ms Alba Verge de los Aires. Flow cytometry was carried out by Ms Chelsea Taylor at the WIMM. Extraction of RNA from PBMCs was performed by Ms Alba Verge de los Aires, Ms Chelsea Taylor and Dr Elise Mahé. Upstream processing of data was performed by Dr Isar Nassiri. I performed all downstream analysis as outlined in Chapters Two to Six.

Acknowledgments

Above all, I would like to thank the patients who consented to be involved in this work; without them this work would not be possible. Thank you also to CRUK, who have provided me with the privilege of time and resources to enable completion of this work. Firstly, I would like to thank my supervisors, Dr Benjamin Fairfax and Prof Mark Middleton for their support and guidance with this project. I would like to thank all current and past members of the Fairfax lab group, for lively debate, comments and advice during laboratory meetings or informal discussions. In particular, I want to thank Dr Isar Nassiri, whose bioinformatic expertise has been of much help during this project. Also, I thank Alba Verge de los Aires for her skillful benchwork help and expertise. Thank you to also Chelsea Taylor, who has been a huge help with experimental troubleshooting, a sounding board for scientific ideas and a fountain of immunological knowledge. Thanks go to Dr Piyush Kumar Sharma for sound advice and support and to Orion Tong, for advice regarding bioinformatic approaches.

I would like to thank my family – my Mum, Dad and sisters Anna and Fiona - for their ongoing support and encouragement over many years. I would also like to thank Jimmy, who believed that this project would be completed almost from the start, even when the experimental and bioinformatic challenges at times seemed vast.

Contents

1 Chapter 1: Introduction 17

1.1 Melanoma and immune checkpoint inhibition 17

1.2 Predictive biomarkers for immune checkpoint blockade 20

1.2.1 Tumour infiltrating lymphocytes (TILs) 20

1.2.2 PD-L1 expression 20

1.2.3 Tumour mutational burden and mismatch repair deficiency 21

1.2.4 Summary of existing predictive markers and the case for exploring the peripheral immune compartment 21

1.3 Monocytes 22

1.3.1 Characterising monocyte subsets 24

1.3.1.1 Classical monocytes 24

1.3.1.2 Intermediate monocytes 24

1.3.1.3 Non-classical monocytes 25

1.3.1.4 Mass-spectrometry approaches to characterise monocyte subsets 27

1.3.1.5 Characterisation of monocyte subsets using dissection at a single-cell level 28

1.3.2 Summary of monocyte subset phenotyping 29

1.3.3 Monocyte differentiation, maturation and fate 30

1.4 Myeloid populations in cancer 32

1.4.1 Macrophages and the tumour microenvironment 32

1.4.2 Myeloid signatures are prognostic in cancer 33

1.4.3 Reprogramming of the myeloid compartment in cancer 33

1.4.4 Myeloid populations and ICB 36

1.4.4.1 scRNA-seq approaches to characterising the myeloid response to ICB 38

1.4.5 Summary 39

1.5 Cell-free DNA 40

1.5.1 Limitations of cfDNA assessment 41

1.6 Methylation 41

1.6.1 Methylation and cancer 42

1.6.2 Methylation cfDNA profiles 43

1.7 5-hydroxymethylcytosine 45

1.7.1 5hmC distribution and function 45

1.7.2 5hmC in cancer 47
1.7.3 5hmC cfDNA profiles 47
1.8 Epigenetic sequencing assays 50
1.8.1 Summary 51
1.9 Conclusions 52
1.10 Key aims 52
2 Chapter 2: Methods 54
2.1 Samples 54
2.1.1 Patient cohort 54
2.1.2 Control samples 54
2.1.3 Clinical outcome data 54
2.2 Peripheral blood mononuclear cell (PBMC) and plasma isolation 55
2.3 Bulk RNA-Sequencing 56
2.3.1 RNA Extraction 56
2.3.2 RNA sequencing 56
2.3.3 Upstream processing of bulk RNA-seq data 57
2.3.3.1 Differential expression analysis 57
2.3.3.2 Pathway enrichment analysis of differentially expressed transcripts 57
2.3.3.3 Building a diagnostic classifier from monocyte transcriptomic profiles 58
2.4 Flow cytometry 58
2.5 Single cell RNA sequencing 58
2.5.1 Patient cohort and scRNA-seq 58
2.5.2 scRNA-Seq library preparation 59
2.5.3 Sequencing 59
2.5.4 Data processing 59
2.5.5 Downstream scRNA-seq analysis 60
2.5.5.1 Quality control 60
2.5.6 Clustering analysis 60
2.5.6.1 Annotation of monocyte clusters 61
2.5.6.2 Identification of cluster-defining markers 61
2.5.6.3 Trajectory analysis 61
2.5.6.4 Differential expression analysis 62
2.5.6.5 Pathway analysis 62
2.5.6.6 Gene module scoring 62

2.5.6.7 Subset scoring in bulk RNA-seq data 62
2.5.6.8 Survival analysis 63
2.6 cfDNA extraction 63
2.7 5hmC cfDNA sequencing 63
2.7.1 5hmC cohort details 63
2.7.2 5hmC cfDNA library preparation 64
2.7.2.1 Addition of a spike-in oligonucleotide sequence for validation of 5hmC pull-down 64
2.7.2.2 End-repair and A tailing 65
2.7.2.3 Adaptor ligation 66
2.7.2.4 5hmC labelling 66
2.7.2.5 Pull-down on streptavidin beads 67
2.7.2.6 Library amplification 67
2.7.2.7 Quantification and QC 68
2.7.3 Sequencing 68
2.7.4 5hmC enrichment validation 69
2.8 5-methylcytosine cfDNA sequencing 69
2.8.1 5mC cohort details 69
2.8.2 cfMeDIP library preparation 69
2.8.3 Library quantification and sequencing 70
2.9 Upstream processing of epigenetic datasets 71
2.10 Downstream analysis of epigenetic cfDNA datasets 71
2.10.1 Identification of epigenetically enriched regions 71
2.10.2 Visualisation of mapped reads 71
2.10.3 Visualisation of epigenetically enriched regions 71
2.10.4 Annotation of epigenetic cfDNA peaks 72
2.10.5 Assessing overlap between epigenetic cfDNA peaksets 72
2.10.6 Assessing enrichment across functional genomic regions and histone modifications . 72
2.10.7 Differential enrichment analysis 73
2.10.8 Pathway enrichment analysis of epigenetic data 73
2.10.9 Exploring association between epigenetic cfDNA modification and transcriptomic expression in circulating immune cell 73
2.11 General statistical analysis 73
3 Chapter 3: Characterising transcriptional profiles in peripheral monocytes in response to immune checkpoint blockade 74

3.1 Background 74

3.2 Aims 76

3.3 Results 76

3.3.1 Distinct transcriptional profiles are seen in peripheral monocytes in melanoma 76

3.3.2 Peripheral monocyte count is associated with transcriptomic monocyte profiles 82

3.3.3 Exploring the utility of monocyte expression profiles as a diagnostic classifier 85

3.3.4 Modulation of transcriptomic profiles in peripheral monocytes with ICB 86

3.3.4.1 Comparing responses to combined ICB (cICB) and single-agent ICB (sICB) 91

3.3.5 ICB is associated with sustained transcriptional modulation in peripheral monocytes 94

3.3.6 Transcriptomic monocyte profiles are associated with clinical response to ICB 96

3.3.7 Exploring peripheral monocyte subset proportions with flow cytometry and clinical outcome 100

3.3.8 Peripheral monocyte subset size is associated with monocyte transcriptomic expression 103

3.4 Summary and Conclusions 105

4 Chapter 4: Characterising peripheral monocyte response to Immune Checkpoint Blockade with scRNA-seq 108

4.1 Background 108

4.2 Aims and Objectives 109

4.3 Results 109

4.3.1 Characterising monocyte subsets and phenotypes using scRNA-seq 109

4.3.2 Annotation of monocyte subsets identifies conventional monocyte clusters 110

4.3.2.1 Visualising clustering trees 111

4.3.2.2 Characterising gene expression of key transcripts across the scRNA-seq dataset. 113

4.3.2.3 Annotation with annotated gene expression sets 114

4.3.2.4 Summary of annotation to identify conventional clusters 116

4.3.3 Characterising monocyte subsets 119

4.3.3.1 Expression profiling peripheral monocytes clusters 119

4.3.3.2 Characterising monocyte clusters with gene expression-correlated module scoring 124

4.3.3.3 Cell cycle annotation and activation status of monocyte clusters 126

4.3.4 Monocyte differentiation and fate explored with trajectory analysis 127

4.3.5 Dissecting monocyte subsets in MM patients versus healthy controls 131

4.3.5.1 Module scoring characterises MM-associated peripheral monocyte profiles 135

4.3.5.2 Trajectory analysis demonstrates differences in maturation states in MM and healthy controls 136
4.3.6 Dissecting differential monocyte subset responses to ICB 137
4.3.6.1 Differential gene expression in monocyte subsets with ICB 137
4.3.6.2 Module expression scoring characterises ICB responses in monocyte subsets 141
4.3.6.3 Trajectory analysis explores maturation in classical monocytes with ICB 143
4.3.7 Monocyte subset proportions 145
4.3.8 Analysis of classical monocytes demonstrates transcriptional and phenotypic heterogeneity 147
4.3.8.1 Charactering heterogeneity within the classical monocyte compartment 148
4.3.8.2 Trajectory analysis demonstrates varying maturation states across the peripheral classical monocyte compartment 155
4.3.8.3 Classical monocyte subsets demonstrate varying responses to ICB 158
4.3.8.4 Classical monocyte subset expression profiles and association with clinical outcome 160
4.3.9 Analysis of $CD16^{+}$ monocytes identifies an intermediate monocyte cluster 162
4.3.9.1 Pseudotime analysis demonstrates dual termination states within the $CD16^{+}$ compartment 171
4.4 Summary and Conclusions 174
5 Chapter 5: Characterising circulating epigenetic cell-free DNA profiles in patients with metastatic melanoma treated with immune checkpoint blockade 179
5.1 Background 179
5.2 Aims and Objectives 182
5.3 Results 182
5.3.1 Genome-wide distribution of epigenetic cfDNA profiles 183
5.3.2 Comparing the genome-wide distribution of 5hmC and 5mC modifications within cfDNA 186
5.3.3 Enrichment of epigenetic modifications across functional regions in cfDNA 191
5.3.4 Enrichment of epigenetic cfDNA modifications across functional regions in metastatic melanoma and with ICB 195
5.3.5 Epigenetic modifications in cfDNA at histone modification sites 198
5.3.6 Characterising differential enrichment of 5hmC and 5mC cfDNA modifications in metastatic melanoma 205
5.3.6.1 5-hydroxymethylcytosine cfDNA profiles in MM 205
5.3.6.2 5-methylcytosine cfDNA profiles in MM 206
5.3.7 Characterising differential enrichment of 5hmC and 5mC modifications in ICB 208

5.3.7.1 Modulation of 5-hydroxymethylcytosine cfDNA profiles with ICB 208
5.3.7.2 Modulation of 5-methylcytosine cfDNA profiles with ICB 209
5.3.7.3 Comparing 5mC cfDNA modulation with combination versus single agent immune checkpoint blockade 214
5.3.8 Characterising association with epigenetic cfDNA profiles with clinical outcome 220
5.3.8.1 5-hydroxymethylcytosine cfDNA profiles and clinical outcome 220
5.3.8.2 5-methylcytosine cfDNA profiles and clinical outcome 227
5.3.9 Exploring association between epigenetic cfDNA profiles and circulating immune expression profiles 232
5.4 Discussion and Conclusions 236
6 Discussion 240

List of Figures

Figure 1: Human Ig Fc gamma receptors (FcγRs) 23

Figure 2: Summary of key monocyte subset functions 26

Figure 3: Summary of key monocyte gene markers 27

Figure 4: Summary of key cancer-associated circulating monocyte profiles 36

Figure 5: Summary of key publications exploring epigenetic cfDNA profiles as cancer biomarkers 49

Figure 6:Table to show antibodies used in flow cytometry to identify monocyte subsets 58

Figure 7: Table to show amplicon and primer sequences for spike-in oligos (from the Song group as previously described[296]). 64

Figure 8: Table to show PCR reagents for creating a spike-in oligo amplicon 65

Figure 9: PCR reaction to create the spike-in oligonucleotide amplicons 65

Figure 10: Reagents used to create bead-bound DNA cleaning solutions 67

Figure 11: PCR amplification for 5hmC-enriched libraries 68

Figure 12: Transcriptomic profiles in peripheral monocytes in melanoma patients 77

Figure 13: Key selected transcripts differentially regulated in MM patients 81

Figure 14: Monocyte expression correlates of baseline monocyte count 84

Figure 15: Area under the curve for linear discriminant analysis classifier 86

Figure 16: Table to show key modulated transcripts with ICB 89

Figure 17: Transcriptional modulation of monocyte profiles in response to ICB 90

Figure 18: Treatment-wise responses to ICB 92

Figure 19: Differential modulation of transcripts with cICB versus sICB 93

Figure 20: Transcriptomic profiles at Cycle 4 (C4) 96

Figure 21: Transcriptional correlates of clinical response in peripheral monocytes 100

Figure 22: Monocyte subsets assessed with flow cytometry 102

Figure 23: Transcriptomic profiles modulated according to peripheral monocyte subset proportions 104

Figure 24: Clustree plots to show monocyte scRNA-seq clusters at varying resolution 112

Figure 25: Expression of key gene markers 114

Figure 26: Annotation of scRNA-seq monocyte subset with annotated expression data 116

Figure 27: Identifying conventional monocyte clusters in scRNA-seq 118

Figure 28: UMAP plot to show Seurat subset assignment by patient ID 119

Figure 29: Characterising monocyte clusters 123

Figure 30: Key hub genes and expression score modules 124

Figure 31: Characterising monocyte clusters 126

Figure 32: Proportion of monocyte cluster cells assigned to cell cycle phase 127

Figure 33: Trajectory analysis of monocyte clusters 129

Figure 34: Correlation between pseudotime and gene module score expression 130

Figure 35: Genes expression modulation as a function of pseudotime 131
Figure 36: Cluster-wise monocyte differential expression in melanoma patients versus healthy controls 133
Figure 37: GOBP pathways enriched in melanoma patients 135
Figure 38: Module scoring by monocyte cluster in melanoma patients and healthy controls 137
Figure 39: Trajectory analysis of monocyte clusters in MM and healthy controls 137
Figure 40: Modulation of transcriptomic profiles in monocyte clusters with ICB 140
Figure 41: GOBP pathway enrichment in monocytes with ICB treatment 141
Figure 42: Transcriptomic response to ICB characterised by module scoring 143
Figure 43: Trajectory analysis characterising ICB responses across monocyte clusters 144
Figure 44: Inter-individual variation in pseudotime following ICB 145
Figure 45: Classical and CD16$^+$ monocyte subset size 146
Figure 46: Mean expression scores in bulk RNA-seq data 147
Figure 47: Clustering levels and assignment across the scRNA-seq dataset 148
Figure 48: Monocyte subsets by sample 149
Figure 49: Characterising transcriptional profiles of classical monocyte subsets 153
Figure 50: Pathway enrichment of classical monocytes subsets 154
Figure 51: Trajectory analysis showing classical monocyte subsets 155
Figure 52: Analysis of trajectory branch points 157
Figure 53: Classical monocyte subset response to ICB 159
Figure 54: Pseudotime across classical subsets 160
Figure 55: Kaplan-Meier curves of overall survival (OS) by classical subset expression score in bulk RNA-seq data 162
Figure 56: Clustering levels and assignment across the scRNA-seq dataset 163
Figure 57: Diverse transcriptional profiles within the CD16$^+$ monocyte cluster 165
Figure 58: CD16$^+$ monocyte cluster proportions 166
Figure 59: Pathway enrichment analysis of CD16$^+$ clusters 167
Figure 60: CD16$^+$ clusters and overall survival (OS) 169
Figure 61: CD16$^+$ cluster expression scores and progression free survival (PFS) 170
Figure 62: CD16$^+$ monocyte subsets and association with autoimmune side effects 171
Figure 63: Pseudotime analysis of CD16$^+$ monocytes 172
Figure 64: Expression heatmap of branch point 2 173
Figure 65: Plot to show distribution of 5hmC and 5mC peaks 184
Figure 66: Metagene profiles across gene bodies to show 5hmC and 5mC distribution 186
Figure 67: Overlap of 5hmC and 5mC enriched regions 188
Figure 68: Pathway enrichment analysis of 5hmC and 5mC enriched genes 190
Figure 69: Enrichment of 5hmC and 5mC across functional genomic regions 193
Figure 70: Enrichment of 5hmC versus 5mC 194

Figure 71: Enrichment of 5mC and 5hmC in cfDNA in MM 196
Figure 72: Enrichment of 5mC and 5hmC across functional regions with ICB 198
Figure 73: Table to show key histone modifications and functional associations 199
Figure 74: Enrichment of 5hmC and 5mC across histone modifications 201
Figure 75: Odds ratios of 5hmC enrichment compared to 5mC cfDNA enrichment at histone modification sites 202
Figure 76: Enrichment of 5mC and 5hmC in cfDNA across histone modifications in MM patients and with ICB 204
Figure 77: Differentially hydroxymethylated regions (DhMRs) in patients with MM 206
Figure 78: Differential 5mC cfDNA profiles in patients with MM 207
Figure 79: DhMRs in patients with ICB 209
Figure 80: Differentially 5-methylated regions (DMRs) with ICB 211
Figure 81: Differentially methylated regions with ICB 212
Figure 82: Cancer testis antigens (CTAs) and DMRs/DhMRs 214
Figure 83: Comparing sICB and cICB responses 216
Figure 84: Key selected genes exhibiting DMRs with ICB 219
Figure 85: Pathway enrichment analysis of genes enriched for 5mC with sICB 220
Figure 86: DhMRs and clinical outcome at baseline 222
Figure 87: DhMRs post-treatment and progression at six months 224
Figure 88: DhMRs post- ICB treatment (C2) and association with irAEs 226
Figure 89: DMRs and six-month progression 228
Figure 90: DMR and irAEs side effects 230
Figure 91: Differentially methylated regions (DMRs) post-treatment and clinical outcome 232
Figure 92: Integration of epigenetic cfDNA profiles with circulating immune expression profiles 235

List of Appendices

Appendix A: Bulk RNA-seq cohort details (1) **248**

Appendix B Bulk RNA-seq cohort details (2) **249**

Appendix C: Bulk RNA-seq cohort details (3) **250**

Appendix D: GOBP pathways enriched in MM patients **252**

Appendix E: GOBP pathways depleted in MM patients **254**

Appendix F: GOBP pathways associated with baseline peripheral monocyte count. **255**

Appendix G: GOBP pathways negatively associated with baseline monocyte count **256**

Appendix H: GOBP pathways enriched in induced transcripts following ICB **259**

Appendix I: GOBP pathways enriched across transcripts suppressed ICB **260**

Appendix J: Transcriptional modulation in monocytes at cycle 4 (C4) with cICB **261**

Appendix K: Transcripts used for MM diagnostic classifier **262**

Appendix L: Monocyte proportion with flow cytometry **263**

Appendix M: scRNA-seq cohort data **264**

Appendix N: Elbowplot to show principal components **265**

Appendix O: *Clustree* plots to show *CD14* and *SELL* expression **266**

Appendix P: *Clustree* plot to show *CDKN1C* and *C1QA* expression **267**

Appendix Q: Classical and CD16^{+} cluster gene markers **268**

Appendix R: Classical subsets 0,1, 2 and 3 gene markers **269**

Appendix S: Classical subset 4,5,6 and 8 gene markers **270**

Appendix T: CD16^{+} cluster 1, 2 and 3 gene markers **271**

Appendix U: Classical and chemokine expression module gene sets **272**

Appendix V: Immaturity and non-classical expression module gene sets **273**

Appendix W: Antigen presentation and complement expression module gene sets **274**

Appendix X: Interferon-signalling and pro-inflammatory module expression gene sets **275**

Appendix Y: Macrophage differentiation and immunosuppression module gene sets **276**

Appendix Z: Boxplots showing classical and CD16^{+} expression scores by treatment **277**

Appendix AA: CD16^{+} cluster-wise enriched pathways **278**

Appendix BB: Percentage mitochondrial and ribosomal genes in classical monocyte subsets **279**

Appendix CC: 5hmC cohort details **280**

Appendix DD: Non-5hmC enriched samples **281**

Appendix EE: 5mC cohort details **282**

Appendix FF: Top 20 most significantly differentially methylated regions (DMRs) with immune checkpoint blockade (ICB) **283**

Appendix GG: Distribution of DMRs across the genome **284**

Appendix HH: Cohort utilised for integration of monocyte expression and 5mC cfDNA profiles **285**

Appendix II: Cohort utilised for integration of CD8^{+} T cell expression and 5mC cfDNA profiles **286**

Table of Abbreviations

Abbreviation	Term
5caC	5-carboxylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ADCC	Antibody-dependent cellular cytotoxicity
APC	Antigen presenting cell
AUC	Area under the curve
bp	Base pairs
C	Control
C1	Cycle 1
C2	Cycle 2
C4	Cycle 4
ccRCC	Clear cell renal cell carcinoma
cfDNA	Cell-free DNA
cfMeDIP	Cell-free DNA 5mC immunoprecipitation
CGI	CpG island
cICB	Combination immune checkpoint blockade
CTA	Cancer testis antigen
ctDNA	Circulating tumour DNA
cyTOF	Mass cytometry
DhMR	Differentially hydroxymethylated region
DLBCL	Diffuse large B cell lymphoma
DMR	Differentially methylated region
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dsDNA	Double stranded DNA
EDTA	Ethylenediaminetetraacetic acid
FACS	Fluorescence-activated cell sorting
FC	Fold change
FCS	Fetal calf serum
FDR	False discovery rate
FSC-A	Forward scatter area
GOBP	Gene Ontology Biological Process
HBSS	Hanks balanced salt solution
HCC	Hepatocellular carcinoma
HD	Healthy donor
HLA	Human leuokocyte antigen
ICB	Immune checkpoint blockade
IF	Immunofluorescence
IFN	Interferon
IHC	Immunohistochemistry
IL	Interleukin
irAE	Immune-related adverse events
LDA	Linear discriminant analysis
LDH	Lactate dehydrogenase

LMR	Lymphocyte to monocyte ratio
LPS	Lipopolysaccharide
Mb	Megabase
M-MDSC	Monocytic myeloid-derived suppressor cell
mAb	Monoclonal antibody
MACS	Magnetic-activated bead sorting
MDSC	Myeloid-derived suppressor cell
MHC	Major histocompatibility complex
MM	Metastatic melanoma
MMR	Mismatch repair
MSI	Microsatellite instability
NSCLC	Non-small cell lung cancer
oligo	DNA oligonucleotide
OR	Odds ratio
OS	Overall survival
P_{adj}	Adjusted p-value
PBMC	Peripheral blood mononuclear cell
PC	Principle component
PCR	Polymerase chain reaction
PDAC	Pancreatic ductal adenocarcinoma
PFS	Progression-free survival
PMN-MDSC	Polymorphonuclear myeloid-derived suppressor cell
QC	Quality control
qPCR	Quantitative polymerase chain reaction
RCC	Renal cell carcinoma
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
scRNA-seq	Single cell RNA sequencing
sICB	Single-agent immune checkpoint blockade
Slan	6-sulfo LacNAc
SSC-A	Side scatter area
TAM	Tumour-associated macrophage
TCR	T cell receptor
TDG	Thymine-DNA glycosylase
TES	Transcriptional end site
TET	Ten-eleven translocation
TIL	Tumour infiltrating lymphocyte
TLR	Toll-like receptor
TMB	Tumour mutational burden
TME	Tumour microenvironment
TNF	Tumour necrosis factor
TSS	Transcriptional start site
UMAP	Uniform manifold approximation and projection
VEGF	Vascular endothelial growth factor

1 Chapter 1: Introduction

In this thesis I characterise the peripheral response to immune checkpoint blockade inhibition, specifically focussing on both circulating myeloid population and epigenetic cfDNA profiles. These aspects will be discussed in this introductory chapter.

1.1 Melanoma and immune checkpoint inhibition

In the UK, melanoma has the fifth highest incidence of all cancer types, with 16,000 new cases diagnosed each year, and 2,300 annual attributable deaths[1]. Whilst most cases are cured by excision, in patients with unresectable or metastatic disease until the last decade therapeutic options have been limited. Melanoma is typically chemoradio-insensitive[2,3], and despite attempts at treatment with interleukin (IL) and interferon (IFN)-based therapies, prognosis for many years remained poor[4]. Prior to 2011, median survival of metastatic melanoma (MM) was nine months, with 18-month mortality approaching 100%[4]. The introduction of immunotherapy has revolutionised the outlook for patients with advanced or metastatic disease.

Melanoma exhibits one of the highest tumour mutational burden (TMB) rates of all malignancies[4,5], driven by the mutagenic properties of ultraviolet (UV) light exposure on sun-exposed sites[5,6]. The mutational heterogeneity characteristic of melanoma is highly immunogenic[7], however subsequent T cell activation is associated with induction of immune checkpoint expression and T cell exhaustion[8]. Simultaneous effects of tumour-associated macrophages (TAMs), myeloid-derived suppressor cell (MDSC) populations and immunosuppressive cytokines also facilitate immune evasion[9]. It is this background which facilitates therapeutic targeting of melanoma with immune checkpoint blockade (ICB).

Physiological immune responses are regulated by a variety of antigen markers which act as 'immune checkpoints'[10]. CTLA-4 was one of the first immune checkpoints to be identified[11]. CTLA-4 is an immunoglobulin cell-surface receptor with prominent inhibitory effects on T cells. CTLA-4 is expressed by cells including $CD8^+$ and $CD4^+$ T cells, and binds to CD80 and CD86 expressed by antigen presenting cells (APCs) with immune

inhibitory effects[12,13]. CTLA-4 is homologous to co-stimulatory receptor CD28, but has greater CD80/CD86 binding avidity[14]. CTLA-4 was the first immune checkpoint to show clinical efficacy following monoclonal antibody (mAb) blockade. Following promising results in clinical trials, in 2011 ipilimumab, an anti-CTLA-4 mAb, was the first licenced immune checkpoint inhibitor to be used therapeutically in the treatment of MM[15]. Phase III clinical trials demonstrated improved median OS with single agent anti-CTLA-4 therapy compared to a pg100 vaccine[16], with dose-dependent efficacy[4,17] . Unlike with anti-PD-1, a role for anti-CTLA-4 monotherapy in the treatment of melanoma has not been established.

In 1992 Ishida et al.[18] identified and cloned PD-1, characterising its role in apoptosis. This group subsequently described the effect of a PD-1 null mutation on T cell activation, establishing the role of PD-1 as an immune checkpoint and immune regulator[19]. Subsequently, the PD-1 ligand and member of the B7 family, PD-L1, was discovered by Freeman et al[20]. PD-L1 was found to be prominently expressed by antigen presenting cells and tumour cells, inhibiting CD28 co-stimulation, T cell proliferation, and cytokine production[20].

Further work demonstrated that blockade of PD-1 - PD-L1 interaction induces $CD8^+$ T cell effector activation[21]. Monoclonal antibodies (mAb) were found to effectively block PD-1-PD-L1 interaction, leading to T cell activation with anti-tumourigenic effects[22] in animals and subsequently humans[4,23,24] The CheckMate 066 study found that the anti-PD-1 mAb nivolumab was associated with response rates of 40% and increased progression free survival (PFS) compared to dacarbazine chemotherapy in patients with MM[25]. Findings from the KEYNOTE-001 trial demonstrated response rates of 52% and extended overall survival (OS) and PFS in patients with locally advanced or MM treated with the anti-PD-1 mAb pembrolizumab[26]. Anti-PD-1 therapy has been shown to induce generalised anti-cancer responses with therapeutic efficacy superior to that of standard-of-care, providing improved clinical outcomes in cancers including non-small-cell lung cancer (NSCLC)[27], renal cell carcinoma (RCC)[28], urothelial carcinoma[29] and head and neck cancers[30].

In MM subsequent trials combining anti-PD-1 and anti-CTLA-4 therapy (combination ICB - 'cICB') found association with enhanced clinical response compared to ICB

monotherapy (single-agent ICB - 'sICB'). Overall response rates (ORR) of 53% were found in phase II studies[4,31] with a subsequent phase III clinical trial demonstrating clinical benefit in excess of that provided by monotherapy regimens[4]. In 2015 a phase II trial showed more favourable ORR and PFS with cICB over either anti-CTLA-4 or anti-PD-1 single agent treatment[26]. Follow-up of this patient cohort demonstrated significantly reduced mortality with cICB versus single-agent ipilimumab, and improved median PFS in those receiving cICB versus single-agent nivoumab[33]. However, 59% of patients who received combination ipilimumab and nivolumab developed a grade three immune-related adverse event[33]. Five-year outcomes from this study demonstrated OS of 52% in those receiving cICB, versus 44% and 26% in those treated with single agent ipilimumab and nivolumab respectively[34].

To date, ICB has been shown to provide clinical benefit in not just melanoma but many cancer types[28,35,36], and clinical trials exploring responses across a diverse range of tumour types are in progress[37]. As the mechanistic understanding of molecular determinants of ICB response expands, broader applications for ICB therapy can be explored. For example, ICB treatment in patients with tumour mismatch repair (MMR) deficiency has been found to be clinically efficacious[36]. ICB is now licenced for use based on MMR deficiency in patients with metastatic solid tumours, irrespective of tissue-of-origin[38].

Despite the promise of ICB as a therapeutic strategy, only a proportion of patients will experience clinical benefit. In addition, of those patients for whom ICB is efficacious, durable clinical response is only experienced by a small minority, with most relapsing[39]. As noted, these drugs are not without adverse effects, and a high incidence of immune-related advserse events (irAEs) are reported, including dermatitis, colitis, gastritis and hepatitis[4,40]. Similar autoimmune toxicity profiles are seen with both PD-1 and anti-CTLA-4 therapies, however toxicity is typically both more frequent and severe in those receiving combination therapy[40]. There is evidence to show that autoimmune side effects are associated with favourable clinical responses[41–43]. In a recently published prospective melanoma cohort early autoimmune side effects were found to be associated with extended PFS and OS[40]. However, balancing the equilibrium between toxicity and clinical response is not straightforward, and treatment effects may prompt

treatment cessation in cases whereby such effects cause considerable morbidity, or prove life-threatening.

1.2 Predictive biomarkers for immune checkpoint blockade

Although the development of ICB has revolutionised the treatment of patients with advanced or MM, there remains the need for robust predictive markers of response to treatment to enable targeting of these therapeutic agents for those most likely to derive clinical benefit, and to identify non-responders. There are multiple biomarkers that have predictive utility in ICB. These include tumour mutational burden (TMB), MMR deficiency, PD-L1 expression and tumour infiltrating lymphocytes (TILs)[44]. Here some of the key predictors are summarised briefly. However, to date these have limited clinical utility in patients with MM, and clinicians are typically dependent on clinical features to guide treatment decisions[45].

1.2.1 Tumour infiltrating lymphocytes (TILs)

Tumour infiltrating lymphocytes (TIL) are independently associated with prognosis in patients with melanoma[46], colorectal cancer[47] and NSCLC[48]. In melanoma, TIL response is categorised as brisk, non-brisk or absent, with a brisk response associated with a more favourable prognosis[49]. A non-brisk TIL response was associated with a 30% reduction in mortality in a retrospective study of new-diagnosed melanomas, whilst a brisk response associated with a 50% reduction[46]. Similarly, Tumeh et al.[50] found that $CD8^+$ T cell infiltration was more frequent in anti-PD-1 responders with melanoma versus non-responders. There has been extensive characterisation of tumour immune microenvironments, however dissecting mechanisms robustly predictive of clinical prognosis and treatment response is challenging, and standardisation of assessment problematic[51,52]. TIL infiltration is not at present utilised as a predictive marker for ICB in patients with melanoma.

1.2.2 PD-L1 expression

PD-L1 is expressed on the surface of antigen presenting cells (APCs) and variably within some solid tumours[53], and its expression in melanoma is associated with prognosis[54]. PD-L1 expression is also predictive of response to anti-PD-1 therapy in patients with

cancers such as NSCLC[55,56]. Assessment of tumour expression of PD-L1 with immunohistochemistry (IHC) has been approved by the FDA in the US and by the National Institute of Health and Care Excellence (NICE) in the UK as a companion diagnostic in NSCLC, and is now part of routine care[57,58]. In melanoma PD-L1 is not a sufficiently robust independent predictive biomarker and PD-L1 expression is not routinely assessed in the clinic[59]. In particular, there are challenges due to tumour heterogeneity with considerable discordance reported between biopsy and resection specimens[60], as well as between primary and metastatic tumour sites[61]. Standardisation of approach is also problematic, with different studies choosing variable PD-L1 expression cut-off proportions with little effect on predictive value[62].

1.2.3 Tumour mutational burden and mismatch repair deficiency

Tumour mutational burden (TMB) defined as the number of somatic mutations per megabase (Mb) of the genome[63]. The high TMB typically exhibited in melanoma[4,5] leads to increased propensity to neoepitope formation, associated with greater immunogenicity and response to ICB[4]. Enhanced response to ICB in patients with high TMB has been demonstrated across tumour types, including NSCLC[64], urothelial carcinoma[65], and small cell lung cancer[51,66]. In the US, pembrolizumab is now licenced for patients with tumours which exhibit high TMB[67]. Similarly, tumour DNA mismatch repair (MMR) deficiency and associated microsatellite instability (MSI) is associated with response to anti-PD-1 treatment[68,69] and these have recently been licensed for use in patients with MMR-deficient metastatic solid organ tumours regardless of tissue-of-origin. However, despite high TMB and MMR deficiency showing promise as predictive biomarkers, both of these approaches lack sufficient efficacy for use in the clinic in patients with MM.

1.2.4 Summary of existing predictive markers and the case for exploring the peripheral immune compartment

In summary, although there are a number of potential prognostic and predictive biomarkers for ICB, many are not suitably sensitive nor specific for use in the clinic in patients with advanced or MM. In addition, existing biomarkers are tumour-based and thus require tissue-based analysis. In many cases this may be problematic, limited by

patient characteristics or tumour site, and is associated with risk of local injury. In addition, core biopsy sampling may not be representative of tumour characteristics due to tumour heterogeneity[70]. Identification of peripheral circulating biomarkers may circumvent these problems and, given ICBs directly target the immune system, have particular promise.

1.3 Monocytes

Previous work characterising peripheral $CD8^+$ T cell responses has demonstrated distinct on-treatment modulation in response to ICB, with expression profiles exhibiting predictive clinical utility[71]. The role played by other cell types is less well characterised, however there is evidence that myeloid cells including circulating monocytes and tissue-based macrophage populations play a role in modulating response to ICB[72,73].

The mononuclear phagocytic system (MPS) comprises monocytes, macrophages and dendritic cells. Monocytes circulate within the peripheral circulation, whilst macrophages are a tissue-based population and can be derived from both circulating monocytes, or from embryonic precursors, according to tissue-type or the presence/absence of inflammation[74]. Whilst both cell types have a key role in phagocytosis and antigen presentation, monocytes play a prominent role in the innate immune response and inflammation, whilst macrophages are involved in tissue homeostasis[75]. To explore the role of this population in modulating response to ICB, a detailed understanding of monocyte phenotype, function and differentiation into macrophages is important. This is explored here in the context of cancer and ICB.

Monocytes are myeloid-derived cells which play a critical role in both innate and adaptive immune responses[76]. Monocytes recognise and bind pathogen-associated molecular patterns (PAMPs) such as lipopolysaccharide (LPS), via pattern recognition receptors (PRRs) such as Toll-like receptors (TLRs)[77,78]. This induces signalling via pathways such as NFκB, MAPK and interferon regulatory factor (IRF) pathways, activating pro-inflammatory cytokine release[77,78]. In addition, cyclic dinucleotides (CDNs), either produced by bacteria or via activation of cyclic GMP-AMP synthase (cGAS) activate stimulator of interferon genes (STING)[78]. This leads to pro-inflammatory cytokine production via NFκB and IRF signalling, inducing T cell activation[78]. Myeloid

cells also exhibit prominent MHC Class II antigen presentation, which can provide co-stimulatory effects for $CD8^+$ and $CD4^+$ T cells[79]. Monocytes also produce components of predominantly classical and alternative complement pathways, some of which have direct bactericidal effects via the membrane attack complex, but which can also promote phagocytosis or have chemotactic effects[80]. Monocytes also exhibit phagocytic activity, and can engulf and remove pathogens and cellular debris[81].

The human Fc receptor family have an important role in regulating the immune response to pathogens[82]. In particular, the human immunoglobulin (Ig) Fc gamma family of receptors (FcγRs) have a key role in facilitating monocyte and macrophage activity[83] (Figure 1). These receptors bind to the Fc part of IgG, inducing phosphoinositide 3-kinase (PI3K), protein kinase C (PKC) and signalling via the Rho GTPase family, causing phagocytosis of IgG-labelled cells, as well as the activation of mitogenic signalling pathways such as mitogen-activated protein kinase (MAPK), leading to chemokine and cytokine production[83]. Whilst most FcγRs have an activating role on this signalling cascade, FcγR2b has an inhibitory role[83]. Monocytes prominently express FcγR1 and FcγR2a, with non-classical monocytes also expressing high levels of FcγR3a (CD16)[84].

Human FC gamma receptor	Antigen name
FcγR1	CD64
FcγR2a	CD32
FcγR2b	CD32
FcγR2c	CD32
FcγR3a	CD16
FcγR3b	CD16

Figure 1: Human Ig Fc gamma receptors (FcγRs)

Initial identification of monocytes was dependent on their distinctive morphological features in peripheral blood films, as mononuclear cells with a characteristic bean-shaped nucleus. However, the development of flow cytometry led to the identification of two distinct monocyte subsets defined by the relative antigenic expression of CD14 and CD16; a $CD14^+CD16^-$ classical subset, and a second distinct $CD16^{++}CD14^{low}$ non-classical subset[85–87]. A third monocyte subset, termed 'intermediate', was subsequently identified, characterised by intermediate CD16 expression and a $CD14^{low}CD16^+$ antigenic profile[88,89].

1.3.1 Characterising monocyte subsets

Classical monocytes are the most abundant of the three monocyte subsets, comprising 80-90% of the circulating monocyte population[89]. Intermediate monocytes comprise approximately 5% of the circulating monocyte population, whilst non-classical monocytes make up approximately 10%[88,90]. Monocyte subsets have distinct antigenic and transcriptional features reflecting distinct functional roles (Figure 2 and Figure 3). These will be discussed as follows.

1.3.1.1 Classical monocytes

Classical monocytes exhibit high expression of chemokine receptors *CCR1, CCR2, CCR5, CXCR1, CXCR2*[91–93] in keeping with their capacity to migrate to sites of inflammation (Figure 2 and Figure 3). Wong et al.[93] performed expression profiling using microarray analysis of fluorescence-activated cell sorting (FACS)-sorted monocyte subsets, and found that classical monocytes were characterised by high expression of *CD14*, *CCR*s and *SELL*, with enrichment of pathways including angiogenesis, wound healing and coagulation. S100 genes were also highly expressed. The S100 gene family bind calcium and have diverse roles in modulating migration, apoptosis and stimulating a pro-inflammatory response[94]. Schmidl et al.[95] utilised HeliScope Cap Analysis of Gene Expression (HeliScopeCAGE)[96], a method which characterises messenger RNA (mRNA) at the 5' end allowing the identification of transcriptional start sites (TSSs)[96], to characterise transcriptomic profiles across flow sorted monocyte subsets. They found that classical monocyte were enriched for genes involved in carbohydrate metabolism, apoptosis, and toll-like receptor (TLR) signalling[95]. The most highly expressed genes in the classical subset included *VCAN, CD36, CLEC4E, CD163, VEGF* and *CCR2*[95]. Classical monocytes have prominent phagocytic activity and secrete high levels of cytokines[97,98] including CCL2, CCL3, IL6 and IL8[90,91]

1.3.1.2 Intermediate monocytes

Intermediate monocytes are defined by intermediate expression of CD16, and high expression of MHC class II[90,95,99,100]. Intermediate monocytes are derived from classical monocytes, and undergo differentiation into non-classical monocytes. Intermediate and non-classical monocytes are closely associated according to hierarchical clustering of

gene expression profiling and antigenic markers[93,101], although evidence regarding this transcriptional similarity is conflicting[102]. This variation may be partly due to user variability in preliminary subset selection using FACS sorting.

Wong et al.[90] demonstrated that intermediate monocytes exhibited intermediate expression of classical and non-classical-associated genes, but also expressed high levels of *GRAFA2, PLAC8* and *CLEC10A.* MHC Class II antigen expression was prominent[90]. Gren et al.[100] found that antigen presenting genes were highly expressed by intermediate monocytes, as well as enrichment of genes associated with differentiation, including transcription factors *IRF5* and *IRF8*. Interestingly, intermediate monocytes have also been shown to express transcripts associated with NK and T cell function such as *ULBP3* and *RAETIE*[100]. Wong et al.[90] also found genes such as *NKG7* were highly expressed within the intermediate group. Schmidl et al.[95] found that intermediate monocytes were characterised by fewer subset-specific genes compared to classical and non-classical monocytes but exhibited high expression of human leukocyte antigen (HLA) genes[95].

Other studies have reported heterogeneity within the intermediate monocyte compartment[103]. In mice, two intermediate monocyte populations are recognised, one of which has a profile intermediate between Ly6C^{hi} and Ly6C^{lo} monocytes, and another characterised by CD209a and MHC class II expression, from which DC are thought to derive ('CD209a^{+} MHCII^{+}Ly6C$^{hi\text{-}to\text{-}int}$')[103]. In humans, diversity within this compartment has also been reported. Villani et al.[104] characterised two monocyte subsets using single-cell RNA sequencing (scRNA-seq), one of which was characterised by a cytotoxic gene signature, although there is some controversy regarding the annotation of these subsets, which will be discussed subsequently (1.3.1.5).

1.3.1.3 Non-classical monocytes

Non-classical monocytes have a pro-inflammatory phenotype, and produce high levels of TNFα in response to LPS stimulation[91,98,105]. Whilst non-classical monocytes produce lower levels of inflammatory cytokine such as IL-6 and IL-1β than their classical counterparts[97,98] they are characterised by prominent anti-viral responses, with high expression of interferon (IFN)-induced genes, such a *IFITM1* and *IFITM2* and classical complement genes such as *C1QA*[90,95,101] (Figure 3). Non-classical monocytes exhibit

enrichment for respiratory chain metabolism pathways rather than the carbohydrate-based metabolic pathways that are enriched in classical monocytes[95]. Ong et al.[106] found that following LPS stimulation, non-classical monocytes exhibited low Ki-67 proliferation index, short telomere length, as well as high levels of reactive oxygen species (ROS) in keeping with a mature, senescent phenotype. Consistent with this, non-classical monocytes exhibit high expression of negative cell cycle regulators such as *CDKN1C* and *MTSS1*[107].

Non-classical monocytes have patrolling behaviour and function and exhibit prominent expression of *CX3CR1*[90], which facilitates binding to the endothelium and promotes migration and endothelial adhesion. Non-classical monocytes also monitor and maintain the integrity of the vascular endothelium[102,108] and show enrichment of genes associated with cytoskeletal rearrangement such a *CD294* and S*IGLEC10*[90].

$CD16^+$ monocytes facilitate antibody-dependent cellular cytotoxicity (ADCC) with tumour- and virus-infected cell killing function. Yeap et al.[109] found that $CD16^+$ monocytes display TNFα mediated-ADCC effects, with CD16-receptor antibody binding stimulating TNFα production and β2-integrin activation[109]. Monocytes lacking CD16 expression failed to exhibit ADCC-effects, and CD16 was induced by IL-10, TGFβ or M-CSF[109], as previously described[110–112]. Cytokines such as LPS, IFN-γ or S100A9 also enhanced the ADCC effect of $CD16^+$ monocytes[109]. CD16-expressing monocytes exhibited tumour-killing behaviour, and were also able to kill virus-infected cells via ADCC mechanisms[109]. CD16 expression was correlated with lysis activity[109].

Monocyte subset	Key function
Classical	Chemokine production Chemotaxis Phagocytosis Angiogenesis Wound healing
Intermediate	Antigen presentation
Non-classical	TNF production in response to LPS Anti-viral responses Respiratory chain metabolism Complement production Antibody-dependent cellular cytotoxicity (ADCC)

Figure 2: Summary of key monocyte subset functions

Monocyte subset	Key marker genes		
	Gene	Gene name	Function
Classical	*CD14*	Cluster of differentiation 14	Pattern-recognition receptor involved in anti-bacterial responses[113]
	CCR2	C-C chemokine receptor type 2	Monocyte chemotactant[114]
	SELL	L-selectin	Leucocyte adhesion[115]
	CD36	Cluster of differentiation 36	Pattern recognition receptor, involved in processes including cell, adhesion, fatty acid transport, platelet adhesion[116]
Intermediate	*HLA-DRA* (MHC Class II genes)	HLA class II histocompatibility antigen, DR alpha chain	Antigen presentation[117]
Non-classical	*FCGR3A*	Low affinity Ig gamma Fc region receptor III-A	Involved in ADCC[109]
	CDKN1C	Cycline dependent kinase inhibitor 1C	Cell cycle regulation[118]
	MTSS1	Metastasis suppressor protein 1	Cell cycle regulation[119]
	CX3CR1	CX3C chemokine receptor 1	Leucocyte adhesion and circulation[120]
	C1QA	Complemenent C1q subcomponent subunit A	Classical complement pathway[121]
	IFITM3	Interferon-induced transmembrane protein 3	Anti-viral function[122]

Figure 3: Summary of key monocyte gene markers

1.3.1.4 Mass-spectrometry approaches to characterise monocyte subsets

Whereas most studies characterising monocyte subsets utilise flow cytometry to define subsets, this approach is limited by the small number of antigen markers that can be simultaneously assessed due to overlapping wavelengths of fluorescent dyes[123]. Mass cytometry (cyTOF) utilises antibodies conjugated to heavy metal ions binding to cellular antigens; subsequent ionisation with argon plasma and resulting charge as electric signals are detected allows for detection of antigenic expression[124]. cyTOF allows phenotyping with a larger number of antigen markers and has been used to further characterise monocyte subsets.

Roussel et al.[125] performed cyTOF with a panel of 38 gene markers. Non-classical monocytes were characterised by low CCR2, CD36 and the FcγR CD64 expression, but higher levels of CD11c and HLA-DR than their intermediate and classical counterparts[125]. CD11c is a transmemrane protein which has a role indiverse range of processes including cell migration and phagocytosis[126,127], whilst HLA-DR is a MHC class II antigen presentation protein. Non-classical monocytes ($CD14^{low}CD16^{+}$) were divided into two groups based on 6-sulfo LacNAc (Slan) expression[125]. Slan is a carbohydrate which binds to P-selectin glycoprotein ligand-1 (PSGL-1)[128], which is expressed by leucocytes with a role in facilitating endothelial adhesion and migration to sites of inflammation[129]. Slan

has previously been shown to have potential utility in distinguishing between intermediate (slan-) and non-classical monocytes[128]. Ravenhill et al.[97] used FACS-sorted monocytes from PBMCs to enrich for classical, intermediate and non-classical monocyte subsets and then performed antigenic profiling using cyTOF. Principal component analysis (PCA) showed that classical monocytes were distinct, whilst intermediate and non-classical monocytes demonstrated a greater degree of phenotypic overlap[97]. Antigenic profiling and transcriptomic expression data revealed good concordance[97]. Further characterisation of monocytes with mass cytometry approaches have explored CD14 and CD16-independent categorisation of monocytes. For example Ong et al.[130] used cyTOF with a panel of 34 markers and found that five markers; HLA-DR, CD64, CD86, CD33 and CCR2, effectively distinguished between the three conventional monocyte subsets.

1.3.1.5 Characterisation of monocyte subsets using dissection at a single-cell level

The development of single cell sequencing (scRNA-seq) technologies has facilitated high-resolution transcriptomic analysis of PBMC subsets across different health and disease states with unsupervised clustering techniques. Villani et al.[104] presented one of the first scRNA-seq analyses of myeloid subsets. This group utilised the high-throughput droplet and plate-based scRNA-seq technology Smart-Seq2[131] to sequence FACS-sorted PBMCs from a single healthy individual[104]. They identify four monocyte subsets, which broadly align to classical ('Mono1'), non-classical ('Mono2') and two intermediate subsets ('Mono3' and 'Mono4'). However, this study is limited by the small total cell numbers (n=337 monocytes) derived from a single donor. It has been subsequently suggested that Mono4, which is characterised by cytotoxic gene expression, is actually representative of CD56-negative NK cell contamination, and that the conventional three-subset convention remains correct[132]. The Mono3 category, denotated as an intermediate population by Villani et al. was characterised by neutrophil-associated transcripts, and it has been suggested that this may in fact comprise a classical subpopulation[133,134].

Notably, a robust intermediate monocyte population has also been challenging to define in the few other scRNA-seq studies which have explored monocyte phenotype and function. Zilionis et al.[134] performed scRNA-seq in myeloid cells derived from blood and

NSCLC in both humans and mice and identified three main monocyte groups, broadly aligning to the Villani et al.[104] Mono1 (classical), Mono2 (non-classical) and Mono3 monocyte groups. The Mono3 population expressed neutrophil-associated transcripts and a $CD14^{+}CD16^{-}$-type transcriptomic profile, and was thought to represent a classical subset, rather than an intermediate subset. Vishwakarma et al. [135] performed scRNA-seq of tumours and peripheral blood samples from patients with RCC. They defined only two monocyte subsets in patients with clear cell RCC (ccRCC), broadly aligned to a $CD14^{+}CD16^{-}$ classical subset, and a CD14-CD16+ non-classical subset. This group found that classical monocytes expressed high levels of chemokine genes, whilst non-classical monocytes expressed genes associated inhibition of apoptosis and NFκB signalling[135]. Similarly, Goudot et al.[136] used Drop-Seq to perform scRNA-seq of 429 monocytes. Unsupervised clustering identified two monocyte subsets; a $CD16^{-}$ classical subset, and a $CD16^{+}$ subset[136]. High-resolution clustering within the $CD16^{-}$ classical subset failed to identify additional levels of heterogeneity within this compartment[136].

1.3.2 Summary of monocyte subset phenotyping

It is evident that monocytes exhibit considerable heterogeneity in terms of transcriptional and antigenic expression and associated function. Whilst there are widely reported gene markers and functions that define each subset (Figure 2) there is considerable variation in the antigenic and transcriptomic monocyte subset profiles described within the literature[133]. This may be due to limitations in study design, including widely using selection based on FACS sorting according to CD14 and CD16 expression which is likely to be affected by user variability, as well as sampling from within small donor cohorts. Studies utilising scRNA-seq provides opportunities for novel dissection of the peripheral monocyte compartment, however current studies are limited by cell number. High-throughput high-dimensional scRNA-seq characterisation of monocyte subsets across multiple individuals and clustering using unsupervised approaches would be of value in characterising peripheral monocyte populations. This is an approach I use in this thesis.

1.3.3 Monocyte differentiation, maturation and fate

Monocyte subsets represent variable states of differentiation and maturation. Monocytes are derived from multipotential progenitor cells within the bone marrow. Here, multipotent haemopoietic stem cells (HSC) differentiate into common lymphoid precursor cells (CLPs)[137] or common myeloid precursor (CMP) cells[103,138]. From CMPs, differentiation occurs into granulocyte and macrophage progenitors (GMPs)[138], from which monocyte/DC progenitors (MDP) derive[139]. Cell from the MDP compartment then differentiate into either unipotent common monocyte progenitors (cMoP) or common DC progenitors (cDC), from which monocytes and DCs are derived respectively[103,140].

Following differentiation from immature progenitor cells, classical monocytes migrate into the blood from the bone marrow, requiring *CCR2* expression for exit into the peripheral bloodstream[141]. They circulate in the blood for approximately one day, wherefrom they travel to tissue in response to chemokines such as CCL2[142]. They then differentiate within tissues into macrophages, or remain in the circulation and develop into non-classical monocytes[103].

Macrophages derived from circulating monocytes acquire transcriptomic and antigenic phenotypes characteristic of their tissue-resident counterparts, however do retain some features that indicate their circulating monocyte ancestry[103,143]. Differentiation into macrophages is not automatic; monocytes can pass into non-lymphoid tissues and then return to the bloodstream[144]. Most organs contain a population of embryonic macrophages derived from fetal precursors that survive into adulthood which make up a variable proportion of the total macrophage population according to tissue[145]. However, monocyte-derived engrafted macrophages are present in most tissues, including the gut[146], heart[147], lung[144] and dermis[148].

Rather than differentiating into macrophages within tissues, a small proportion of classical monocytes undergo transition into non-classical monocytes[149]. The relative size of the classical and non-classical monocyte components is affected by production of monocytes from the stem cell compartment, tissue extravasation of monocytes, and differentiation into monocyte-derived macrophages[103]. These factors are influenced by process such as such as inflammation or infection. In humans, lipopolysaccharide (LPS) stimulates an initial monocytopenia, with a subsequent increase in classical, then

intermediate and subsequently, non-classical monocytes[149]. Patel et al.[149] used in vivo deuterium labelling to show that intermediate monocytes survive for about four days, whilst non-classical monocytes approximately seven days. Transition between monocyte subsets, from classical to intermediate and then non-classical, was observed[149]. Only a small proportion (approximately 1%) of classical monocytes eventually differentiate into non-classical monocytes, whilst most move into tissues and differentiate into macrophages, or undergo apoptosis[149]. Differentiation of non-classical monocytes is dependent on colony stimulating factor 1 receptor (*CSF1R)* expression[150,151]. CSF1R is a CSF1 receptor has a key regulatory role in macrophage differentiation[152].

The linear differentiation trajectory between monocyte subsets is supported by studies which have identified markers of maturity in non-classical monocytes[88]. For example, Ancuta et al.[107] found that non-classical monocytes expressed high levels of cell cycle regulation genes, such a *CDKN1C* and *MTSS1*. Pathway enrichment analysis of non-classical-defining gene markers have demonstrated enrichment of differentiation-associated pathways[91,107]. Non-classical monocytes also display a senescent phenotype, in keeping with a mature, terminally differentiated cell group[106].

Outstanding questions and challenges regarding myeloid differentiation remain. For example, mechanisms by which classical monocytes are marked for a particular lineage pathway have not yet been characterised[103]. scRNA-seq has shown that the relative proportion of classical and non-classical monocytes (or their Ly6C^{hi} and Ly6C^{low} murine counterparts) remain relatively constant[153]. However, it has been suggested that experimental assays may not be representative of monocyte subset proportions as Ly6C^{low} monocytes, and their human non-classical counterparts, have strong endothelial adhesive properties. Non-classical monocytes may therefore not be well-represented in peripheral sampling[103].

Whilst monocytes show distinct heterogeneity in terms of functional activity and antigenic profiles, epigenetically monocyte subsets appear similar and utilise the same promotors, demonstrating a shared origin and phenotype[103,153]. Although much of the existing literature regarding monocyte discusses discrete subsets, it may be more appropriate to consider monocytes as a population comprising states within a

differentiation continuum[154,155]. Trajectory analysis using scRNA-seq techniques would be of value in characterising differentiation states across the monocyte compartment. This will be explored using scRNA-seq approaches later in this thesis.

1.4 Myeloid populations in cancer

There is increasing evidence to demonstrate that both circulating and tumour-based myeloid populations have a role in modulating tumourigenesis. Here, current understanding regarding myeloid populations in cancer, and their role as prognostic markers, is explored.

1.4.1 Macrophages and the tumour microenvironment

Cells of the mononuclear phagocytic system (MPS) play a key role in modulating tumorigenesis. Under the influence of the T_H1 cytokines IFN-γ and LPS, macrophages within the tumour microenvironment (TME) are polarised towards a pro-inflammatory phenotype with anti-tumourigenic effects. These macrophages are termed classically activated or 'M1' macrophages[156]. Conversely, T_H2 cytokines IL-3 and IL-4 are associated with polarisation towards an immunosuppressive phenotype and production of cytokines such as IL-10 and arginase, with pro-tumourigenic effects. These cells are termed alternatively activated or 'M2' monocytes[156]. In cancer, tumour-associated macrophages (TAMs) are typically polarised towards an M2 phenotype, with upregulation of tissue remodelling and wound healing properties[157]. TAMs produce pro-angiogenic cytokines, and facilitate tumour cell invasion and metastasis. TAMs induce an immunosuppressive tumour phenotype characterised by production of prostaglandins, arginase, IL-10 and TGF-β, which stimulate T_{reg} responses[158,159], and expression of immune checkpoints such as PD-L1 and B7-H4[160]. More recent it has been suggested that a binary classification is not a suitable method of fully describing the plasticity seen within the TAM compartment due to overlap in stimuli responses and profiling between the two groups, and that macrophage nomenclature should reflect the appropriate stimulus (e.g. M(IL-4), or M(IFN-γ)[161]. Despite this, this nomenclature remains widely used.

1.4.2 Myeloid signatures are prognostic in cancer

Expanded circulating monocyte populations have been reported across different cancer types, and reflect induction of myelopoiesis[162]. Increasing peripheral monocyte counts are negatively prognostic in cancers including melanoma[40,163], breast[164] and colorectal[165]. In particular, an elevated lymphocyte-to-monocyte ratio (LMR) is associated with a more favourable prognosis across cancer types[166–170], although this does not appear to be prognostic in melanoma. Consistent with these observations, myeloid signatures and tumour infiltration are associated with poor prognosis in cancers types including breast cancer[171], NSCLC[172] and bladder cancer[173,174].

Whilst much of the literature has focussed on dissecting the predictive role of TAM populations; circulating monocytes display phenotypic alteration in the context of cancer and also have prognostic and predictive utility.

1.4.3 Reprogramming of the myeloid compartment in cancer

In cancer, monocytes undergo transcriptional and antigenic reprogramming acquiring distinct cancer-associated profiles[155] (Figure 4). Cancer is associated with polarisation of haematopoiesis towards myeloid populations via tumour-based production of cytokines such as GM-CSF and M-CSF[175]. Chemotaxis and migration markers such as *CX3CR1* and *CCR2* are typically upregulated in monocytes in cancer[176–179]. These markers facilitate monocyte migration to tumour sites, which may then differentiate into M2-polarised TAM populations with pro-tumourigenic effects. The CCR2-CCL2 axis facilitates monocyte recruitment from the bone marrow to tumour sites, and blockade of this axis reduces monocyte infiltration in tumours, and impairs tumour metastasis[178]. In colorectal cancer, monocytes have been found to exhibit high levels of *CCL2*, a key protein facilitating monocyte chemotaxis and recruitment, and in endometrial and breast cancer peripheral monocytes are enriched for expression of transcripts associated with cell migration, as well as angiogenesis and apoptosis pathways. Qian et al.[178] showed that classical monocytes were recruited to tumour sites via the CCL2-CCR2 axis, facilitating growth of breast cancer metastases.

Broadly, it is thought that cancer-associated monocyte profiles are defined by an immunosuppressive phenotype[162]. For example, Gordon et al.[180] found that monocytes

from patients with ovarian cancer exhibited reduced facility for antibody-mediated phagocytosis and cytotoxicity[180]. This study found this effect across both a $CD14^+$ population and a $CD14^{bright}CD16^-$ subset of monocytes[180], and it can be hypothesised that this effect may be due to variable expression of FcγRs in monocytes from patients versus healthy donors. Circulating monocytes in cancer patients may show impaired differentiation and skew towards a more immature phenotype[136,181]. For example, Goudot et al.[136] found that monocytes from breast cancer patients exhibited reduced *MAFB* expression, which is required for monocyte-macrophage differentiation. Evasion of monocyte phagocytic function in cancer has been demonstrated. For example, many tumour cells exhibit a TNF-related apoptosis-inducing ligand (TRAIL) resistant-phenotype[182] and upregulate CD47, facilitating tumour cell evasion of phagocytic mechanisms[183]. Monocytes derived from cancer patients have been found to have reduced ADCC activity [180].

Despite upregulation of immunosuppressive function, enrichment of pro-inflammatory anti-viral pathways in peripheral monocytes has also been demonstrated in some studies. For example. Ramos et al.[177] performed microarray expression profiling of monocytes from healthy donors and breast cancer patients, and found that cancer-associated profiles included enrichment of IFN-α and IFN-γ, and downregulation of TNFα signalling. In addition, a number of studies have demonstrated elevated monocyte *HIF1A* expression in cancer patients, associated with reduced response to hypoxia[176], thus facilitating tumour growth. Dysregulation of physiological metabolic processes have also been identified. For example, Krieg et al.[73] identified reduced glucose and amino acid transporter expression in monocytes from melanoma patients versus healthy individuals.

Perhaps the most widely documented change in peripheral circulating myeloid populations in cancer is the expansion of monocytic myeloid-derived suppressor cells (M-MDSCs). There are two main types of MDSC; polymorphonuclear (PMN-MDSC) and monocytic (M-MDSC)[184]. The former express $CD11b^+CD14^-CD15^+$ or $CD11b^+CD14^-CD66b$[185], whilst the latter are morphologically similar to monocytes and are characterised by expression of CD14, CD11b, CD33 and low HLA antigen presenting gene expression[184]. The hallmark of MDSCs is their immunosuppressive phenotype, with

the capacity for suppression of CD3/CD28 driven T-cell proliferation or inhibition of IFN-γ production the gold standard to confirm classification[185]. M-MDSCs are subject to controversy both in terms of classification and nomenclature, but also regarding cell-of-origin. It has been proposed that M-MDSCs are derived from the circulating monocyte pool as opposed to unique progenitor cells[186]. However, current opinion is that MDSCs share a common progenitor with monocytes[184]. In the context of pro-inflammatory conditions including cancer, there is a shift towards a MDSC phenotype, with both monocytes and M-MDSC existing simultaneously in the circulation as a dual population[184].

MDSCs have been found to be more abundant in patients with cancer versus healthy individuals across many cancer types including renal[187], breast[188], colorectal cancers[189] and melanoma[190,191], and are associated with poor prognosis. For example, Gonda et al.[188] found that breast cancer patients exhibited an expanded MDCS population compared to healthy controls. Those with MDSCs comprising >0.1% of all PBMCs had shorter OS than those with MDSC populations comprising <0.1% of all PBMCs[188]. In a meta-analysis, Zhang et al.[190] found that MDSCs were predictive of reduced OS and PFS across many different solid organ malignancies.

M-MDSC are a controversial cell group due to the challenges associated with classification and cell-of-origin. Perhaps reflecting the lack of consensus regarding MDSC phenotyping and classification, a number of studies have explored the prognostic and predictive utility of myeloid populations characterised by $CD14^{+}HLA\text{-}DR^{lo/-}$, without more detailed characterisation of other M-MDSC markers. Downregulation of MHC class II antigens, such as HLA-DR, and the costimulatory receptor CD86[155], have also been observed in circulating monocytes in patients across a wide range of cancer types. Ugurel et al.[179] used flow cytometry to demonstrate reduced expression of HLA class II and CD86/B7-2 on peripheral monocytes in patients with melanoma. Expansion of this subset was associated with poor prognosis[179]. It is difficult to determine whether these studies describe the reduced facility of circulating monocytes for antigen presentation via MHC class II, and/or the effect of an expanded M-MDSC population.

It may be that differential enrichment of key biological functions in monocytes in cancer may be attributable to monocyte subset-wise profiles, or change in monocyte subset

proportion. However, monocyte subset-wise profiling in the context of cancer has been limited. Expansion of the non-classical monocyte compartment has been documented in cancer patients[176]. For example, Cassetta et al.[176] found that the peripheral non-classical population was expanded in breast and endometrial cancer patients. Similarly, Chavan et al.[192] demonstrated a significantly smaller pool of classical ($CD14^{+}CD16^{-}$) monocytes in patients with MM versus healthy controls. These studies conflict with a broadly more immature and classical-type monocyte expression profiles in cancer observed in bulk RNA-seq studies. Further work exploring monocyte subset-specific populations in cancer would be of value in elucidating these profiles and their associated function.

The distinct transcriptomic profiles in peripheral monocyte populations have been shown to have potential diagnostic utility. Hamm et al.[193] identified a monocyte expression profile with robust diagnostic utility in colorectal cancer. However, the clinical applicability of such studies has yet to be fully explored. Existing transcriptomic studies, including those in melanoma[73], have used low-throughput approaches within small cohorts[155].

Function	Example references
Polarisation of hemopoiesis towards myeloid population and expansion of peripheral monocyte populations	Wu et al.[175]; Kiss et al.[155]
Upregulation of chemotaxis pathways	Cassetta et al.[176]; Ramos et al.[177]; Qian[178]; Ugurel et al.[179]
Reduced response to hypoxia	Cassetta et al.[176]
Altered metabolic pathways	Cassetta et al[176]; Kireg et al.[73]
Expansion of M-MDSC (or M-MDSC-like) populations	Gustafson et al.[187]; Gonda et al.[188]; Sun et al.[194]; Meyer et al.[191]; Ugurel et al.[179]
Impaired cytotoxicity function and phagocytosis	Gordon et al.[180]
Polarisation towards an immature phenotype	Goudot et al.[136], Almand[181]
Induction of anti-viral signalling pathways	Ramos et al.[177]

Figure 4: Summary of key cancer-associated circulating monocyte profiles

1.4.4 Myeloid populations and ICB

There is an expanding body of evidence demonstrating that myeloid populations have utility as predictive markers in patients treated with ICB[195]. Here, key findings regarding the association of circulating monocytes with modulation of ICB are described.

Circulating myeloid cells have predictive utility in the context of both anti-PD-1 and anti-CTLA-4 therapy. Baseline monocyte count has been shown to be predictive in the

context of ICB. For example, Ye et al.[40] found that baseline monocyte count was associated with poor prognosis in patients with MM subsequently treated with ICB. Similarly, Chasseuil et al.[196] found that an elevated baseline monocyte count was associated with reduced OS and shorter PFS following anti-PD-1 therapy in melanoma[196]. Soyano et al.[197] also found that elevated baseline monocyte count was associated with relapse in patients with advanced NSCLC treated with anti-PD-1 therapy. Baseline monocyte count was negatively associated with OS, mortality and disease progression[197]. Strauss et al.[198] also recently confirmed the key role of peripheral myeloid populations in modulating ICB responses. They found that a murine model characterised by myeloid-specific loss of PD-1 exhibited increased T effector memory cell population with anti-tumour effect, in spite of retained lymphoid PD-1 expression[198]. This demonstrates the critical role that myeloid populations have in modulating response to ICB.

M-MDSC populations are robustly predictive of clinical response to ICB. At baseline, a smaller M-MDSC population is associated with more favourable clinical response to both anti-PD-1 and anti-CTLA-4 ICB in melanoma[199,200]. For example, Gibney et al.[200] found that lower baseline levels of M-MDSCs was associated with reduced risk of relapse following nivolumab treatment. Similarly, a small $CD14^{+}HLA^{-}DR^{low/-}$ M-MDSC population was associated with reduced relapse-free survival in advanced melanoma following ipilimumab treatment[201]. ICB-associated modulation of MDSC populations has also been demonstrated. deCoaña et al.[202] reported a reduction in the proportion of M-MDSCs with ipilimumab in patients who responded to treatment, whilst expansion of post-treatment M-MDSC populations was associated with reduced survival. As previously outlined, some of these studies explore the predictive role of $HLA\text{-}DRA^{low}$ populations without more detailed characterisation to confirm M-MDSC phenotype. For example, Meyer et al.[191] reported higher levels of $HLA\text{-}DR^{-}$ monocytes in those with MM versus healthy controls, and HLA-DR- monocytes were expanded in non-responders to ipilimumab. Whether these studies reflect expansion of a true MDSC population in cancer patients with prognostic utility, or a polarisation towards a MDSC-like $HLA\text{-}DRA^{low}$ phenotype in a true monocyte population, is difficult to determine.

Whilst circulating myeloid populations have typically been found to be predictive of poor prognosis, evidence is conflicting. Krieg et al.[73] used mass spectrometry to explore PBMC subsets in a melanoma cohort treated with anti-PD-1 therapy and found the peripheral $CD14^{+}CD16^{-}$monocyte population to be associated with a favourable clinical response[73]. Whilst the $CD14^{-}CD16^{+}$ population was elevated in patients compared to healthy controls, it was not associated with clinical outcome[73]. However, selection of only $CD14^{+}$HA-DR^{hi} classical monocyte populations and HLA-DR^{lo} non-classical monocytes for this analysis are likely to have removed a $CD14^{+}$HLA-DR^{low} population associated with a MDSC phenotype from this analysis, which are typically associated with poor prognosis.

Despite the diversity of monocyte populations, work characterising the role and modulation of monocyte subsets with ICB is limited. Non-classical monocytes have been shown to have predictive utility in the context of anti-CTLA-4 therapy. Romano et al.[203] found that in patients with melanoma, a higher baseline peripheral non-classical $CD16^{+}$ monocyte population was associated with response to ipilumumab[203], with stimulation of ADCC-mediated T_{reg} lysis modulated by $CD16^{+}$, but not $CD16^{-}$ monocytes[195,203]. Further work characterising subset-wise roles in this context may potentially be insightful, particularly given the diverse phenotypic and functional roles within the monocyte compartment.

1.4.4.1 scRNA-seq approaches to characterising the myeloid response to ICB

There has been limited work utilising scRNA-seq approaches to characterise myeloid responses to ICB. Gubin et al.[204] utilised high dimensional profiling techniques CyTOF and scRNA-seq to assess the response to ICB in subpopulations of intratumoural haemopoietic cells in mice. In a mouse sarcoma model, macrophage and monocyte populations were found to undergo dynamic changes following ICB. Using scRNA-seq they identified a tissue-based myeloid subset 'Mac_s1' characterised by expression of *CCR2, ITGA4* and *ILR2,* which decreased in size with cICB, but increased with sICB. They also identified an anti-inflammatory subset characterised by expression of *CXCR3, CCL2* and *MRC1* and oxidative phosphorylation signatures. The proportion of $CD45^{+}$ cells that this anti-inflammatory subset ('Mac_s2') comprised decreased following ICB treatment[205]. Three myeloid subsets, Mac_s3, Mac_s4 and Mac_s5, expanded following ICB, more prominently with cICB. This included one subset (Mac_s5) which exhibited an

IFN-γ and NFκB signalling response. This demonstrates that ICB may be associated with dynamic subset-wise changes in the myeloid compartment. However, a focussed analysis of wither peripheral or tumour-based monocytes and ICB responses in humans has not yet been presented. This will be explored in this thesis.

1.4.5 Summary

Monocytes are a heterogeneous population characterised by diverse range of biological functions, and transcriptional and antigenic profiles. This introduction outlines the key role that both monocyte and macrophage populations have in modulating tumourigenic mechanisms, and their potential clinical utility as diagnostic, prognostic and predictive biomarkers across a wide range of cancers including MM, and also within the context of ICB. Despite this, characterisation of circulating monocyte profiles in large melanoma cohorts is lacking, and is description of on-treatment modulation with ICB. In view of the diversity exhibited within conventional monocyte populations and the robust association with clinical outcome, there has been surprisingly little exploration of subset-wise responses in the context of malignancy and ICB. scRNA-seq techniques provide new opportunities for high-resolution analysis of monocyte subsets and treatment responses.

1.5 Cell-free DNA

Cell-free DNA (cfDNA) are short double-stranded fragments of DNA within body fluid including plasma or urine[206,207]. Whilst most cfDNA in plasma is derived from leukocytes, cfDNA is also derived from a wide range of tissues, including predominantly erythrocyte progenitors, endothelial cells and hepatocytes[208]. cfDNA is typically 160-180 base pairs (bp) in length[209,210], corresponding to DNA wrapping around a nucleosome core[210,211]. The half-life of cfDNA is approximately a few minutes to 1-2 hours, allowing for near real-time genomic profiling[212].

In patients with cancer, a significant proportion of cfDNA may be tumour-derived (circulating tumour DNA, 'ctDNA')[206]. This can range from less than 0.1% to more than 90% of the total cfDNA component, with ctDNA concentration associated with disease burden[213]. Analysis of ctDNA has shown promise as a diagnostic and prognostic biomarker, and for monitoring on-treatment response. Elevated ctDNA levels are associated with poor prognosis in patients across cancer types including gastro-oesophageal adenocarcinoma[214], melanoma[215] and NSCLC[216]. In addition, ctDNA has shown utility for monitoring minimal residual disease (MRD) in cancers including lung, breast and colon cancer[206]. Tie et al.[217] found that detectable ctDNA following surgery for patients with stage II colon cancer was a significant predictor of relapse both with and without adjuvant chemotherapy. ctDNA concentration has also been shown to be predictive of response in patients treated with ICB[218]. Bratman et al.[218] reported that baseline ctDNA concentration was associated with PFS in patients treated with pembrolizumab, with depletion of ctDNA following ICB associated with sustained clinical benefit.

Mutational ctDNA profiles have been shown to be concordant with mutational landscapes within solid tumours[219,220]. Tissue core biopsy is the gold standard for assessment of tumour parameters. However, biopsy sampling is associated with risk of local injury, and in some instances may not be clinical feasible. In addition, core biopsy of conventional tumour sampling may not represent mutational heterogeneity[221]. Analysis of ctDNA may therefore better facilitate representative sampling of tumour

DNA and mutational profiles. Assessment of ctDNA as a 'liquid biopsy' overcomes many of these issues.

Variant analysis in ctDNA has shown promise as a diagnostic and predictive biomarker. McEvoy et al.[222] demonstrated that detection of mutations in genes such as *BRAF, NRAS* and *TERT* in ctDNA were predictive of subsequent relapse in patients with melanoma. ctDNA has also been utilised to monitor cancer evolution and development of resistance to treatment[223]. Murtaza et al.[223] performed exome sequencing of cfDNA at intervals throughout the treatment of patients with lung, breast and ovarian cancer, and identified acquisition of mutations in genes including *PIK3CA, RB1* and *MED1,* associated with treatment resistance. Cutts et al.[224] used exome sequencing of cfDNA to monitor mutational evolution in melanoma following vemurafenib and ipilimumab treatment, and observed on-treatment mutational selection[224]. Monitoring single-gene mutations may therefore guide treatment selection and monitor on-treatment response[206].

1.5.1 Limitations of cfDNA assessment

Despite the clear opportunities provided by cfDNA for clinical diagnosis, prognostication, guiding treatment and monitoring response in patients with cancer, there are notable limitations[206]. ctDNA is typically present only at low concentrations, particularly in those with low tumour burden and at early disease stage [213]. In some patients ctDNA may not be detectable and rare tumour mutations not readily identified[206]. cfDNA is typically short and fragmented which provides technical challenges for library preparation and sequencing[213].

Whilst ctDNA quantification and variant assessment has a promising role in diagnosis and prognostication, profiling 5-methylcystine (5mC) and 5-hydroxymethylcytosine (5hmC) cfDNA modifications may offer an alternative approach for non-invasive diagnosis and prognostication.

1.6 Methylation

Methylation of cytosines at the 5' position is a key mechanism through which gene expression is regulated. Methylation is typically enriched within CG-dense regions,

termed CpG island (CGI) sites[225]. Most promoter regions contain CGIs, and whilst between 60-80% of the 28 million CpG sites in humans exist in a methylated state[226], those within promoter regions are typically unmethylated[227]. Methylation at CGI sites within promoter regions facilitate regulation of gene expression. DNA methylation can affect transcription by either directly inhibiting transcription factor binding, or facilitating the binding of methyl-binding proteins (MBPs) which are then able to bind to chromatin re-modellers which can stimulate or block transcription[228–230]. Methylation is also associated with heterochromatin, a marker of transcriptional inactivity and repression[231]. Although the role of methylation is typically repressive with methylation at promoter regions associated with gene silencing[232], within gene bodies, DNA methylation is typically permissive of gene transcription[232,233].

Methylation is a dynamic modification[234] and plays an important regulatory role across a wide range of biological processes including development[232], haematopoiesis[235] and immune responses[236]. For example, methylation regulates $CD8^+$ T cell and macrophage differentiation[236]. Methylation has also been shown to regulate immune activation and phenotype. For example, loss of methylation in T cells at the *FOXP3* gene promoter is associated with acquisition of a T_{reg} phenotype[237].

DNA methylation is regulated via DNA methyltransferase enzymes (DNMT1, DNMT2 and DNMT3) which facilitate methylation of unlabelled 5' cytosines[238]. Whilst DNMT1 commonly acts at hemi-methylated CpG sites, DNMT3 (DNMT3A and DNMT3B) act in an unbiased fashion, methylating both unmethylated and partially methylated regions[231,239]. Typically, DNMT1 acts to preserve and maintain methylation profiles during cell division, whilst DNMT3A and DNMT3B establish DNA methylation patterns during development[238,239]. Methylation is also regulated by the actions of methyl CpG binding proteins and regulators of chromatin remodelling[234].

1.6.1 Methylation and cancer

Aberrant methylation is a key feature of malignancy[232]. Cancer is associated with global hypomethylation, which contributes to genomic instability, driving tumourigenesis[231,234]. Localised hypomethylation commonly occurs in malignancy[240] whilst methylation of tumour suppressor genes can lead to inactivation and constitutive

deregulation of pro-tumourigenic pathways such as mitogenic and growth signalling pathways[231,241].

Methylation of the promotor region within tumour suppressor genes is a key driver of tumour development[238]. For example, tumour suppressor genes including *PTEN, BRCA1/2* and *RB* acquire loss-of-function mutations as a key step in tumourigenesis, but also may exhibit aberrant promoter hypermethylation and silencing[238]. Some genes are infrequently mutated, but commonly undergo inactivation via promoter hypermethylation, such as the *MGMT* gene involved in DNA repair[242], and cell cycle regulator *CDKN2B*[238,243]. Similarly, silencing of the *MLH1* MMR gene via promoter methylation leads to MSI and drives tumorigenesis in a subset of colorectal and endometrial cancers[244,245]. Mutational events can also contribute to aberrant methylation profiles in malignancy. For example, mutations in enzymes responsible for maintaining physiological methylation patterns such as DNMT[246], TET enzymes[247], or IDH1/2[248] are reported in cancers including colorectal cancer[238,249] and glioma[250,251]

In view of the relative stability and abundance of DNA methylation within the genome and tissue-specific nature of methylation profiles, DNA methylation markers are attractive diagnostic, prognostic and predictive candidate biomarkers[252]. There are a number of tumour-based methylation-based assays which are used in clinical practice. These include tests which examine the methylation of genes including *APC, GSTP1* and *RASSF1* in prostate cancer diagnostics, *SEPT9* in colorectal cancer, and *MGMT* to predict response to temozolamide in patients with glioblastoma[252].

1.6.2 Methylation cfDNA profiles

cfDNA exhibits epigenetic modification reflective of tissue-of-origin[208]. 5mC cfDNA profiles are promising diagnostic biomarkers and there are many recent studies describing their utility in this context across a range of cancer types including pancreatic[253], intracranial[254], renal[255,256], HCC[257], ovarian[258], lung[259] (see Figure 5). A key challenge in epigenomic profiling of cfDNA is the sensitivity of conventional epigenetic profiling techniques. cfDNA typically exists at low concentrations, and conventional bisulphite sequencing (BS) techniques involve loss of a large proportion of input DNA[260]. The development of new approaches such as immunoprecipitation-based methods

make assessment of cfDNA 5mC profiling feasible without bisulfite degradation[261]. These techniques have been shown to be of use in the detection of even early-stage disease. For example, in 2018, Shen et al.[261] developed a cfDNA immunoprecipitation protocol 'cfMeDIP' which is suitable for low input cfDNA quantities via use of a carrier lambda DNA amplicon. They utilised this method to show that that cfDNA methylation profiles could be used to detect early-stage cancer across a range of tumour types including breast cancer, lung cancer, renal cell carcinoma (RCC) and acute myeloid leukaemia (AML)[261].

The utility of cfDNA methylation profile analysis is not limited to tumour detection, but also in determining disease stage and distinguishing between pathological subtypes. Nuzzo et al.[255] utilised cfMeDIP to analyse methylated cfDNA profiles in both the plasma and urine of patients with RCC. This approach was found to be a sensitive method for detecting patients with RCC or urothelial carcinoma versus healthy controls, but also for classifying patients according to disease stage. Nassiri et al.[254] used cfMeDIP to demonstrate that plasma cfDNA methylation profiles could be used for diagnosis of intracranial tumours and successfully distinguish between different intracranial tumour types, which can be challenging on imaging. Recently Lasseter et al.[256] used cfMeDIP to show that 5mC cfDNA profiling was a more sensitive diagnostic marker than cfDNA variant detection in RCC.

Whilst most studies have focussed on exploring the epigenetic cfDNA profiles for diagnostic purposes, there is emerging evidence to show that 5mC profiling may have utility as a prognostic and predictive biomarker. In particular, studies demonstrate that 5mC profiles are prognostic in patients with DLBCL[262], and predictive of response to chemotherapy in patients with ovarian cancer[258].

Analysis of cfDNA methylation profiles provides a unique opportunity for non-invasive diagnostics, as well as for guiding treatment and deducing prognosis. The body of evidence is still rapidly expanding, yet to date there has been little or no work exploring on-treatment profiles epigenetic cfDNA modifications in patients with cancer. Although the importance of methylation in regulating immune function is recognised, and a large proportion of cfDNA is leukocyte-derived, there is little work exploring immune profiles or ICB-responses via cfDNA epigenetic profiling.

1.7 5-hydroxymethylcytosine

5-hydroxymethylcytosine (5hmC) is an epigenetic DNA modification first identified by Wyatt and Cohen in 1952[226,263]. Over the next 50 years, it was largely ignored by the scientific community. However, in 2009, Kriaucionis and Heintz[264] demonstrated the presence of 5hmC in Purkinje neurones in the cerebellum using thin-layer chromatography. Simultaneously, Tahiliani et al.[265] demonstrated that 5hmC was generated from the oxidation of 5mC by the ten-eleven translocation (TET) family of enzymes. Whilst DNA can undergo demethylation in a passive manner during DNA replication in the absence of the DNA methyltranferase DNMT1, it can also undergo active demethylation via oxidation by TET enzymes forming 5hmC[266]. Subsequent intermediates in the active replication-independent oxidation process include 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Thymine-DNA glycosylase (TDG) is able to remove 5caC from DNA[267], whilst base excision repair then generates an unmodified cytosine[268]. It was thought that 5hmC exists at too high a level to be a merely a non-functional modification in the demethylation process[269] and is noted to be relatively stable[270]. Whilst approximately 10% of cytosines exist in a methylated state, between approximately 0.1-0.7% of the genome is hydroxymethylated[266,271]. This is far in excess of the frequency of 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), which are found at $20x10^{6}$% and $3x10^{6}$% of all cytosines respectively[271].

1.7.1 5hmC distribution and function

5hmC was first identified and characterised within the CNS and embryonic stem cells (ESCs), but 5hmC labelling has been found across all mammalian organ systems[270,272,273]. Tissue type is one of the key variables affecting 5hmC enrichment[274], and it has been suggested that the relative distribution and 5hmC profiles across tissues may infer cell of origin or tissue-type[271,275]. 5hmC is found at highest levels in the mammalian brain, and in particular within neuronal cells[264,268], but is also found at high levels in the kidney, colorectal tissue and liver (0.4-0.65% cytosines), whilst organs including lung, heart, breast, blood and placenta are more poorly enriched for 5hmC (0.05-0.06% cytosines)[275–277].

5hmC is typically enriched within the gene body and depleted at intergenic regions[275,278,279]. Enrichment is typically seen within promoter and enhancer regions and depleted at the transcriptional start site (TSS)[278–281] keeping with a role in regulation of gene expression. 5hmC is broadly a marker of gene activation, with higher levels of 5hmC found within actively transcribed genes[275,278,280,281]. Some proteins directly bind the 5hmC modification to affect gene transcription[276]. For example, Mecp2 Is found at high levels within the brain and binds directly to 5hmC to promote gene expression in neuronal cells[280].

Another key way in which gene expression is regulated is via post-translational histone modification. 5hmC is associated with histone modifications indicative of active transcription, euchromatin and enhancer regions, such as H3K4me1 and H3K27ac[279,282]. A key role for 5hmC in embryonic development and differentiation has also been described[283]. Serandour et al.[284] demonstrated that 5hmC enrichment was required for neuronal and adipocyte differentiation in mice. Williams et al.[285] found that 5hmC was enriched at TSS in ESCs and that 5hmC enrichment comprised the initial step in enhancer activation[284]. 5hmC and the TET enzymes family have also been shown to have a role in regulating immune function and activation. Reduced TET activity in HSCs with associated 5hmC depletion has been shown to divert differentiation towards myeloid populations[286]. 5hmC is also enriched at sites associated with transcription factors involved in regulating differentiation of naïve $CD4^+$ T cells into Th1, Th2 and T_{reg} cells[287,288].

Despite a broadly positive relationship between 5hmC enrichment and gene expression, this association is gene-specific. For example, Lin et al.[276] characterised 5hmC enrichment and gene expression across CNS and liver tissue in mice, and found that >75% of genes lacking 5hmC labelling were expressed; including mitochondrial and ribosomal genes. 5hmC was also poorly enriched within sex chromosomes[276].

Dynamic modification of 5hmC has been observed following environmental exposures[289]. Thomson et al.[290] described dynamic changes in both 5mC and 5hmC profiles in the liver of mice treated with the carcinogen phenobarbital to induce tumorigenesis, with rapid alteration in 5hmC levels and demethylation occurring in tandem with the acquisition of 5hmC at the same sites[291]. In activated genes, 5hmC

levels increased within promotor regions, with a reduction in 5mC. This suggests that 5hmC could have potential utility as a biomarker of a variety of treatment responses[289]. However, limited work has been performed to explore on-treatment changes in 5hmC profiles in response to drug and other environmental stimuli, including in the context of ICB.

1.7.2 5hmC in cancer

5hmC has been shown to be a potential diagnostic tumour biomarker. In malignancy, global hypo-5-hydroxymethylation is typical and has been reported in cancers including melanoma, lung, colon, kidney and brain[275,289,292–294]. There have been many studies demonstrating 5hmC depletion in melanoma. Lian et al.[295] used IHC and immunofluorescence (IF) to demonstrated prominent nuclear 5hmC staining in benign melanocytic lesions such as naevi, whilst 5hmC was depleted in malignant melanoma. Loss of 5hmC was associated with disease progression and poor prognosis, with higher 5hmC-staining associated with greater likelihood of survival[295].

1.7.3 5hmC cfDNA profiles

5hmC modification is present within cfDNA. As previously discussed, cfDNA represents DNA derived from predominantly leukocytes, but in patients with cancer may be tumour-derived. 5hmC cfDNA profiles have therefore been explored as potential cancer biomarkers (Figure 5). In a proof-of-concept study, Song et al.[296] demonstrated the potential utility of 5hmC cfDNA profiles as a diagnostic cancer biomarker, characterising cfDNA 5hmC profiles across seven cancer types, and identifying distinct 5hmC cfDNA profiles in cancer patients versus healthy controls.

Many groups have successfully utilised the hMe-Seal approach[296] to characterise 5hmC cfDNA profiles and demonstrate the utility of 5hmC as a potential diagnostic biomarker[296–299] (see Figure 5). Distinct 5hmC cfDNA profiles with diagnostic utility have been demonstrated in cancers including PDAC[253,297], colorectal[298,299], gastric[298], liver[298], thyroid[298], oesophageal[300], NSCLC[301] and hepatocellular carcinoma (HCC)[302] (Figure 5). In pancreatic cancer, Guler et al.[297] identified thousands of differentially 5-hydroxymethylated regions (DhMRs) in cfDNA in patients with PDAC versus healthy controls within genes related to the pancreas and cancer development. These DhMRs

were used to develop a regression model with a AUC of 0.92-0.94[297], demonstrating the utility of 5hmC cfDNA profiles as a potential diagnostic biomarker.

Studies have also demonstrated the utility of 5hmC cfDNA profiles as prognostic biomarkers. For example, Chen et al.[303] recently demonstrated that 5hmC cfDNA profiles were predictive of response to R-CHOP in patients with DLBCL (Figure 5). To date, there are no studies that have described the predictive role of 5hmC cfDNA profiles in the context of ICB, nor that have described on-treatment 5hmC cfDNA modification. This will be explored in this thesis.

Study	Cancer type	Modification	Study aim	Epigenetic sequencing approach	Comments
Cao et al. [253]	PDAC	5hmC/5mC	Diagnostic	cfMeDIP	
Nassiri et al. [254]	Intracranial	5mC	Diagnostic	cfMeDIP	
Nuzzo et al. [255]	RCC	5mC	Diagnostic	cfMeDIP	5mC cfDNA as a urinary biomarker
Guler et al. [297]	PDAC	5hmC	Diagnostic	hMe-Seal	
Li et al. [298]	Colorectal, gastric, PDAC, liver, thyroid	5hmC	Diagnostic	hMe-Seal	
Song et al. [296]	Lung, HCC, PDAC	5hmC	Diagnostic	hMe-Seal	
Tian et al. [300]	Oesophageal	5hmC	Diagnostic	nano-hmC-Seal	
Zhang et al. [301]	NSCLC	5hmC	Diagnostic	nano-hmC-Seal	
Gao et al. [299]	Colorectal	5hmC	Diagnostic	hMe-Seal	
Chiu et al. [304]	DLBCL	5hmC	Prognostic	nano-hmC-Seal	
Cai[302]	HCC	5hmC	Diagnostic	nano-5mC-Seal	
Holmila et al.[257]	HCC	5mC	Diagnostic	BS across a small gene panel	
Widschwendter et al.[258]	Ovarian	5mC	Diagnostic, predictive	Methylation bead chip array and RRBS	5mC cfDNA profiling as a diagnostic marker, and predictive of response to chemotherapy.
Barault et al.[305]	Metastatic colorectal cancer	5mC	Predictive	Genome-wide methylation arrays	
Gallardo-Gomez et al.[306]	Colorectal carcinoma	5mC	Diagnostic	MethylationEPIC array	
Chen et al. [303]	DLBCL	5hmC	Predictive	hMe-Seal	5hmC cfDNA profiles predictive of response to R-CHOP.
Kristensen et al.[262]	DLBCL	5mC	Prognostic	Promoter methylation of a small gene panel with BS	5mC cfDNA profiles are poor prognostic marker in DLBCL
Takahashi et al.[307]	Breast	5mC	Predictive	Methylation-specific PCR (single gene)	5mC cfDNA assessment of RASSF1A promoter was associated with neoadjuvant chemotherapy response
Lasseter[256]	RCC	5mC	Diagnostic	cfMeDIP	Demonstrated that cfMeDIP-seq was more sensitive as a diagnostic marker than variant detection in ctDNA, but that variant detection is preferable for monitoring on-treatment response.
Hulbert et al. [259]	Lung	5mC	Diagnostic	Methylation-specific PCR	Promoter methylation of a limited gene panel in sputum and plasma

Figure 5: Summary of key publications exploring epigenetic cfDNA profiles as cancer biomarkers

1.8 Epigenetic sequencing assays

Whilst the potential of epigenomic cfDNA profiling as a diagnostic and prognostic biomarker has been recognised, there are a lack of prospective studies confirming clinical utility, and reproducibility is also problematic. This may be in part due to the variation in assays used by different groups. It is therefore pertinent to briefly consider the different epigenetic sequencing approaches. Common approaches to profiling the epigenome landscape DNA include chemical conversion with bisulphite, immunoprecipitation-based methods and those that involve labelling with β-glucosyltransferase.

Bisulphite sequencing (BS) methods utilise the conversion of unmethylated cytosines to uracil[308]. Following PCR, this conversion is then detected as thymine during sequencing, and the percentage of methylated cytosines can be determined during post-sequencing analysis. Oxidative bisulfite sequencing ('OxBS-seq') allows for discrimination between 5mC and 5hmC[309]. Whilst BS-based methods provide high levels of single-base resolution, per-sample costs are typically relatively high and between 84-96% of input DNA is lost[310], which poses problematic in low-input samples such as cfDNA. In addition, bisulphite conversion can cause issues with alignment in following sequencing[311].

Another key approach for epigenomic cfDNA profiling is the utilisation of methylated DNA immunoprecipitation-based techniques ('MeDIP'). These methods utilise capture of 5mC or 5hmC-labelled DNA with a mAb[312]. This method has the advantages of 5mC-specificity, and a lower per-sample cost than BS[313]. Conversely, as MeDIP involves pull-down of 5mC labelled DNA without marked loss of DNA fragments, success can be achieved with relatively low input DNA amounts. Shen et al.[261] developed an approach which allows cfDNA inputs as low as 1 ng, and which relies on the addition of a lambda DNA filler to increase the DNA quantity prior to immunoprecipitation of 5mC enriched DNA. This method produces broadly comparable results to those obtained with BS, but with much lower input DNA concentrations[314]. However, immunoprecipitation-based methods have lower base resolution than BS, and care must be taken to check antibody specificity to prevent non-specific binding and pull-down[313]. In addition,

immunoprecipitation-based methods exhibit preferential sequencing of hypermethylated regions[313].

The popularity of hMe-Seal is also notable and has been used by the majority of studies exploring 5hmC cfDNA profiles as diagnostic biomarkers[296–299,303]. This technique involves labelling 5hmC with an azide-modified glucose via β-glucosyltransferase and subsequent pull-down on streptavidin beads[296]. Captured DNA is then amplified following clean-up. This enrichment-based method provides genome-wide coverage with high sensitivity and specificity[315], and may exceed the specificity available with immunoprecipitation-based enrichment techniques[316]. Nano-hmC-Seal is a method developed by Han et al.[282] and used successfully for 5hmC sequencing with very low input DNA quantities[300–302,304].

1.8.1 Summary

cfDNA provides new opportunities for non-invasive diagnostics, prognostication and for monitoring response to treatment and relapse, both via monitoring ctDNA quantities, and by tumour-associated variant detection. 5hmC and 5mC epigenetic modifications have distinct roles in regulating gene expression and cellular activity. cfDNA, which is typically predominantly leukocyte derived but which may represent ctDNA in patients with cancer, exhibits epigenetic modification. The potential of 5hmC and 5mC cfDNA profiles as diagnostic and prognostic biomarkers is established. Although challenges remain, new epigenetic library preparation techniques now facilitate genome-wide assessment of the cfDNA epigenome allowing for low cfDNA input quantities. Despite this, characterisation of epigenetic cfDNA profiles in melanoma are lacking but potentially may have diagnostic, prognostic and predictive utility. Additionally, ICB-dependent modulation of epigenetic cfDNA profiles have yet to be characterised. As ICB is associated with prominent transcriptomic activation in circulating immune cells, it may be that characterisation of epigenetic cfDNA profiles reflect immune activation and may correlate with clinical response.

1.9 Conclusions

This introduction highlights that whilst ICB has undoubtedly revolutionised the treatment of patients with MM and other cancer types, clinically-applicable predictive markers are limited and tissue-based. This provides limitations in terms of clinical feasibility, and issues associated wtith tumour heterogeneity. Peripheral predictive markers of clinical response to ICB are required.

Myeloid populations have emerged as key determinants of ICB treatment-response and clinical outcome in patients with cancer. Despite this, melanoma-associated peripheral monocyte phenotypic profiling has yet to be performed across large prospective patient cohorts. Whilst the predictive utility of monocyte populations is evident, on-treatment characterisation of circulating monocyte responses with ICB is limited. Profiling of myeloid populations may characterise mechanisms of immune modulation, as well as identify prognostic and predictive markers of response to ICB in this patient group.

The novel opportunities afforded by characterisation of the cfDNA as a 'liquid biopsy' are also clear. Epigenetic modifications 5hmC and 5mC may both represent and regulate immune function and tumourigenesis. In recent years, epigenetic cfDNA profiles have been shown to have diagnostic utility in patients with cancer. In melanoma, the predictive utility of these markers, and on-treatment modulation with ICB, have not yet been characterised.

1.10 Key aims

The key aim of the work in this thesis was to characterise peripheral responses to ICB, with a view to identifying markers that might have clinical applicability. In view of the promising, but yet to be fully characterised, role that circulating myeloid profiles have in ICB modulation, and the emerging evidence demonstrating the diagnostic and prognostic utility of epigenetic cfDNA, I decided to explore the circulating myeloid compartment, and epigenetic cfDNA profiles. Broadly, my key questions were as follows.

- Do circulating monocytes exhibit distinct MM-associated profiles?
- Do circulating monocytes exhibit on-treatment modulation following ICB?

- Do circulating monocytes exhibit transcriptomic correlates of clinical response to ICB?
- Can scRNA-seq be used to provide high-resolution characterisation of monocyte subsets?
- Do circulating monocytes exhibit subset-wise modulation in MM and ICB?
- Are epigenetic cfDNA profiles (both 5mC and 5hmC) distinct in MM patients and are these modulated in ICB?
- Do epigenetic cfDNA profiles have predictive utility in MM patients treated with ICB?

In the following chapters I present my experimental findings addressing these key questions.

2 Chapter 2: Methods

2.1 Samples

2.1.1 Patient cohort

Peripheral blood samples were obtained from patients >= 18 years old with metastatic melanoma (MM) treated with immune checkpoint blockade (ICB). This included patients treated with ipilimumab/nivolumab (anti-CTLA-4/PD-1) combination immune checkpoint blockade (cICB), and with single-agent pembrolizumab or nivolumab (anti-PD-1) immune checkpoint blockade (sICB). All patients provided prior written informed consent for samples to be donated to the Oxford Radcliffe Biobank (09/H0606/5 + 5, Oxford Centre for Histopathology Research ethical 250 approval 16/A019, 18/A064) to be used for research (n = 180). Peripheral blood samples were obtained prior to the first treatment dose (cycle 'C1'), and then just before the second dose (cycle 2, C2), and then fourth treatment dose (cycle 4, C4). ICB treatment regimes were according to standard treatment protocols. For patients receiving cICB this involved administration of ipilimumab and nivoumab every three weeks for up to four doses. For patients treated with sICB, pembrolizumab was administered once every three weeks, or every four weeks for nivolumab. Up to 50 ml of peripheral blood was collected per sample in ethylenediaminetetraacetic acid (EDTA) tubes.

2.1.2 Control samples

Control samples were obtained from the Oxford biobank after informed written consent was obtained. Samples were collected with local ethical approval (REC 06/Q1605/55).

2.1.3 Clinical outcome data

Clinical outcome data across the cohort was obtained, including parameters such as progression at six months, time-to-progression, overall survival (OS), progression free survival (PFS) and death status. The development of immune-related adverse events ('irAEs') prior to treatment cycle 5 was noted as per the National Cancer Institute Common Terminology Criteria for Adverse Events (CTCAE)[317]. OS was defined as the time

between the start date of ICB treatment to death, whilst PFS was defined as time between the first ICB dose and disease progression according to either death or disease progression on imaging.

2.2 Peripheral blood mononuclear cell (PBMC) and plasma isolation

PBMCs were isolated from peripheral blood from patients, typically within 1-2 hours. Whole peripheral blood was diluted in an equal volume of Hanks' balanced salt solution (HBSS) (BioWhittaker) at room temperature. The blood was then slowly layered over 12.5 mls of Ficoll (GE Healthcare) in a 50 ml Falcon tube at room temperature. The samples were centrifuged at 400g for 30 minutes at room temperature with slow acceleration and deceleration off. Plasma (diluted 1:1 in HBSS) was collected from the separated whole blood and then centrifuged at 16,000g for 20 min at 4°C. Plasma was stored at -80°C.

PBMCs were isolated from the separated blood and placed in a 50 ml Falcon tube. From this point PBMCs were kept on ice. Cold HBSS was added to the PMCSs to a total volume of 50 mls. The PBMCs were spun at 300 g for 10 mins at 4°C in a centrifuge with acceleration and deceleration set to maximum. The supernatant was then removed and the PBMC pellet was resuspended in 30 ml of HBSS on ice. The PBMCs were then centrifuged again at 300 g for 10 mins at 4°C with acceleration and deceleration set to maximum. The supernatant was then removed and the pellet resuspended in 20 mls of HBSS on ice .10 μl of PBMC suspension was taken and mixed with 10 μl of trypan blue (Sigma Life Science) were quantified using a haemocytometer (Hausser Scientific). PBMCs for magnetic bead sorting were taken and topped up to a final volume of 30 mls with cold HBSS and spun once more for 10 mins at 300 g at 4°C.

To separate PBMCs of interest the MACS magnetic bead system (Miltenyi Biotec) was used according to manufacturer's protocols. MACS buffer was prepared by mixing 50 ml HBSS with 250 μl FCS (Gibco) and 200 μl EDTA (Sigma-Aldrich). The PBMCs were resuspended in the appropriate volume of MACS buffer (80 μl per 10^7 PBMCs). To isolate $CD14^+$ cells, 20 μl per 10^7 cells of $CD14^+$ microbeads (Miltenyi Biotec) were added to the PBMCs. The beads and PBMCs were mixed by pipetting, and then samples were

incubated for 15 mins at 4°C. Following incubation, 1 ml of MACS buffer was added to each sample and the sample was spun for 10 mins at 300 g at 4°C. Following spinning the supernatant was removed and the pellet was resuspended in 500 μl of MACS buffer. The cell suspension was then applied to the MS column (Miltenyi Biotec). The sample was left to run through the column with flow-through collected for negative selection of other PBMC cell types. The column was washed three times with 500 μl MACS buffer. The PBMCs were flushed from the column with 500 μl of MACS buffer into a 1.5 ml microtube. The cells were then quantified the Countess II FL Automated Cell Counter (ThermoFisher Scientific) according to manufacturer's protocol. Cells were then spun at 300 g for 10 mins at 4°C. The supernatant was removed and then cells were lysed by adding 350 μl of either buffer RLT (Qiagen) with 1% beta-mercaptoethanol, or buffer RLT with dithiothreitol (DTT) (Sigma Aldrich). The lysed cell sample was then stored at -80°C.

2.3 Bulk RNA-Sequencing

2.3.1 RNA Extraction

PBMC samples isolated using the MACS magnetic bead system were lysed in RLT and stored at -80°C. The QIAshredder (Qiagen) was used to homogenise samples, and RNA extraction was carried out with the AllPrep DNA/RNA/miRNA kit (Qiagen) according to manufacturer's protocols[318]. 35 μl of RNAase free water was used for RNA elution. RNA was quantified using Qubit RNA HS Assay kit (ThermoFisher Scientific). RNA samples were stored at -80°C until required for library preparation and sequencing.

2.3.2 RNA sequencing

Bulk RNA-seq libraries were sequenced in three batches. The first RNA cohort was sequenced using 75 bp paired-end sequencing on an Illumina Hiseq-4000 platform. The second cohort was sequenced using the Illumina Novaseq platform with 150 bp paired-end sequencing. The third batch of samples was run on a NovaSeq6000 platform using 150 bp paired-end sequencing. The bulk RNA-seq was performed at the Oxford Genome Centre, Wellcome Centre for Human Genetics.

2.3.3 Upstream processing of bulk RNA-seq data

HISAT2[319] was used to align reads in FASTQ files to the GRCh38/hg38 genome build. Adequately mapped reads were identified according to MAPQ score with *bamtools*[320]. *Picard*[321] was used to remove duplicate reads. *HTseq*[322] was used to generate read count data.

2.3.3.1 Differential expression analysis

Differential expression analysis of bulk RNA-seq data was carried out using *DESeq2*[323]. *DESEq2*[324] was used to normalise read counts. The *ComBat-seq*[325] package was used to remove the effect of sequencing batch across bulk RNA-seq cohort.

For differential expression analysis of transcripts expressed across healthy controls versus pre-treatment samples, variables including sex and sex were included in the *DESEq2* design argument (i.e. design = ~ sex + age + cycle). For analysis of differentially expressed transcripts between untreated and treated samples (including C1 versus C2, and C1 versus C4 analysis), pairwise analysis was performed. Transcripts were filtered and only those with a mean count > 10 were selected for differential expression analysis. Differential expression analysis was performed using a binomial Wald test. The Benjamin-Hochberg (BH) method was used to correct for multiple testing to identify significantly differentially expressed transcripts ($P_{adj}<0.05$).

To identify transcripts in $CD14^{+}$ cells associated with clinical outcome parameters, differential expression analysis was used, dichotomising clinical outcome data, either into those who had melanoma progression at six months, or those who did not progress; and those who experienced irAEs prior to the fifth treatment cycle, and those that did not. *DESEq2* was used to perform differential expression analysis, controlling for treatment state where applicable, as well as age and sex.

2.3.3.2 Pathway enrichment analysis of differentially expressed transcripts

The *XGR*[326] package was used for pathway enrichment analysis using interrogation of the GOBP database. Unless otherwise stated, induced/suppressed or positively/negatively associated transcripts were analysed separately.

2.3.3.3 Building a diagnostic classifier from monocyte transcriptomic profiles

To create a diagnostic classifier the top 40 most significantly differentially expressed genes between MM patients at baseline (C1) and healthy donors were taken (see Appendix K for gene set) and the *DESeq2*-normalised expression values (following adjustment for batch with *ComBat-seq*) of these transcriptswere used as the input to a training and validation set. The *MASS* package[327] in R was used to perform linear discriminant analysis.

2.4 Flow cytometry

Flow cytometry was performed by Ms Chelsea Taylor from the Fairfax laboratory at the WIMM, University of Oxford. Following separation, PBMCs were stored in liquid nitrogen in 90% FCS and 10% dimethyl sulfoxide (DMSO). Samples were thawed and then $1x10^6$ PBMCs were placed into a solution of HBSS with 5% FCS on ice. After 30 mins cells were fixed in 2% formaldehyde. Samples were stained using a fixable amine-reactive viability dye (LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit). The appropriate antibodies were added to the sample as outlined below (Figure 6). The flow cytometry was carried out with a LSRII (Becton Dickinson) platform. FlowJo software was used to analyse cells blinded to clinical outcome data. For gating, myeloid cells were identified using FSC-A and SSC-A, excluding doublets. Dead/dying cells were removed via the selection of viability stain negative cells. Monocytes were then selected based on CD14 expression, and then classical, intermediate and non-classical monocytes were gated according to relative CD14/CD16 expression.

Antibody/stain	Fluorochrome	Clone	Dilution	Source
Live/Dead cell stain kit	Near-IR	n/a	1000	ThermoFisher Scientific
CD14	PE/Cy7	M5E2	50	BioLegend
CD16	FITC	3G8	50	BioLegend

Figure 6:Table to show antibodies used in flow cytometry to identify monocyte subsets

2.5 Single cell RNA sequencing

2.5.1 Patient cohort and scRNA-seq

scRNA-seq library preparation and sequencing was performed in two batches. In the first, $CD8^+$ and $CD14^+$ cells were isolated from PBMCs from peripheral blood from four

patients treated with ipilimumab/nivolumab combination-immune checkpoint blockade (cICB) and four treated with pembrolizumab single-immune checkpoint blockade (sICB). Samples were obtained before (C1) and after (C2) the first treatment cycle. Healthy donor samples from three individuals were pooled prior to library preparation and sequenced in a second batch. See Appendix M for cohort/sample details.

2.5.2 scRNA-Seq library preparation

The Chromium 10x platform was used for scRNA-seq library preparation. scRNA-seq data was prepared in two batches. In the first batch, 6000 of both $CD8^+$ and $CD14^+$ cells were combined in suspension (total number 12,000 cells). In the second batch, 2000 $CD14^+$ cells from each of three healthy donors were mixed in suspension in a single well. 5' transcriptome processing was performed using the Chromium 10x system. Samples were processed according to standard manufacturer's protocols[328]. This involved initial barcoding with a unique tag, reverse transcription, cDNA amplification and library preparation.

2.5.3 Sequencing

For the first batch, 5' transcriptome library sequencing was performed using a HiSeq platform with 75bp paired-end sequencing with a depth of at least 50,000 reads per cell. The control samples were sequenced using a NovaSeq 6000 platform with 150 bp paired-end sequencing with a mean sequencing depth of 78,000 reads per cell.

2.5.4 Data processing

The *Cellranger* package was used to process scRNA-seq data[329]. The Cellranger *mkfastq* was used to produce FASTQ files from the Illumina BCL files. Cellranger *count* was then applied to each FASTQ file to produce a feature barcoding and gene expression library. Cellranger *mkfastq, count* and *aggr* were used to combine samples for merged downstream analysis.

2.5.5 Downstream scRNA-seq analysis

2.5.5.1 Quality control

For quality control (QC), the *scater* package[330] was used to remove outliers, and *scran* package[331] was used to remove doublets. Cells were then selected based on *CD14* expression, with those cells expressing *CD8A, CD3D, CD3E, CD3G, CD56, CD19* or *CD20* removed from downstream analysis to remove possible contaminants. QC was then performed by removing cells that had fewer than 300 different features (genes), fewer than 500 UMIs (transcripts) and those whereby more than 20% of expressed genes were mitochondrial. Genes expressed in fewer than five cells were also removed from downstream analysis.

2.5.6 Clustering analysis

The *Seurat*[332] package in R was used for downstream analysis of the scRNA-seq data. The raw gene expression matrix was normalised and then subject to *FindVariableFeatures* function[332]. Both untreated and treated samples were merged following normalisation. Integration anchors between MM patient and Healthy donor (Control) data sets were identified using *FindIntegrationAnchors* and integration of the patient and Control datasets was performed with *IntegrateData* function to control for batch effect. Data was then scaled, and then a principal component analysis was run using the first 13 dimensions. This level of dimensionality was chosen as it was the level at which the variation between consecutive principal components was minimal as seen on an *Elbowplot*. Unsupervised clustering was performing using the *FindClusters* function and the clusters visualised using a uniform manifold approximation and projection (UMAP) plot.

To explore heterogeneity within the $CD16^{+}$ compartment unsupervised clustering was performed using *FindClusters*, using 13 principal components. Principal component number was identified by visualising variance accounted for across successive principal components on an *Elbow plot*.

2.5.6.1 Annotation of monocyte clusters

The *clustree*[333] package was used to visualise monocyte clustering trees at resolutions of 0.1, 0.2 0.25, 0.5, 0.75, 1,2 and 3 (at 13 dimensions). To explore the expression of key marker genes across monocyte subset, the Seurat *Featureplot and Dotplot* functions were used to visualise gene expression across the scRNA-seq dataset.

To perform annotation of monocyte clusters with published gene sets the *singleR*[334] package was used to compare the scRNA-seq data to annotated expression datasets including the Human Primary Cell Atlas dataset[335], MonacoImmune dataset[336], Immune Cell Expression data set[337] and scRNA-seq data published by Villani et al[104].

2.5.6.2 Identification of cluster-defining markers

To identify genes differentially expressed in monocyte cluster the *FindAllMarkers* function was used to find differentially expressed genes between a single subset and the remaining data, selecting only positively expressed transcripts (Wilcoxon signed-rank test, $P_{adj}<0.05$).

2.5.6.3 Trajectory analysis

Trajectory analysis was performed across the scRNA-seq dataset using the *monocle* R package[338]. For the analysis across both classical and $CD16^+$ subsets the dataset was randomly sub-sampled to a population of 2000 cells. All $CD16^+$ cells (n=857 cells) were used for pseudotime analysis within the $CD16^+$ subset. Following the creation of a large *CellDataSet* object using the *newCellDataSet* function, size factors and dispersion estimates were calculated using the *estimateSizeFactors* and *estimateDispersions* functions. Genes used for ordering were selected in an unsupervised way, according to those with greatest variance across the dataset, with those with a mean expression of >= 1 and empirical dispersion of >= 0.5 within the *CellDataSet* object. Cells were ordered using the *setOrderingFilter* function. Following dimensionality reduction with *reduceDimension*, cells were ordered and pseudotime maps visualised using the *plot_cell_trajectory* function.

To identify gene that showed branch-dependent expression, the *BEAM* function in *monocle* was used (q-value < 1e-34). To identify genes that changed as a function of pseudotime the *differentialGeneTest* function was used.

2.5.6.4 Differential expression analysis

Differential expression analysis within clusters across conditions was performed using the *Seurat FindMarkers* function with the default Wilcoxon signed-rank sum test. The logFC threshold was set to -Inf and a method based on Bonferoni adjustment was used to correct for multiple testing. To control for different subset sizes between conditions and monocyte clusters, the analysis was limited to the smallest subset available across the relevant condition, except where otherwise stated.

2.5.6.5 Pathway analysis

Pathway analysis of gene sets from the scRNA-seq dataset (e.g. canonical markers genes or differentially expressed genes) was performed using *XGR* interrogation of the GOBP database.

2.5.6.6 Gene module scoring

To calculate the phenotype scores (e.g. immaturity, complement, antigen presenting scores), the Seurat *AddModuleScore* function was utilised to assign a score indicating the average expression of a defined gene set per single cell. The gene sets used for module scoring applied to each cell were identified by selecting the top 50 genes most significantly correlated with hub genes of interest such as *CD36, C1QA,* and *HLA-DR* expression. To identify significantly correlated transcripts, Pearsons correlation analysis was performed across all single cells using the *Hmisc* package in R. Hub genes (as outlined in Figure 30) were selected based on association with a particular phenotype or function. For example, *FCGR3A* was selected as a hub gene indicative of non-classical phenotype.

2.5.6.7 Subset scoring in bulk RNA-seq data

The most significant differentially expressed genes (showing positive gene expression) for each scRNA-seq subset that had been identified using *FindAllMarkers* were

identified. The geometric mean value of these genes *DESeq2*-normalised bulk-RNA expression data per sample was calculated. This expression score per sample in the bulk RNA-seq data was then used to explore the association with clinical outcome, either using a Wilcoxon signed rank test, or using the *survminer*[339] and *survival*[340] packages to perform survival analysis.

2.5.6.8 Survival analysis

The *survminer*[339] and *survival*[340] packages were used to perform survival analysis (exploring association with OS and PFS) and visualise Kaplan-Meier curves. Samples were dichotomised into two subsets according to variable of interest e.g expression score, being above or below the median value. A log-rank test was used to determine statistical significance.

2.6 cfDNA extraction

cfDNA was extracted from plasma (diluted 1:1 in HBSS) using the Circulating Nucleic Acid Kit (Qiagen) according to the standard protocol[341]. cfDNA was eluted in 30 μl of EB buffer and stored at -20°C. dsDNA was quantified using Qubit (ThermoFisher Scientific) and dsDNA size of approximately 180bp was assessed by running on an E-Gel Agarose Electrophoresis system (ThermoFisher Scientific).

2.7 5hmC cfDNA sequencing

2.7.1 5hmC cohort details

Samples were sequenced in two batches including paired untreated and treated samples, unpaired C1 and C2 samples and healthy donors. Two non-5hmC enriched controls were also sequenced, composed of cfDNA from six pooled samples (C1 and C2 respectively) that had also been included in the 5hmC enriched paired cohort (see Appendix CC).

2.7.2 5hmC cfDNA library preparation

2.7.2.1 Addition of a spike-in oligonucleotide sequence for validation of 5hmC pull-down

To confirm 5hmC enrichment a spike-in oligo sequence was added to each cfDNA samples prior to library preparation. This DNA oligonucleotide (oligo) sequence comprised three amplicon sequences, one with high 5hmC-modification, a second with 5mC modification and a third unmodified sequence added in equal quantities. The spike-in oligo was provided by the Song group, at the Target Discovery Institute, University of Oxford. The ratio of these sequences in raw sequencing files were assessed following sequencing as discussed below in section 2.7.4 to confirm successful 5hmC enrichment.

Spike-in	Sequence	Primers
Control ('C')	CGTTTCCGTTCTTCTTCGTCATAACTTAATGTTTTTATTTAAAATACCCTC TGAAAAGAAAGGAAACGACAGGTGCTGAAAGCGAGGCTTTTTGGCCT CTGTCGTTTCCTTTCTCTGTTTTTGTCCGTGGAATGAACAATGGAAGTC AACAAAAAAGCAGCTGGCTGACATTTTCGGTGCGAGTA	C_Fwd CGTTTCCGTTCTTCT TCGTC C_Rev TACTCGCACCGAAAA TGTCA
5hmC	TGAAAACGAAAGGGGATACGGGAAAACGTAAAACCTTCGTAAACACC ACGCTCGGTGAGACGTGGGAGGCGAAAATTGGCGAACGTCCGGATGC TGAAGTGATGGCAGAGCGGAAAGAGCATTATTCAGCGCCCGTTCCTG ACCGTGTGGCTTACCTGACCGCCGGTATCGACTCCCAGCTGGAC	HMC_Fwd TGAAAACGAAAGGGGATAC G HMC_Rev GTCCAGCTGGGAGTCGATAC
5mC	TGAAAACGAAAGGGGATACGGGAAAACGTAAAACCTTCGTAAACACC ACGCTCGGTGAGACGTGGGAGGCGAAAATTGGCGAACGTCCGGATGC TGAAGTGATGGCAGAGCGGAAAGAGCATTATTCAGCGCCCGTTCCTG ACCGTGTGGCTTACCTGACCGCCGGTATCGACTCCCAGCTGGAC	MC_Fwd GTGGCGGGTTATGATGAACT MC_Rev CATAAAATGCGGGGATTCAC

Figure 7: Table to show amplicon and primer sequences for spike-in oligos (from the Song group as previously described[296]).

Spike-in controls were created from lambda DNA using the primers in Figure 7 (as previously published[296]). The PCR reaction described in Figure 8 and conditions described in Figure 9 was then performed. The PCR product was confirmed by running 2 μl of the oligo on a 2% agarose gel (E-gel EX) and a single band seen at 180 bp for the control and 5mC control amplicons, and 185bp for the 5hmC amplicon. The reaction was then purified using the Zymo-Spin IC column according to the manufacturer's instructions. Quantification of the PCR product was then carried out using a Qubit

dsDNA assay Kit. Please note that this protocol was carried out by the Song group and spike-in oligos used in this assay produced by that group.

Reagent	Volume (μl)
Lambda DNA (1 ng/ul)	1
10x Standard Taq polymerase buffer	5
Oligo primer F	1
Oligo primer R	1
dNTP mix	1
Taq polymerase (5000 U/ml)	0.25
Nuclease-free water	40.75
TOTAL	**50**

Figure 8: Table to show PCR reagents for creating a spike-in oligo amplicon

Stage	Temperature (°C)	Time	No. cycles
Initial denaturation	95	30 secs	1
Denaturation	95	30 secs	25
Annealing	51 (C-control oligo), 49 (5hmC and 5mC oligo)	30 secs	
Extensions	68	1 min	
Final extension	68	5 min	1
Hold	4	α	

Figure 9: PCR reaction to create the spike-in oligonucleotide amplicons

2.7.2.2 End-repair and A tailing

5hmC library preparation was carried out according to a previously described protocol[298] using the KAPA Hyper Prep Kit (KAPA Biosystems). A minimum of 5 ng of DNA was taken forward into 5hmc library preparation. A spike-in oligonucleotide sequence as described in section 2.7.2.1 was added to each sample (0.01 pg per 10 ng of DNA). 3.5 μl of end-repair and A-tailing buffer and 1.5 μl of end-repair and A-tailing enzyme from the KAPA Hyper Prep Kit was added to each cfDNA sample and samples were then incubated at 20°C for 30 minutes and then 65°C for 30 minutes in a thermal cycler with the lid temperature set to 85°C.

2.7.2.3 Adaptor ligation

Adapter ligation was performing by adding 1.5 μl (30 uM) of unique index adapters (NEXTFLEX DNA Barcodes, PerkinElmer) (making sure the adapters were kept on ice) per 10 ng of DNA with 3.75 μl of PCR-grade water, 15 μl of ligation buffer and 10 μl of DNA ligase. Following vortexing and briefly spinning in a microcentrifuge, the samples were incubated in a thermal cycler at 20°C for 30 minutes.

A 0.8X bead-based sample clean-up of adapter-ligated cfDNA with KAPA pure beads (KAPA Biosystems) was performed. On adding 44 μl of beads to each sample, the beads and sample were mixed thoroughly by pipetting. Samples were incubated at room temperature for 10 minutes. The samples were then placed on a DynaMag^{TM}-96 Side magnet (ThermoFisher Scientific) until the supernatant was clear. The supernatant was removed and discarded and then the beads were washed with 200 μl of freshly prepared 80% ethanol, ensuring that the beads did not dry out in between washes. Following the second wash the ethanol was removed, the samples spun down and the beads were left for 3-5 minutes to allow excess liquid to evaporate. Beads were then resuspended in 16 μl nuclease-free water, vortexed and then incubated for 2 mins. The samples were then placed on the magnet at room temperature and clear supernatant removed and placed in a fresh PCR tube. At this point samples were stored at -20°C overnight.

2.7.2.4 5hmC labelling

DNA was labelled with UDP-6-N3-Glc (Jena Bioscience). 0.5 μl of UDP-6-N3-Glc was added to each sample, with 2 μl of 10X buffer and 2 μl of β-GT. Samples were then incubated at 37°C for 2 hours in a thermal cycler with lid temperature set to 50°C. Samples were then biotinylated with DBCO-PEG4-biotin by adding 1 μl of DBCO-PEG4-biotin to each sample and then incubating for 37°C for two hours. Immediately following this, 10 μg of salmon sperm was added to each sample (ThermoFisher Scientific) and then samples were purified using Micro Bio-Spin^{TM} P30 Gel Columns (Bio-Rad) according to manufacturer's protocol.

2.7.2.5 Pull-down on streptavidin beads

During the final incubation Dynabeads™ M-270 Streptavidin (ThermoFisher Scientific) were washed. 0.5 μl of beads suspension was used per sample and beads were then washed three times with 50 μl of buffer 1 (see Figure 10) on a DynaMagTM-2 magnet (ThermoFisher Scientifc). The Dynabeads™ M-270 Streptavidin were then resuspended in 50 μl of buffer 1. 10 μl of salmon sperm DNA (10 mg/ml) (ThermoFisher Scientific) was added to the beads and then samples were incubated on a rota at room temperature for 30 minutes. The beads were then washed three times with 50 μl of buffer 1 and then resuspended in the required volume of 2x buffer 1 (50 μl per sample) (Figure 10). 50 μl of washed M-270 bead suspension was then added to each sample following P30 column purification. The samples and beads were incubated on the rota at room temperature for 30 minutes to allow for 5hmC-pull-down.

Following incubation on the rota the beads and sample mixture were placed in a PCR strip tube on a DynaMag™-96 Side Magnet and washed three times with 50 μl of each buffer solution described in Figure 10. The beads were then washed with PCR-grade water three times and then resuspended in 22.5 μl water.

Buffer	Reagents
Buffer 1	5 mM Tris pH =7.5, 0.5 mM EDTA, 1 M NaCl, 0.2% Tween 20
Buffer 2	5 mM Tris pH = 7.5, 0.5 mM EDTA, 0.2% Tween
Buffer 3	5 mM Tris pH = 9.0, 0.5 mM EDTA, 1M NaCl, 0.2% Tween 20
Buffer 4	5mM Tris pH = 9.0, 0.5 mM EDTA, 0.2% Tween 20
2xBuffer 1	10mM Tris pH =7.5, 1 mM EDTA, 2M NaCl, 0.4% Tween 20

Figure 10: Reagents used to create bead-bound DNA cleaning solutions

2.7.2.6 Library amplification

For library amplification a reagent mix was prepared including 25 μl of KAPA Hifi HotStart ReadyMix (2x), 5 μl of KAPA Library Amplification Primer Mix (10x) which was added to the adapter ligated library on ice. The samples were mixed with a pipette before being spun briefly in a microcentrifuge. The PCR reaction was run in a thermal cycler as outlined in Figure 11.

Step	Temperature	Duration	Cycle
Initial denaturation	98	45 second	1
Denaturation	98	15 seconds	14/15
Annealing	60	30 second	
Extension	72	30 seconds	
Final extension	72	1 minute	1
Hold	4	-	1

Figure 11: PCR amplification for 5hmC-enriched libraries

Post-amplification purification was carried out by performing a 1X bead-based clean-up. 25 μl of KAPA Pure Beads was added to each sample. Samples and beads were then mixed by pipetting and then incubated at room temperature for 10 minutes. The samples in PCR tubes were then placed onto a DynaMag^{TM}-96 Side Magnet until the supernatant was clear. The supernatant was then removed and the beads were washed twice with 200 μl of 80% ethanol. The ethanol was then removed from the sample, samples were spun briefly in the microcentrifuge and then left to dry at room temperature for between 3-5 minutes until the liquid had evaporated. The beads were then resuspended in 20 μl of EB buffer (Qiagen). The sample was then placed back on the magnet and once the supernatant was clear it was transferred to a new PCR tube strip.

2.7.2.7 Quantification and QC

The DNA quantity of each sample library was assessed using the Qubit dsDNA HS Assay kit (Life Technologies) according to the standard protocol[342]. 2 μl of each sample was also run on a 1-2% agarose gel using the E-Gel Agarose Electrophoresis system (ThermoFisher Scientific) to confirm correct library length of approximately 300bp. Libraries were pooled to a concentration of 4 nM and concentration confirmed using the KAPA Library Quantification Kit (KAPA Biosystems) as per manufacturer's protocol.

2.7.3 Sequencing

The 5hmC cfDNA libraries were sequenced using paired-end 75bp sequencing on the NextSeq 500 platform in two batches.

2.7.4 5hmC enrichment validation

To confirm 5hmC pull-down the ratio of 5hmC oligo to both the 5mC-modified and control spike-in oligo was determined according to number of reads in FASQ files. Samples with both a 5hmC:5mC and 5hmC:Control oligo enrichment ratio of two or above were taken forward for downstream analysis.

2.8 5-methylcytosine cfDNA sequencing

2.8.1 5mC cohort details

5mC cfDNA library sequencing was carried out in three library batches. Samples were obtained from the cohort described in section 2.1 and are described in Appendix EE.

2.8.2 cfMeDIP library preparation

5mC cfDNA libraries were produced according to a previously described protocol published by Shen et al.[261] using 5mC immunoprecipitation ('cfMeDIP'). In brief, this protocol utilises a 5mC mAb for immunoprecipitation of 5mC-labelled cfDNA. To facilitate low input of cfDNA, this protocol uses a 'filler DNA' comprising variably-methylated and unmethylated lambda DNA amplicons. These were made using the primers as per the protocol described by Shen et al[261] adapted from Taiwo et al.[343].

A minimum input cfDNA of 5 ng was used per library. Library preparation was carried using the KAPA HyperPrep Kit (Roche) out with end-repair and A-tailing and adapter ligation with NEBNext adaptors (New England Biolabs). Following adaptor ligation, cfDNA was quantified using Qubit and filler lambda DNA was added to make a total quantity of 100 ng of DNA per sample. Filler lambda DNA was added in the following proportions: 50% unmethylated 20LCpG amplicon, 50% methylated DNA to include 10% 1CpG, 10% 5CpG, 10% 15CpG and 10% 20SCpG amplicons.

To carry out immunoprecipitation of 5mC-enriched DNA, the MagMeDIP kit (Diagenode) was used as described in the protocol published by Shen et al[261]. In brief, to calculate the percentage methylated and unmethylated cfDNA recovery and 5mC immunoprecipitation specificity a bacterial *A. thaliana* DNA spike-in was added to the adapter-ligated cfDNA and lambda DNA mixture. The DNA was incubated with 5mC

antibody (from the MadMeDIP kit, Diagenode) and Magbeads at 4°C for 17 hours at a concentration of between 0.06 μg/μl - 0.2 μg/μl. Following wash steps with reagents from the IPure kit v2 (Diagenode), to determine the percentage recovery of methylated and unmethylated DNA and to calculate the specificity of pull-down, a qPCR reaction was performed as described by Shen et al.[261]. A proportion of the sample was taken aside prior to immunoprecipitation as non-enriched control samples ('IC' samples), ensuring that the IC sample was 10% of the volume of the immunoprecipitated ('IP') sample. The cycle threshold (Ct) values for both 5mC-immunoprecipitated IP and non-5mC immunoprecipitated IC samples were determined, which were used to calculate the percentage recovery of both methylated and unmethylated bacterial *A. thaliana* spike-in. From this, the specificity of the immunoprecipitation step was calculated. Samples taken forward to sequencing had a minimum 5mC-pulldown specificity of 87% (see Appendix EE for 5mC sample details).

To determine the appropriate number of PCR amplification cycles for each library, a qPCR reaction across both the IP and IC samples was performed. The Ct values obtained were rounded to the nearest whole number and this number indicated the appropriate number of cycles required for library amplification. Library amplification was carried out as per protocol, and then a bead-based library clean-up was performed with Agencourt AMPure XP beads (1X). Dual size selected was then performed by a 0.6X bead-based size selection.

2.8.3 Library quantification and sequencing

The final 5mC-enriched cfDNA libraries were quantified using a dsDNA Qubit as per standard manufacturer's protocols. The 5mC library size was assessed using the Agilent TapeStation using the D1000 ScreenTape Assay according to previously published protocols[344].

5mC cfDNA libraries were pooled to a final library concentration of 4nM and sequenced using a NextSeq500 platform with paired-end 75 bp reads.

2.9 Upstream processing of epigenetic datasets

Following sequencing, FASTQ files were aligned to the GRCh38/hg38 genome buid with *HISAT2*[345]. The *samtools view* function was used to convert SAM to BAM files. *Samtools sort* was applied to sort BAM files. Properly mapped reads were selected and singetons removed by filtering with *MarkDuplicates* and *bamtools filter*. *Samtools*[346] was used to calculate statistics. To create expression matrixes *Stringtie* was used to generate read count data.

BAM files from multiple indivduals were merged using the *samtools merge* function to create merged pre- (C1) and post- (C2) treatment, healthy donor (HD), and non-enriched samples.

2.10 Downstream analysis of epigenetic cfDNA datasets

Unless stated otherwise, the following analysis was performed across both 5hmC and 5mC datasets.

2.10.1 Identification of epigenetically enriched regions

The *MACS2*[347] package was used for identifying enriched regions ('peaks') in both 5hmC and 5mC datasets across the genome ('peak-calling'). Broad peak calling was used with a bandwidth of 300 and q-value cut-off of 5.00e-02. Peaks were called per sample specifying the enriched sample as 'treatment' and the background 'control' as the merged (C1 and C2) non-enriched samples.

2.10.2 Visualisation of mapped reads

Visualisation of read counts per million mapped reads were plotted across the gene body using the *ngs.plot* package[348] with 20kb flanking regions.

2.10.3 Visualisation of epigenetically enriched regions

To visualise epigenetically enriched regions the *ChipSeeker* function *covplot* was used to show the distribution of epigenetic peaks (called by *MACS2*) across chromosomal regions. Non-standard chromosomes were removed with the *BRGenomics*[349] function *tidyChromosomes*.

2.10.4 Annotation of epigenetic cfDNA peaks

To identify the genes within which 5hmC and 5mC enriched regions were situated, MACS2-generated peak sets were annotated with the UCSC hg38 build information using the *ChipSeeker*[350] package. Non-standard chromosomes with the *BRGenomics* function *tidyChromosomes*.

2.10.5 Assessing overlap between epigenetic cfDNA peaksets

Permutation analysis using the *regioneR*[351] package was used to determine whether 5hmC and 5mC peaksets showed significant overlap.

2.10.6 Assessing enrichment across functional genomic regions and histone modifications

To identify epigenetic enrichment across functional regions peak sets were annotated according to locations retrieved from publicly available data from the Roadmap Epigenomics project[352]. Genomic locations of functional regions (e.g. exons, introns, promoters) were retrieved from the UCSC hg38 with the *annotatr* package[353]. The *regioneR*[351] package was used to perform permutation analysis to determine the enrichment of a peakset of interest across a particular function region e.g. exonic regions, compared to a randomly generated peakset. Prior to running permutation analysis non-standard chromosomes were removed using the *BRGenomics*[349] function *tidyChromosomes*.

The same approach was also used to determine enrichment of 5hmC and 5mC peaksets across histone modification sites. Histone modification sites were obtained from data available through the NIH Roadmap Epigenomics project. Histone modification sites (broad peak calling files) for H3K27ac, H3K27me3, H3K36me3, H3K4me1, H3K4me3 and H3K9me3 were obtained from peripheral T cells (cell ID E034) and monocytes (cell ID E029).

To compare enrichment of 5hmC and 5mC across functional genomic regions and histone modification sites, odds ratios were calculated comparing overlap of the 5hmC/5mC peakset of interest to a set of known defined genomic regions (e.g. exonic regions).

2.10.7 Differential enrichment analysis

To perform differential enrichment analysis of 5hmC and 5mC enriched cfDNA the *DiffBind*[354] package was used in R. Samples were loaded into the R environment, and a consensus peakset composed of peaks shared between a specified number of samples (default = 2, used unless otherwise stated) was created. The *dba.count* function was then used to create a matrix of reads within this consensus peakset, and the *DiffBind dba.analyze* wrapper used to perform differential enrichmentr analysis across reads of the consensus peakset with both *edgeR* and *DESeq2*. Multiple testing was corrected for and significant peaks defined as those with FDR <0.05.

2.10.8 Pathway enrichment analysis of epigenetic data

To explore biological function of epigenetically enriched peaks, pathway enrichment analysis was performed using the *ChipSeeker*-annotated genes and interrogation of the GOBP database with *XGR* (FDR<0.05).

2.10.9 Exploring association between epigenetic cfDNA modification and transcriptomic expression in circulating immune cell

To explore association between epigenetic enrichment in cfDNA and transcriptomic expression in circulating immune cells, a score was identified composed of the top 20 most highly enriched 5hmC ('H-5hmC-S') or 5mC regions ('H-5mC-S)' (see Figure 92). Reads were assigned to genic regions using *StringTie* to create an enrichment matrix as described in 2.9. The geometric mean enrichment of these 20 genes was calculated per sample, and *DESEq2* was used to identify transcriptomic profiles in peripheral monocytes and $CD8^+$ T cells in the bulk RNA-seq data that were associated with the H-5hmC-S or H-5mC-S score.

2.11 General statistical analysis

All statistical analyses described in this thesis were performed within the R package (v 4.0.4). Plots were created using *ggplot2.* Boxplots demonstrate the upper and lower quartiles, with the line in the middle indicating the median, and Tukey's whiskers indicating the range.

3 Chapter 3: Characterising transcriptional profiles in peripheral monocytes in response to immune checkpoint blockade

3.1 Background

Myeloid cells have a role in the modulation of tumourigenesis, with tumour associated macrophage (TAM) infiltration a marker of poor prognosis across multiple cancer types, including melanoma[173,355,356]. Macrophage polarisation towards a pro-inflammatory ('M1'), or anti-inflammatory ('M2') phenotype is similarly associated with anti- or pro-tumourigenic effects respectively[357].

Whilst tumour-based myeloid profiles have been relatively well-characterised, the role of the peripheral myeloid compartment in modulating tumorigenesis is not fully understood, yet also appears to have predictive utility. Expansion of the peripheral monocyte compartment has been observed in patients across many cancer types[155], and is reflective of generalised induction towards myelopoiesis[155]. An expanded circulating monocyte population is typically associated with poor prognosis[163–165], and has been shown to have predictive utility in the context of ICB[196,197,358,359]. Despite this, mechanisms by which these effects are modulated have yet to be fully characterised.

In cancer, circulating monocytes undergo reprogramming across a diverse range of biological functions[155]. Characterisation of these profiles may provide insights into biological mechanisms through which monocytes have prognostic effects. Typically, monocytes in patients with cancer undergo polarisation towards an immature phenotype with impaired differentiation and maturation[181] with upregulation of chemotaxis and migration-associated genes such as *CCR2*, and *CX3CR1*[177], and enrichment of pro-angiogenic pathways[360]. Dysregulation of inflammatory cytokine production has also been observed, although directional effect is divergent across different cancer types. Broadly it is thought that the peripheral myeloid compartment in cancer is polarised towards an immunosuppressive phenotype, accompanied by

expansion of the myeloid-derived suppressor cell (MDSC) population[185]. Transcriptional reprogramming also reflects changes in metabolic function[73].

Both circulating monocyte and TAM populations have been shown to be predictive biomarkers in both anti-PD-1 and anti-CTLA-4 therapy[73,197,203,361]. A large circulating monocyte population, or high monocyte:lymphocyte ratio, are negatively prognostic[362,363] and associated with lack of response to ICB across multiple tumour types including melanoma[197,358,364–366]. Tumour-associated TAM profiles and TAM infiltration have also been shown to be indicators of poor prognosis[367,368], and predictive of limited treatment response in patients receiving ICB[361,369]. Evidence is conflicting however, with high classical circulating monocyte population reported to be predictive of clinical response in anti-PD-1 treatment of MM[73]. Similarly, higher baseline levels of non-classical monocytes have been reported to be predictive of response to anti-CTLA-4 therapy in melanoma[203], whilst a smaller post-treatment non-classical monocyte population was associated with prolonged overall survival (OS) in PD-1 treated glioblastoma[365]. Mechanistic explanation for these observations has not yet been fully determined.

Much of the existing work profiling transcriptomic and antigenic expression of circulating monocytes in cancer patients was performed in small cohorts and/or using a limited panel of candidate genes[177,193,360,370], and there has been little characterisation within melanoma cohorts[73]. Myeloid responses to ICB have also not been fully characterised. High-throughput dissection of monocyte transcriptomics in large melanoma cohorts would therefore be of value to provide insights into cancer-related immune dysfunction and responses, which may underpin efforts to identify prognostic and predictive biomarkers. Transcriptional correlates of response in peripheral monocytes may reveal mechanisms by which monocytes have modulatory effects in ICB.

In this chapter bulk RNA-seq of peripheral monocytes was used to characterise transcriptomic expression profiles of monocytes in MM patients versus healthy controls, to explore on-treatment modulation following ICB, and to identify transcriptional correlates of clinical response.

3.2 Aims

The aims of my work were as follows:

- Characterise transcriptomic profiles in peripheral monocytes in a large MM cohort to explore transcriptional reprogramming of peripheral monocytes in MM
- Characterise on-treatment modulation of transcriptomic profiles in peripheral monocytes following ICB
- Compare responses to cICB versus sICB in monocytes
- Explore clinical correlates of response with monocyte expression profiles in MM patients

3.3 Results

To characterise expression profiles of monocytes in melanoma patients and transcriptomic modulation with ICB, bulk RNA-seq analysis of $CD14^+$ cells isolated from PBMCs was performed. Samples were obtained before treatment (C1, 114 patients), prior to the second treatment dose (C2, 98 patients), and prior to the fourth dose (C4, 22 patients). $CD14^+$ cells were also obtained from healthy donor samples (HD, 45 samples). A total of 279 transcriptomes, with a total of 14,071,879,692 reads and 47,284,508 median reads per sample, were utilised for analysis (see Appendix A, Appendix B and Appendix C).

3.3.1 Distinct transcriptional profiles are seen in peripheral monocytes in melanoma

To explore whether baseline transcriptomic profiles in peripheral monocytes from patients with MM were distinct from those of healthy individuals, differential expression analysis with *DESeq2*[324] was performed across $CD14^+$ cells from patients with MM versus healthy donors (HD) (C1 n=114 samples; HD n=45 samples). Differential expression analysis of monocyte transcriptomic profiles identified 1774 significantly differentially regulated transcripts (adjusted p-value (P_{adj}) <0.05) (Figure 12). Of the significantly modulated transcripts, 871 were induced in patients versus controls, and 903 were suppressed (Figure 12B). Although the proportion of transcripts induced and suppressed

were similar, there was greater magnitude of induction effect in MM patients (Figure 12A).

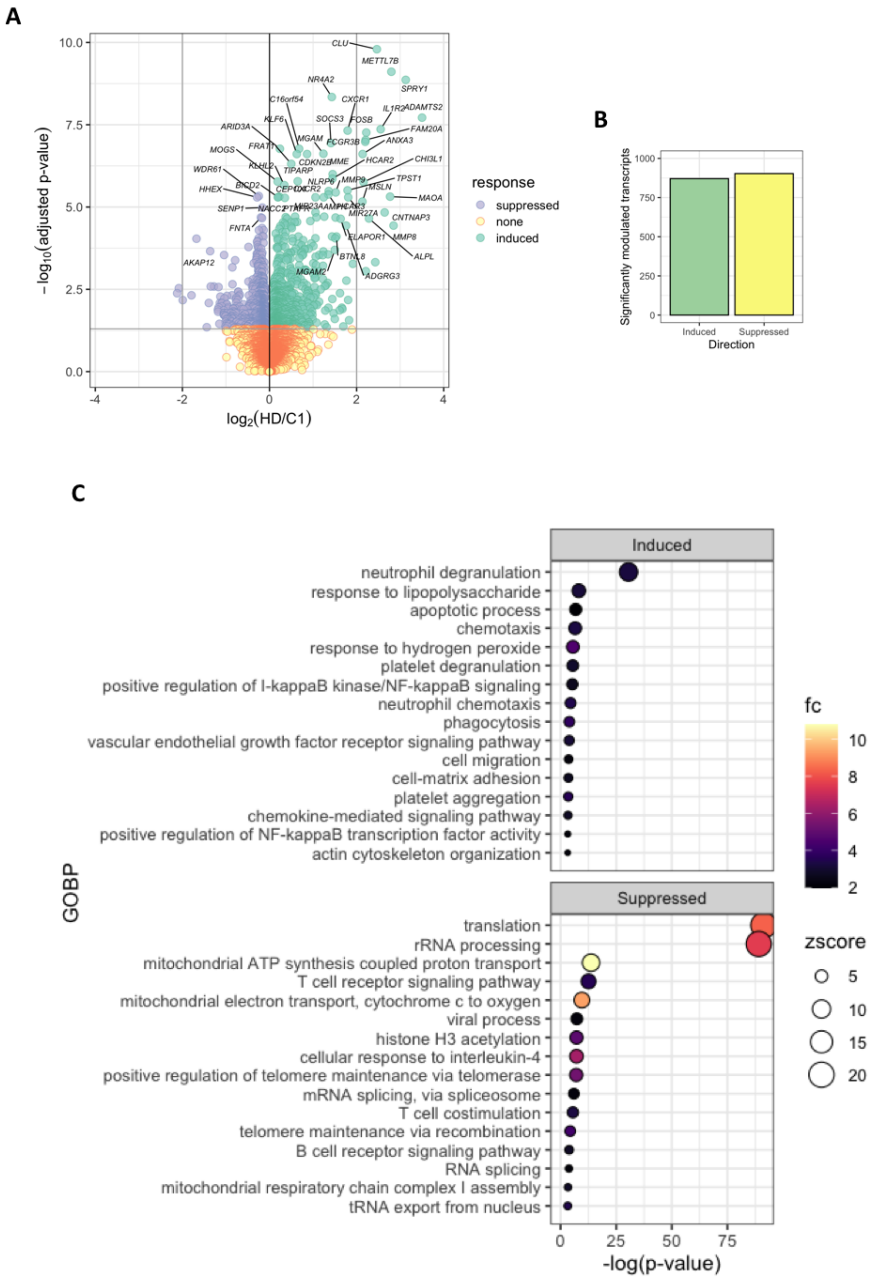


Figure 12: Transcriptomic profiles in peripheral monocytes in melanoma patients
Volcano plot to show significantly differentially expressed genes between patients at baseline (C1) and healthy donors (HD) (114 C1 samples, 45 healthy controls, DESEq2 analysis controlling for age and sex and for batch effect with ComBat-seq). B) Bar plot to show number of significantly transcripts induced in monocytes in patients ('induced') and downregulated ('suppressed') (P_{adj}<0.05). C) Dot plot to show key significantly enriched and suppressed GOBP pathways in monocytes in patients versus healthy controls (P_{adj} <0.05).

Modulation of transcripts associated with a diverse range of biological functions were seen in MM patients (Figure 13). The most significantly differentially modulated transcripts included *FCGR3B* (P_{adj}= $1.0x10^{-7}$), upregulated in patients versus controls. *FCGR3B* encodes the CD16 antigen, a type III Fcγ receptor for IgG, which defines the non-classical and intermediate monocyte subset[109]. Other highly significantly induced genes in patients included the chemokine receptor *CCR1* (P_{adj}= 0.044), pro-inflammatory cytokine *IL1B* (P_{adj} = 0.00084), as well as other cytokines and cytokine receptors including *IL18R1* (P_{adj}= 0.011) and *IL17RA* (P_{adj}= 0.00016). Chemokine receptors *CXCR1* and *CXCR2* which bind to the pro-inflammatory cytokine and neutrophil chemoattractant IL-8 and which have recently been found to be associated with poor response to ICB[371], were found to be highly significantly induced in MM patients (P_{adj}= $4.6x10^{-8}$; P_{adj}= $4.9x10^{-6}$) . There was increased expression of the immunosuppressive cytokine *IL10* (P_{adj}= 0.01)[372] in MM patients versus HDs. Transcripts with reduced expression in MM patients included the cytokine *TNF* (P_{adj}= 0.018), the complement component *C1QBP* (P_{adj}= 0.0274) and MHC class II antigen presenting genes including *HLA-DMA* (P_{adj}= 0.011), *HLA-DPA1* (P_{adj}= 0.019) and *HLA-DPB1* (P_{adj}= 0.0031).

To explore differential functional profiles in peripheral monocytes in patients versus healthy controls, pathway enrichment analysis was performed using *XGR*[326] interrogation of the GOBP database[373] (Figure 12C). There were 68 pathways enriched in MM patients versus healthy controls (P_{adj}<0.05) (Figure 12C). These include pathways such as chemotaxis (GO:0006935, P_{adj}= 0.0014; GO:0060326, P_{adj}= 0.0096) and cell migration (GO:0016477, P_{adj}= 0.026). Induction of chemotaxis pathways in monocytes facilitates migration to sites of inflammation and promotes myeloid infiltration into tumours which may differentiate into TAM with pro-tumourigenic effects. Indeed, systemic and tumour expression of chemokine *CCL2* has been associated with poor prognosis across different cancer types[374].

Other enriched pathways induced in MM patients include those associated with a pro-thrombotic state and angiogenesis, including platelet degranulation (GO:0002576, P_{adj} = 0.0041) and aggregation pathways (GO:0070527, P_{adj}= 0.031), and vascular endothelial growth factor (VEGF) signalling (GO:0048010, P_{adj}= 0.02). Innate immune response pathways were upregulated, including response to lipopolysaccharide (LPS)

(GO:0032496, P_{adj} = 0.00027) and phagocytosis (GO:0006909, P_{adj} = 0.019). Nuclear factor kappa light chain enhancer of activated B cells (NFκB) signalling (GO:0051092, P_{adj}= 0.039) and mitogenic signalling pathways (GO:0000187, P_{adj}= 4.8e-02) were also enriched in MM. In myeloid cells the NFκB transcription factor induces expression of numerous pro-inflammatory cytokines including TNF-α, IL-1β and IL-6[375].

A total of 69 GOBP pathways were downregulated in patients versus controls, including pathways involved in transcription and translation such as RNA splicing (GO:0008380, P_{adj}= 0.023), rRNA processing (GO:0006364, P_{adj} = $2.1x10^{-39}$) and tRNA nuclear export (GO:0006409, P_{adj}= 0.039) (Figure 12C and Appendix D). Pathways associated with mitochondrial metabolism were downregulated, including mitochondrial ATP synthesis (GO:0042776, P_{adj} = $3.6x10^{-8}$), mitochondrial respiratory chain assembly (GO:0032981, P_{adj}= 0.035) and mitochondrial electron transport (GO:0006123, P_{adj} = $6.8x10^{-05}$; GO:0006120, P_{adj}= 0.012) (Figure 12 and Appendix E). T cell receptor signalling (GO:0050852, P_{adj} = $3.4x10^{6}$) and T cell co-stimulation (GO:0031295, P_{adj} = 0.0039) pathways were also downregulated. Pathways including telomere maintenance (GO:0000722, P_{adj}= 0.013) and telomerase activity (GO:0032212, P_{adj}= 0.00087) were downregulated, suggesting dysregulation of normal cellular ageing mechanisms. B cell signalling pathways (GO:0050853, P_{adj}= 0.021) were also downregulated, as well as the response to virus pathway (GO:0009615, P_{adj}= 0.028).

It is notable that *CXCR1* and *CXCR2* expression was prominently upregulated. CXCR1 and CXCR2 are receptors for the cytokine IL-8, which mediates neutrophil chemotaxis and degranulation and which has pro-tumourigenic effects via promotion of angiogenesis, invasion and tumour myeloid infiltration[371,376]. Serum IL-8 is associated with poor response in the context of ICB[371]. Upregulation of *CXCR2* in myeloid cells has been reported in patients with pancreatic cancer (PDAC), with CXCR2 inhibition shown to successfully enhance the effect of anti-PD-1 treatment in murine PDAC models[377]. In addition, it has been demonstrated that CXCR1 and CXCR2 expression on circulating $CD8^+$ T cells is associated with the absence of autoimmune side effects in patients treated with ICB[40]. Evidence is however conflicting, with recent profiling of a $CD14^+CD16^+HLA\text{-}DR^{hi}$ classical monocyte population reporting elevated expression of the CXCR2 ligand *CXCL2* in patients with MM who responded to anti-PD-1 treatment[73].

Further work to assess serum IL-8 levels in this cohort would be informative to characterise the mechanism and role of the CXCR1/CXCR2–IL8 axis in this context.

Induction of pro-inflammatory cytokine IL-1β expression has previously been reported in RCC patients, with IL-1β expression in circulating monocytes shown to play a key role in tumorigenesis. Chittezhath et al.[360] demonstrate that IL-1R signalling induces TAMs towards a pro-tumourigenic phenotype, with blockade of the IL-1/IL-1R axis halting tumour growth. They also found that IL-1β expression in the tumour was associated with advanced clinical stage. Upregulation of IL-1β expression in monocytes in MM patients in this dataset is therefore in keeping with a pro-tumourigenic phenotype.

These observations demonstrate distinct reprogramming within the peripheral monocyte compartment in patients with MM, towards a population enriched for pro-tumourigenic pathways including upregulation of chemotaxis, cell-matrix adhesion and angiogenesis-related pathways. Pathways associated with a pro-thrombotic state such as VEGF signalling and platelet aggregation were also enriched, suggesting that the circulating myeloid population may play a role in facilitating the hypercoagulable state characteristic of malignancy. Disruption of immune response pathways were evident, with antigen-presentation via MHC class II and T cell co-stimulation pathways downregulated. Modulation of pathways associated with cell maturation and replicative potential were noted, as well as cellular processes such as metabolism and replication. These observations are broadly in keeping with previously described cancer-associated monocyte profiles. However, it is notable that whilst previous studies have demonstrated reduced response to LPS and phagocytic activity in cancer[378], these data I present in this chapter demonstrate upregulation of phagocytosis and induction of response to LPS, in keeping with activation of innate immune response pathways.

It is notable that many pathways induced in MM in this cohort are those typically enriched within classical monocytes, whilst transcripts and pathways in keeping with intermediate and/or non-classical monocyte profiles, such as antigen presentation, anti-viral responses and mitochondrial metabolism, were downregulated. In addition, pathways associated with RNA translation and telomere maintenance were downregulated in MM. This is suggestive of a non-dividing and mature cell population,

in keeping with a non-classical monocyte phenotype. Thus, these transcriptional profiles may reflect polarisation towards a more immature, classical-type phenotype in MM.

Upregulated			Downregulated		
Function	**Gene**	**P_{adj}**	**Function**	**Gene**	**P_{adj}**
Mitogenic signalling	*DUSP6*	0.0042	MHC Class II antigen presentation	*HLA-DPA1*	0.019
	HGF	0.0040		*HLA-DPB1*	0.0031
	IL1B	0.00084		*HLA-DMA*	0.011
	ITGA1	0.0017		*HLA-DMB*	0.0087
	MAPK14	0.011			
	MAP2K6	0.024			
Chemokine signalling	*CCR1*	0.045	Complement	*C1QBP*	0.027
	CCRL2	0.038			
	CXCL1	$7.6x10^{-5}$			
	CXCL8	0.0031			
	CXCR1	$4.6x10^{-8}$			
	CXCR2	$4.9x10^{-6}$			
Pro-thrombotic	*GP6*	0.0011	Protein expression (Transcription, translation, ribosomal processing)	*RPL23*	0.0044
	GP9	0.025		*RPL27*	0.0025
	PPBP	0.017		*RPL36*	0.0015
	GTPBP2	0.048		*RPS21*	0.00023
	PROS1	0.0077		*DARS*	0.024
				EIF2A	0.044
				METAP1	0.0027
Angiogenesis	*CRK*	0.0040	TCR signalling and costimulation	*HLA-DPA1*	0.019
	NCF1	0.022		*HLA-DPB1*	0.0031
	NCF4	0.0087		*MALAT1*	0.018
	PTK2B	0.023		*PRKCQ*	0.0065
	CRK	0.0037		*TNF*	0.018
	CYBA	0.031		*CD3D*	0.030
				CD3E	0.024
				CD5	0.0036
Response to LPS	*FOS*	0.011	Telomere maintenance	*CCT2*	0.036
	PTGFR	0.0019		*CCT4*	0.0023
	ADM	0.043		*CCT6A*	0.011
	CXCL1	$7.6x10^{-5}$		*FBXO4*	0.00051
	C5AR1			*RPS7*	0.013
				RPL5	0.015
GTPase signalling	*ARHGAP29*	0.0040	B cell receptor signalling	*CD19*	0.027
	ARHGAP6	0.0065		*CD79A*	0.025
	ARHGEF11	0.0018			
	TSC2	0.028			
	GMIP	0.0030			
Phagocytosis	*ICAM5*	0.0025	DNA replication	*POLA1*	0.0097
	GAS6	0.011		*POLE3*	0.00018
	ANXA1	0.015		*RRM1*	0.015
				SET	0.0092
NFκB signalling	*CLU*	$1.5x10^{-10}$	Base excision repair	*COPS2*	0.033
	IL1B	0.00084		*COPS4*	0.037
	IKBKG	0.0077		*COPS7B*	0.018
	TLR2	0.0013		*RBX1*	0.012
				RPS27A	0.00089
				POLD2	0.00356
				RFC5	0.0082

Figure 13: Key selected transcripts differentially regulated in MM patients
Table to show key selected transcripts differentially expressed in MM patients at baseline (C1) versus healthy donors (HD).

3.3.2 Peripheral monocyte count is associated with transcriptomic monocyte profiles

In keeping with the existing literature, in this MM cohort an elevated monocyte count at baseline has been found to be predictive of adverse clinical outcome[40,71]. To explore the relationship between clinical peripheral monocyte count at baseline and monocyte gene expression, differential expression analysis was performed with *DESeq2* (controlling for age and sex). There was prominent association between peripheral monocyte count at baseline (C1) and transcriptomic expression, with a total of 1,344 transcripts from pre-treatment samples associated with baseline peripheral monocyte count including 694 positively correlated, and 650 anti-correlated with peripheral monocyte count (88 C1 samples, $P_{adj}<0.05$) (Figure 14A).

Transcripts significantly positively associated with peripheral monocyte count included the matrix metalloproteinase genes *MMP8* (P_{adj}= 0.00013), *MMP9* (P_{adj}= 0.0063), *MMP19* (P_{adj}= 0.0051) and *MMP25* (P_{adj}= 0.0022). Proliferation index marker *MKI67* was also highly significantly positively associated with increasing monocyte count, in keeping with greater proliferative activity reflecting a larger peripheral monocyte population (P_{adj}= $7.3x10^{-06}$). Platelet-associated genes such as *PPBP* (P_{adj}= 0.0051) and *GP9* (P_{adj}= 0.028) were also positively associated with monocyte count, in keeping with a pro-thrombotic state. Expression of immunosuppressive cytokine *IL10RB* (P_{adj}= 0.0048), *IL1R2* decoy receptor which inhibits IL-1 binding and function (P_{adj}= $2.6x10^{-7}$), and *IL4* (P_{adj}= 0.0079), were also positively correlated with monocyte count.

Transcripts that showed negative association with peripheral monocyte count include *C3* (P_{adj}= 0.00055), a key gene in initiation of both classical and alternative complement cascades; as well as pro-inflammatory cytokine *IL12A* (P_{adj}= 0.045), which stimulates IFN-γ production by T cells[379], and *CXCL10* (P_{adj}= 0.00057), which is induced by IFN-γ and which has prominent chemoattractant properties[380]. Consistent with this, monocytye expression of MHC class II genes, including *HLA-DRA* (P_{adj}= 0.00050), *HLA-DMA* (P_{adj}= 0.0063), *HLA-DPA1* (P_{adj}= 0.00057) and *HLA-DPB1* (P_{adj}= 0.0010), which are typically induced by IFN-γ, were negatively associated with peripheral monocyte count at baseline.

Pathway enrichment analysis was performed across transcripts positively and anti-correlated with peripheral monocyte count at baseline using *XGR* (Appendix F and Appendix G). Many of the pathways positively associated with monocyte count included those typically associated with a classical monocyte phenotype, including cell migration (GO:0030335, P_{adj}= 0.004), cell adhesion (GO:0007155, P_{adj}= $6.7x10^{-4}$), chemotaxis (GO:0006935, P_{adj}= 0.04), VEGF signalling (GO:0048010, P_{adj}= 0.0029), response to wounding (GO:0009611, P_{adj}=0.033) and phagocytosis (GO:0006909, P_{adj}=0.027) (Figure 14B). The previously described monocyte cancer-associated pathways in section 3.3.1 were positively associated with monocyte count, including platelet activation (GO:0070527, P_{adj} = 0.0053), neutrophil degranulation (GO:0043312, P_{adj} = $4.6x10^{-20}$) and blood coagulation (GO:0007596, P_{adj}= 0.047). The response to LPS pathway (GO:0032496, P_{adj} = 0.048) was also positively associated.

Pathways negatively associated with monocyte count included DNA damage repair (GO:0006281, P_{adj}= 0.024) (Figure 14B), antigen presentation via MHC class II (GO:0019886, P_{adj}= 0.014), T cell co-stimulation (GO:0031295, P_{adj}= 0.041), viral processes (GO:0016032, P_{adj}= 0.0053) and intracellular viral transport (GO:0075733, P_{adj}= 0.0018) pathways. This suggests that the co-stimulatory function of the circulating monocyte population may be reduced in patients with an expanded circulating monocyte population. This may reflect polarisation towards an immature classical phenotype.

We have previously found that elevated baseline monocyte count is a marker of unfavourable clinical outcome in patients receiving ICB in this cohort[40], and the prognostic utility of peripheral monocyte count has been reported widely across cancer types. The observations in this chapter indicate previously undescribed mechanisms through which the prognostic effects of peripheral monocytes and elevated counts are mediated. Predictive utility of a large circulating monocytes population may be conferred via myeloid-effected pro-tumourigenic effects. This indicates that this effect is not simply quantitative, but shows that a high monocyte count is qualitatively associated with expression profiles in monocytes themselves. For example, enrichment of pathways across transcripts positively associated with monocyte population size

revealed pro-thrombotic pathways, suggesting that the circulating myeloid population may play a role in inducing a pro-thrombotic and hypercoagulable state in cancer.

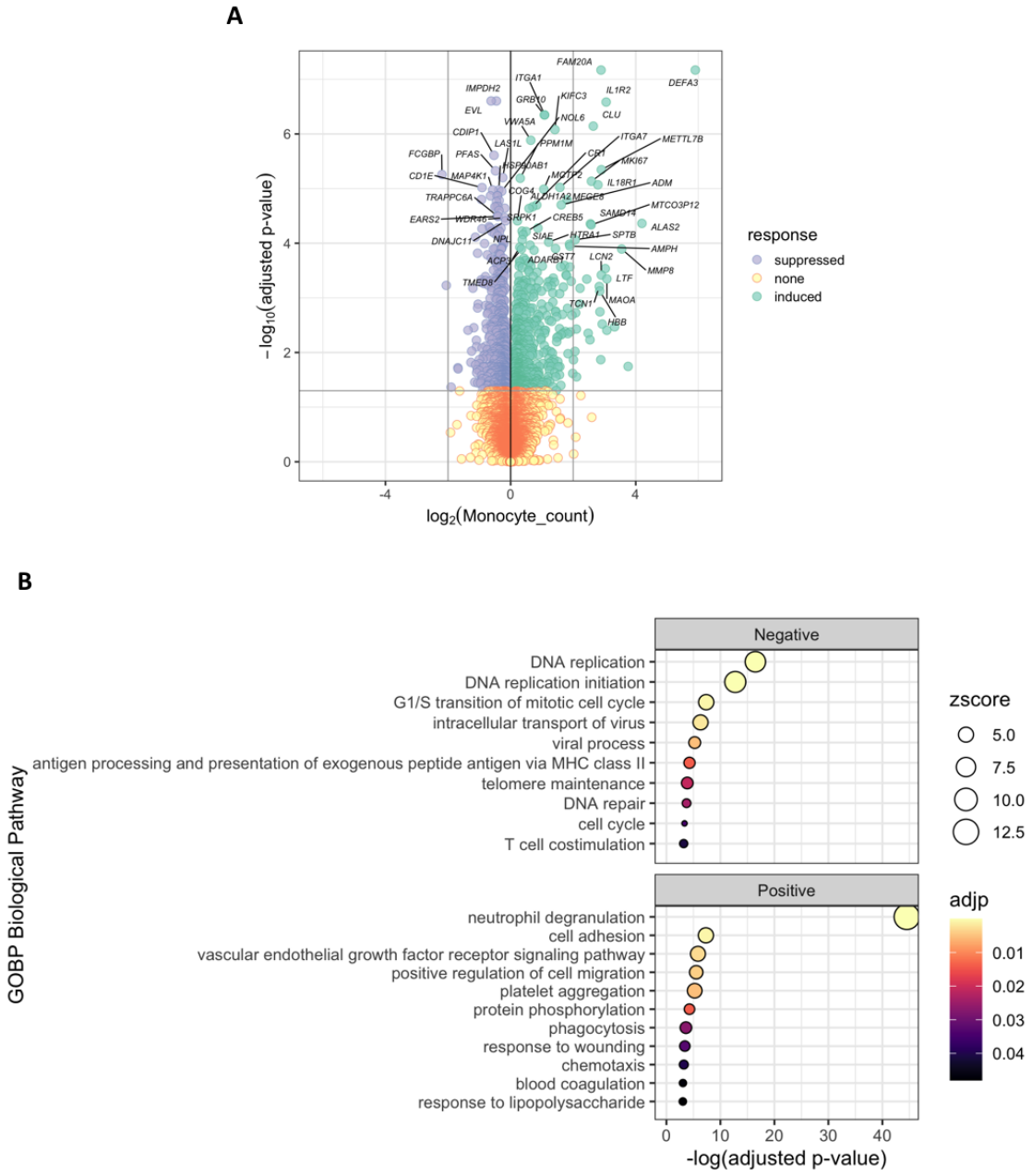


Figure 14: Monocyte expression correlates of baseline monocyte count
A) Volcano plot to show monocyte transcripts associated with baseline peripheral monocyte count (DESEq2 analysis, controlling for age and sex and using ComBat-seq to control for batch effect, 88 C1 samples, P_{adj}<0.05). B) Pathway enrichment analysis of key GOBP pathways showing key positively and negatively associated pathways with baseline monocytes count (P_{adj}<0.05).

3.3.3 Exploring the utility of monocyte expression profiles as a diagnostic classifier

It has previously been demonstrated that transcriptional expression profiles in monocytes may have diagnostic utility in patients with cancer. Hamm et al.[193] identified a gene expression marker panel in peripheral monocytes that successfully identified patients with colorectal cancer from healthy donor samples with a high level of diagnostic accuracy. To explore whether a similar monocyte gene expression panel could be used to identify MM patients versus healthy controls, a linear discriminant analysis (LDA) was performed to create a predictive model. The top 40 most significantly differentially expressed genes in MM patients at baseline versus healthy controls were taken and normalised expression values used as the classifier input across patients at baseline (n=115) and healthy donors (n=45) (see Appendix K for gene set). A random forest (RF) classifier was used to determine the importance of the gene markers in the model. The *MASS* package[327] was then used in R to perform a linear discriminant analysis to create a predictive model with an area under the curve (AUC) of 0.75, demonstrating a good degree of diagnostic accuracy (Figure 15). This demonstrates the potential that peripheral immune-based expression profiling may have as a clinical diagnostic biomarker. This tool could potentially be used in patients with stage three disease to detect those at higher risk of relapse, and to guide investigation and treatment by prioritising patients for imaging and adjuvant ICB treatment. Additionally, where applicable the diagnostic utility of the antigenic expression of these markers could be explored which might underpin the development a small diagnostic flow panel.

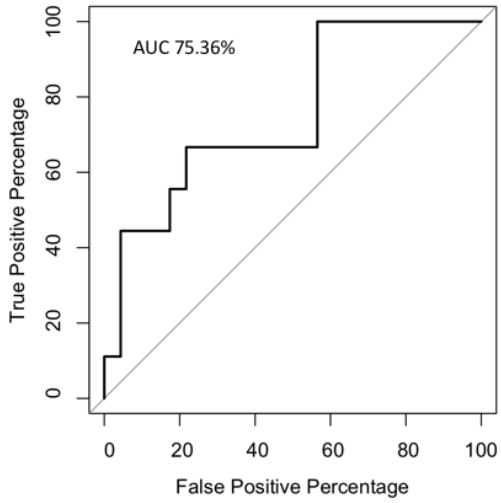


Figure 15: Area under the curve for linear discriminant analysis classifier
Plot to show false positive and false negatives as per the linear discriminant analysis predictive model to distinguish between baseline C1 and healthy donors.

3.3.4 Modulation of transcriptomic profiles in peripheral monocytes with ICB

Next, I explored whether monocyte transcriptomic profiles were modulated by ICB. To do this I performed pairwise differential expression analysis across 96 patients with MM between baseline (C1) and following treatment (C2). This included 45 patients who had received single agent anti-PD-1 treatment (sICB) and 51 patients treated with combination anti-PD-1/anti-CTLA-4 (cICB) treatment. Across all 96 paired samples, pairwise analysis revealed a total of 2,983 significantly differentially modulated transcripts with ICB including 1,587 that were induced, and 1396 that were downregulated (P_{adj}< 0.05) (Figure 17).

Prominent amongst transcripts most significantly induced with ICB included those involved in both type I and II interferon signalling response, including guanylate-binding protein genes (including *GBP1* (P_{adj}= 7.7x10^{-18}), *GBP2 (*P_{adj}= 7.0x10^{-17}) and *GBP3* (P_{adj}= 2.6x10^{-10}) (Figure 16). JAK-STAT pathway genes were also induced, including *STAT1-3* (P_{adj}= 5.6x10^{-15}, P_{adj}= 8.7x10^{-08}, P_{adj}= 0.011 respectively), and *JAK2* (P_{adj}= 6.2x10^{-9})). Interferon-inducible transcripts were also upregulated with ICB, including *IFITM3* (P_{adj}= 9.2x10^{-6}), *IFITM1* (P_{adj}= 0.002) and *IFI27* (P_{adj}= 9.7x10^{-8})). Immune axis transcripts *CXCL9* (P_{adj}= 6.974172x10^{-17}), *CXCL10* (P_{adj}= 9.6x10^{-14}) and *CXCL11* (P_{adj}= 3.6x10^{-9}) were induced. *CXCL9, CXCL10* and *CXCL11* are induced by IFN-γ and are involved in CXCR3-dependent

immune cell chemotaxis and T_H1 polarisation[380,381]. Induction of these cytokines has previously been associated with the development of autoimmune side effects following ICB, with levels at baseline anti-correlated and post-ICB levels positively correlated with autoimmune side effects[382].

Classical complement genes *C1QA* (P_{adj}= 1.7x10^{-10}), *C1QB* (P_{adj}= 5.5x10^{-10}) and *C1QC* (P_{adj}= 7.2x10^{-10}); numerous MHC class II genes such as *HLA-DRA* (P_{adj}= 2.6x10^{-8}), *HLA-DRB5* (P_{adj}= 4.0x10^{-7}) and *HLA-DPB1* (P_{adj}= 3.2x10^{-12}); and the myeloid pro-inflammatory marker *CD38* (P_{adj}= 1.2x10^{-16}) were also induced[383] (Figure 16). Notably, induction of *CD274* (P_{adj}= 2.0x10^{-11}) which encodes the immune checkpoint PD-L1 receptor[384,385], was observed.

Transcripts downregulated with ICB included the pro-inflammatory cytokine receptor *IL1R1* (P_{adj}= 4.2x10^{-9}) and thrombospondin gene *THBS1* (P_{adj}= 4.0x10^{-6}), the latter of which has a role in modulating angiogenesis and endothelial cell adhesion[386] (Figure 16). *PPARGC1A* (P_{adj}= 6.2x10^{-13}), which interacts with PPARγ facilitating transcription factor binding and which has a key role in regulating mitochondrial synthesis[387], was amongst the most significantly downregulated transcripts with ICB.

Pathway enrichment analysis of GOBP pathways showed significant induction of anti-viral pathways and those associated with T cell activation (Figure 17 and Appendix H). These include induction of IFN-γ-signalling (GO:0060333, P_{adj}= 2.3x10^{-27}) and type I interferon-mediated signalling pathways (GO:0060337, P_{adj}= 8.3x10^{-20}), antigen presentation via MHC class II (GO:0019886, P_{adj}= 0.00077), and T cell receptor signalling (GO:0050852, P_{adj}= 7.2x10^{-15}) and co-stimulation pathways (GO:0031295, P_{adj}= 0.0022) (Figure 17 and Appendix H).

IFN-γ is a cytokine and type II interferon produced predominantly by $CD8^+$ T cells, T_H1 cells and NK cells, which has prominent anti-viral effects and which exerts multiple diverse biological effects on innate immune responses[388]. It acts via its receptors IFNGR1 and IFNGR2, which are predominantly expressed on myeloid cells, and which activate the JAK-STAT1 signalling cascade[389,390]. Its effects on myeloid cells and antigen presenting cells (APCs) include induction of MHC Class I and II[391] and co-stimulatory receptors such as CD80[392] expression, with T cell co-stimulatory effects. IFN-γ also

induces nitric oxide (NO) and reactive oxygen species production[393], phagocytic activity[394] and polarisation of macrophages towards a M1 anti-tumourigenic phenotype[395]. It also inhibits the development of a MDSC phenotype in myeloid cells[396]. However, it also has an immune regulatory role via induction of PD-L1 expression in macrophages, monocytes and on tumour cells[384,397]. Type I IFN effects include, like IFN-γ, both immune stimulatory and inhibitory effects, stimulating CCL2 production by monocytes enhancing monocyte recruitment[398], and enhancing co-stimulatory marker CD40, CD80 and CD86 and MHC class II expression on APCs with T cell co-stimulatory effects[399,400] whilst in the context of chronic inflammation induce APC-expression of immunosuppressive molecules such as IL-10 and PD-L1[388,401,402]. Induction of interferon pathways in peripheral monocytes following ICB may be reflective of interferon production by activated CD8 T^+ cells following ICB treatment. Monocyte IFN-γ signalling is inversely correlated with *CSF1R* expression, which has a key role in monocyte recruitment to tissues and macrophage differentiation and TAM infiltration[403] It may be that ICB exerts effects on myeloid populations via induction of IFN-γ signalling pathways with resultant anti-TAM effects, however these effects have yet to be fully explored.

Pathways associated with cellular activation and turnover, including MAPK signalling (GO:0000165, P_{adj}= 0.00058), ATPase activity (GO:0046034, P_{adj}= 0.034) and cell proliferation (GO:0008283, P_{adj}= 0.0041) were also upregulated with ICB (Figure 17B). Notably, the negative regulation of the IL-10 signalling pathway was upregulated (GO:0032693, P_{adj}= 0.0014), in keeping with ICB-associated downregulation of an immunosuppressive phenotype characteristic of myeloid profiles in cancer. Pro-inflammatory signalling pathways including TNF (GO:0033209, P_{adj}= $3.5x10^{-11}$) and NFκB signalling (GO:0038061, P_{adj}= $3.5x10^{-11}$) were also induced. GOBP pathways downregulated with ICB included protein phosphorylation (GO:0006468, P_{adj}= $6.3x10^{-6}$), GTPase activity (GO:0051056, P_{adj}= $1.3x10^{-5}$) and angiogenesis (GO:0001525, P_{adj}= 0.008) (Figure 17B and Appendix I).

Induced			Suppressed		
Pathway	**Gene**	**P_{adj}**	**Pathway**	**Gene**	**P_{adj}**
Antigen presenting genes (MHC Class II)	*HLA-DPB1*	$3.2x10^{-12}$	Angiogenesis	*NOV*	0.00021
	HLA-DPA1	$5.7x10^{-11}$		*PRKC1*	0.0067
	HLA-DMA	$7.9x10^{-11}$		*VEZ1*	0.047
	HLA-DMB	$2.3x10^{-10}$		*ITGAV*	0.046
	HLA-DQB1	$8.2x10^{-10}$			
	HLA-DQA1	$1.6x10^{-9}$			
	HLA-DOA	$1.9x10^{-8}$			
	HLA-DRA	$2.6x10^{-8}$			
Classical complement cascade	*C1QA*	$1.7x10^{-10}$	GTPase activity	*ITSN1*	$1.8x10^{-9}$
	C1QB	$5.5x10^{-10}$		*ARHGAP19*	0.0073
	C1QC	$7.2x10^{-10}$		*ARHGAP24*	0.015
	C1QL3	0.015		*ARHGAP29*	0.00060
	C2	$4.3x10^{-15}$		*ARHGAP6*	0.00028
				ECT2	0.00016
				NCAM1	0.036
				NF1	0.012
				IRS2	0.00068
Anti-viral/ response to IFN	*APOBEC3A*	0.0018	Phosphorylation	*AKT3*	0.00060
	APOBEC3C	$2.6x10^{-8}$		*PIK3R1*	0.0034
	APOBEC3D	$6.6x10^{-9}$		*MAP3K6*	0.020
	APOBEC3F	$1.2x10^{-9}$		*MAPK5*	0.039
	IFITM3	$2.7x10^{-7}$		*ADAM10*	0.031
	IFIT3	0.00018		*APP*	0.024
	STAT1	$5.6x10^{-15}$		*TGFBR1*	0.0027
	STAT2	$8.7x10^{-08}$		*BRAF*	0.00063
	GBP1	$7.7x10^{-18}$			
	GBP2	$7.0x10^{-17}$			
T cell activation	*PSMA2*	0.0020	Cell-cell adhesion	*BCL2*	0.048
	PSMA3	$3.3x10^{-14}$		*CD2*	0.026
	PSMA4	$2.3x10^{-13}$		*CD93*	$1.6x10^{-7}$
	PSMA5	$2.0x10^{-12}$		*ICAM3*	0.0091
	PSMA7	0.046		*PKD1*	0.0019
	HLA genes				

Figure 16: Table to show key modulated transcripts with ICB

Key transcripts modulated with ICB using DESeq2 analysis across 96 paired patient samples (P_{adj}<0.05).

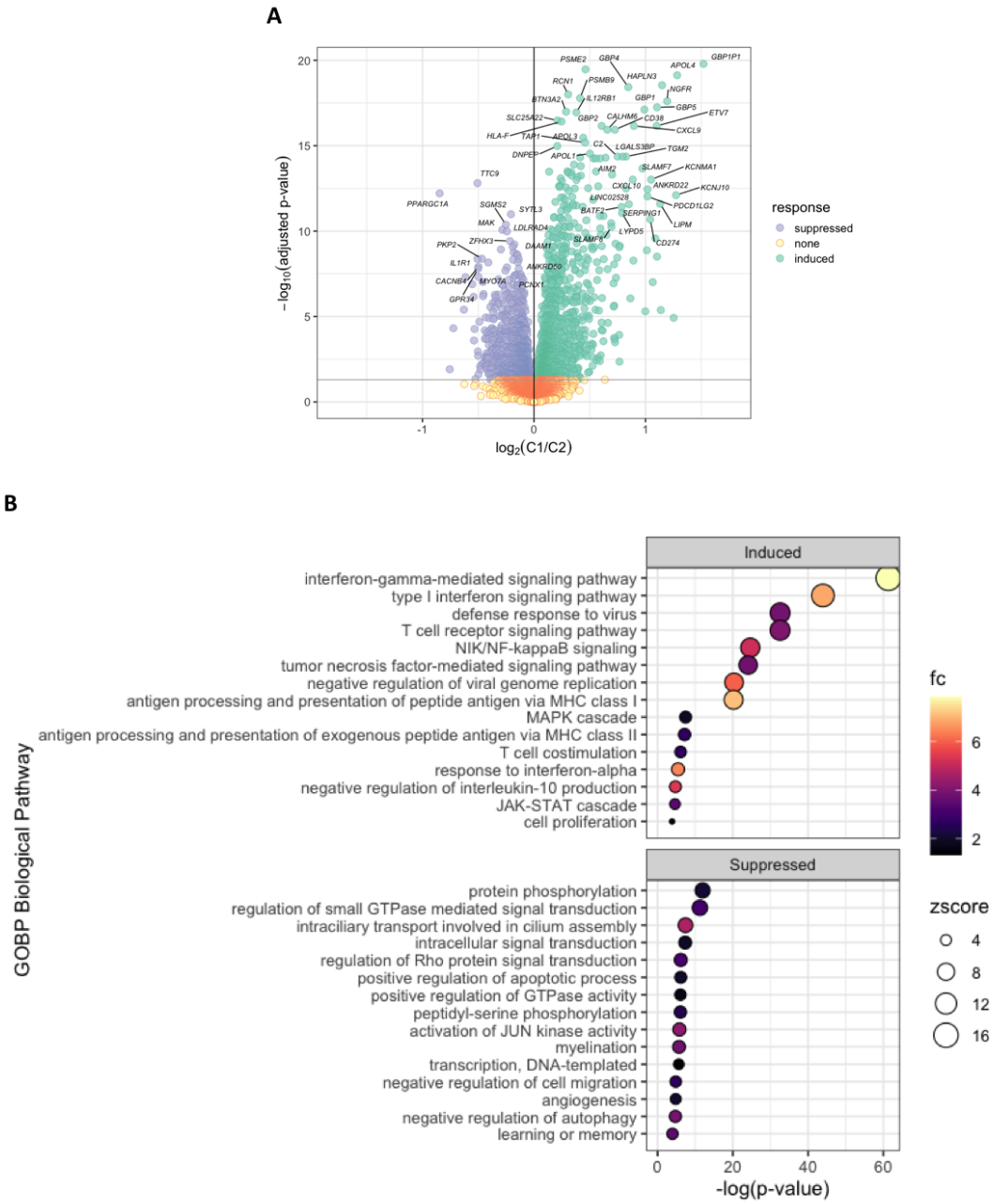


Figure 17: Transcriptional modulation of monocyte profiles in response to ICB

A) Volcano pot to show significantly modulated transcripts (2983, P_{adj}<0.05) with ICB (pairwise analysis across 96 patients). B) Dot plots to show key significantly enriched GOBP pathways induced and suppressed with ICB across all patients (P_{adj}<0.05).

3.3.4.1 Comparing responses to combined ICB (cICB) and single-agent ICB (sICB)

Combination anti-PD-1 and anti-CTLA-4 treatment (cICB) therapy is associated with improved OS and PFS compared to single-agent anti-PD-1 therapy (sICB)[32], but also a higher incidence of immune-related adverse-events (irAEs[40]). Within the $CD8^+$ T cell compartment, cICB induces a greater magnitude of transcriptomic response compared to sICB, although the phenotypic effect is equivalent[71]. To explore differences in modulation of monocyte transcriptomic profiles between the two treatment regimens, I performed treatment-wise differential expression analysis. This was performed via pairwise differential expression analysis with *DESeq2* across those receiving sICB and cICB, in turn. Differential expression analysis across those treated with sICB identified only 85 transcripts differentially modulated with treatment (*DESeq2* pairwise analysis, n=45 patients). In patients receiving cICB, a total of 2777 transcripts were differentially expressed (*DESeq2* pairwise analysis, n=51 patients). These included 1,423 transcripts that were induced, and 1,354 suppressed with cICB (Figure 18). A total of 85 transcripts were modulated in monocytes following sICB (*DESEq2* pairwise analysis, n=45 patients). There was near complete overlap of the transcripts modulated by the two treatment regimens, with 80/85 transcripts significantly modulated by sICB also significantly modulated by cICB (Figure 18C, Figure 3D).

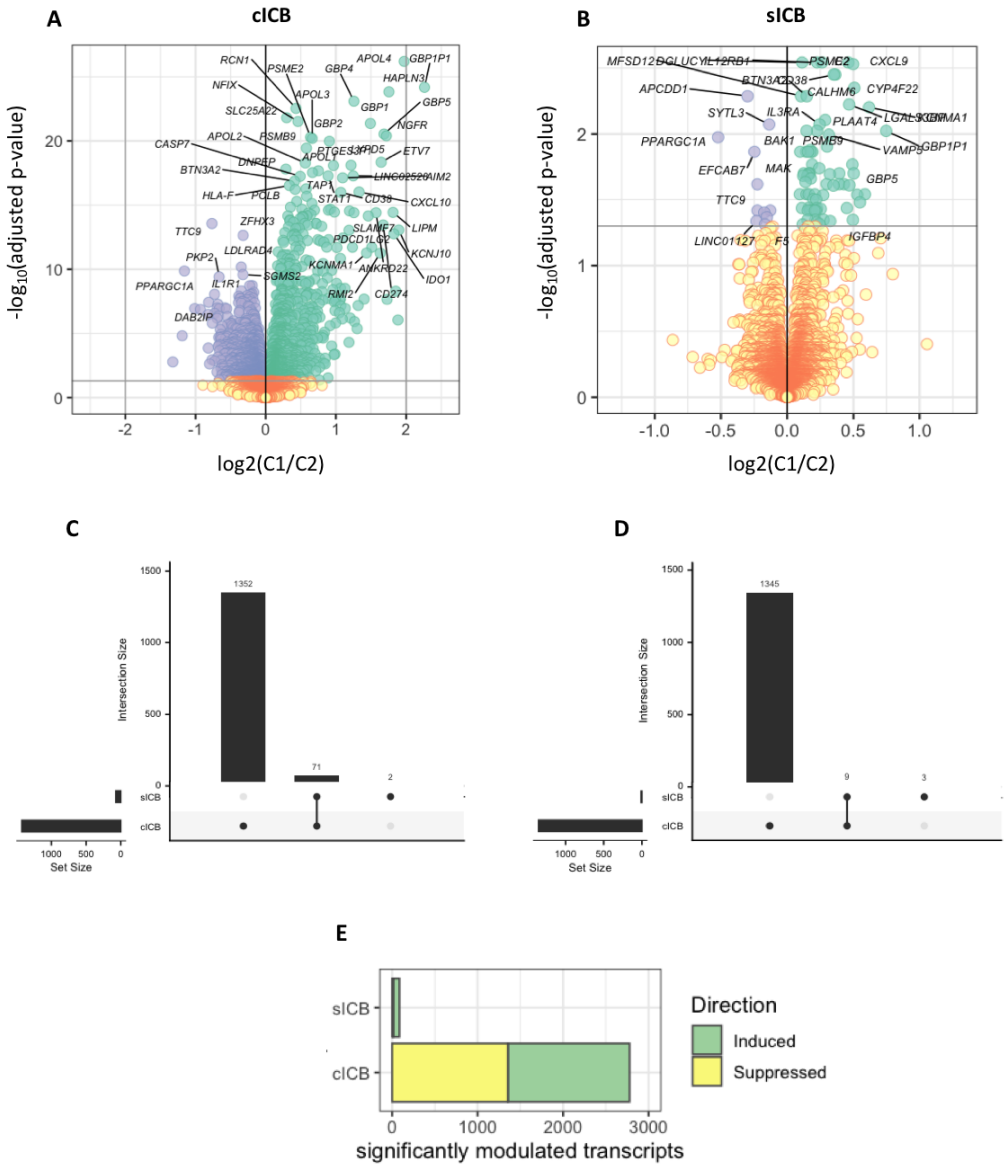


Figure 18: Treatment-wise responses to ICB

A) Volcano plot to show significantly differentially regulated transcripts in response to cICB (DESEq2 pairwise analysis, 51 patients, P_{adj}<0.05). B) Volcano plot to show significantly differentially regulated transcripts with sICB (DESEq2 pairwise analysis, 45 patients). C) Upset plot to show overlap between significantly induced transcripts between cICB and sICB. D) Upset plot to show overlap between significantly suppressed transcripts in cICB and sICB datasets. E) Bar plot to show number of significantly induced and suppressed transcripts identified with pairwise analysis in patients receiving sICB and patients receiving cICB (P_{adj}<0.05).

Comparison of sICB and cICB responses demonstrated a shared direction of effect. Although many more transcripts were differentially expressed with cICB than sICB and the magnitude of effect was greater with cICB, the direction of effect was strongly

correlated (Figure 19A). Using interaction terms, differential expression analysis using *DESeq2* was performed to identify transcripts preferentially modulated by cICB. A total of 352 transcripts were differentially regulated by cICB compared to sICB, including 240 preferentially induced. Pathway enrichment analysis of preferentially induced versus suppressed genes between cICB and sICB showed enrichment of anti-viral type response pathways including type I IFN and IFN-γ -related signalling, as well as T cell receptor signalling (Figure 19B). This demonstrates that as in peripheral $CD8^+$ T cells, cICB has a similar but more pronounced effect on transcriptional modulation of peripheral monocyte populations than sICB.

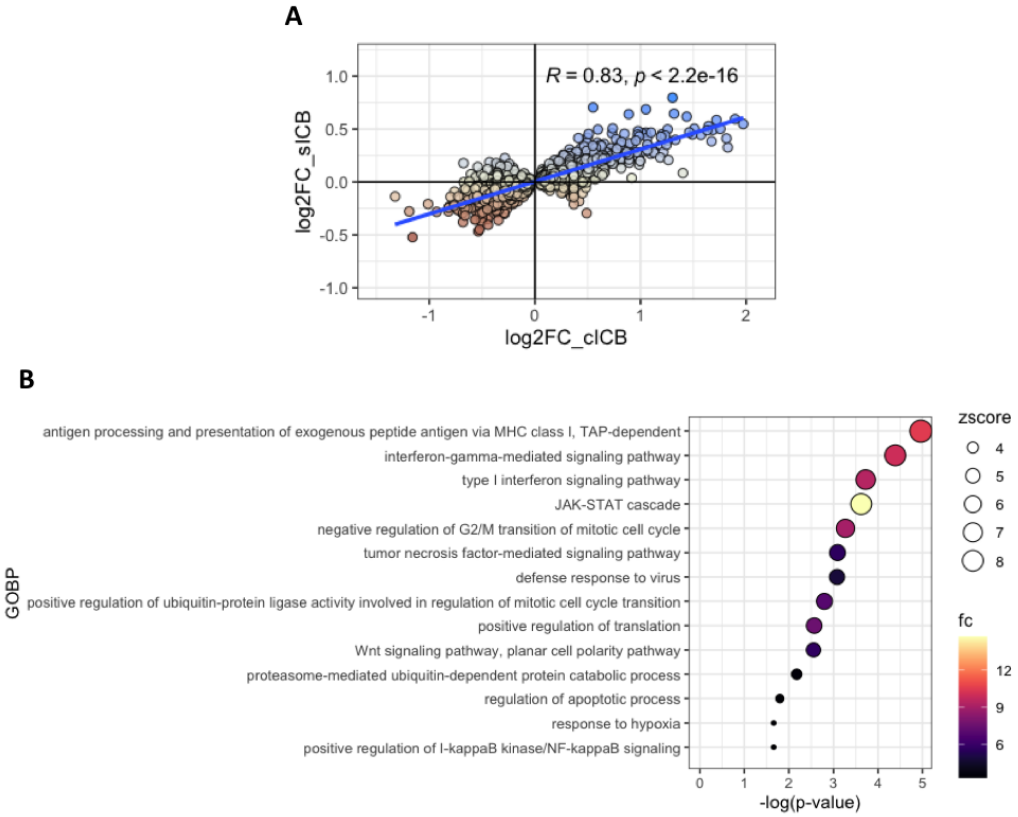


Figure 19: Differential modulation of transcripts with cICB versus sICB
A) Plot to show correlation between log_2 fold change (FC) of transcripts significantly differentially expressed with cICB ('log2FC_cICB') (n= 45 patients) versus the log_2 fold change with sICB ('log2FC_sICB') (n=51 patients) (Pearson's correlation). B) Dot plot to show key preferentially enriched GOBP pathways with cICB (P_{adj}<0.05).

3.3.5 ICB is associated with sustained transcriptional modulation in peripheral monocytes

To explore whether modulation of transcriptional profiles was sustained beyond the second treatment cycle, differential expression analysis was performed between paired samples at baseline (C1) pre-treatment and just prior to the fourth treatment dose (C4), which for most patients will be around day 64 after commencing ICB. A total of 242 transcripts were significantly differentially expressed at C4 (*DESEq2* pairwise analysis across 22 patients, $P_{adj}<0.05$).

Genes that were induced at C4 included complement genes *C1QA* (Padj= 0.0068)*, C1QB* (P_{adj} = 0.020), *C1QC* (P_{adj} = 0.022) and *C2* (P_{adj} = 0.011), the non-classical monocyte marker gene *FCGR3A* (P_{adj} = 0.0090) and antigen presenting genes *HLA-DPA1* (P_{adj} = 0.042) and *HLA-DPB1* (P_{adj}= 0.049). Transcripts associated with interferon responses were also upregulated, including *GBP1* (P_{adj}= 0.036), *GBP2* (P_{adj}= 0.027) and *GBP5* (P_{adj}= 0.044), and those involved in the JAK-STAT pathway signalling including *STAT1* (P_{adj}= 0.042) and *JAK2* (P_{adj}= 0.043).

The *CD177* transcript, which has been shown to be involved in neutrophil proliferation[404], was most significantly downregulated at C4 (P_{adj} = 0.0068). Other downregulated transcripts included the pro-inflammatory cytokine receptor *IL7R (*P_{adj} = 0.022) which is induced in response to LPS[405]. *IL7R* is a risk locus for autoimmune susceptibility[405]. Chemokine *CCL5* (P_{adj} = 0.0068), chemokine receptor *CCR7* (P_{adj} = 0.021) were also downregulated. A number of T cell markers were downregulated (CD8A (P_{adj} = 0.011)*, CD8B* (P_{adj}= 0.014)*, CD3D (*P_{adj}= 0.041)*, CD3E (*P_{adj}= 0.037) and *CD3G* (P_{adj}= 0.0495), which might be reflective of differential monocyte-$CD8^+$ T cell adhesion or 'stickiness'.

Pathway enrichment analysis demonstrated enrichment of IFN-γ-mediated signalling (GO:0060333, $P_{adj} = 1.6x10^{-9}$), type I interferon signalling pathway (GO:0060337, $P_{adj} = 1.4x10^{-6}$), defence response to virus (GO:0009615, $P_{adj} = 1.6x10^{-3}$) and TNF-mediated signalling (GO:0033209, $P_{adj}= 3.2x10^{-10}$). Angiogenesis (GO:0001525, P_{adj} = 0.01), innate immune response (GO:0045087, P_{adj} = 0.012) and cell surface receptor signalling (GO:0007166, P_{adj} = 0.0001) pathways were downregulated (Figure 20C).

Differential expression analysis was carried out to explore treatment-wise responses at this later treatment cycle. No significant differential regulation of transcripts was seen in patients receiving cICB, although sample size was small (n=7 paired samples). In those receiving sICB (n=14 paired samples) 161 transcripts were differentially modulated (see Appendix J). Although differential modulation failed to reach the level of significance in the cICB cohort, there was correlation between the transcript expression fold change across the significantly differentially modulated sICB gene set in both sICB and cICB cohorts, reflecting the same directional effect (Appendix J). Differential sICB and cICB responses are likely to be due to lack of power due to a small cohort size. Comparison of cICB and sICB responses with interaction terms in *DESeq2* revealed only a single modulated transcript (*RN7SL138P*, upregulated in sICB versus sICB). However, it should be noted that analysis at C4 may be confounded by exclusion of both those who develop irAEs (typically those receiving cICB), or those who progress (typically those reciving sICB).

Overall, this demonstrates that at the fourth treatment cycle (C4), distinct transcriptomic profiles can be seen within the peripheral monocyte compartment, reflecting ongoing and sustained interferon and viral-type responses and downregulation of inflammation and myeloid cancer-associated pathways such as chemotaxis and angiogenesis pathways. These effects may reflect both direct and indirect responses to both ICB and cancer.

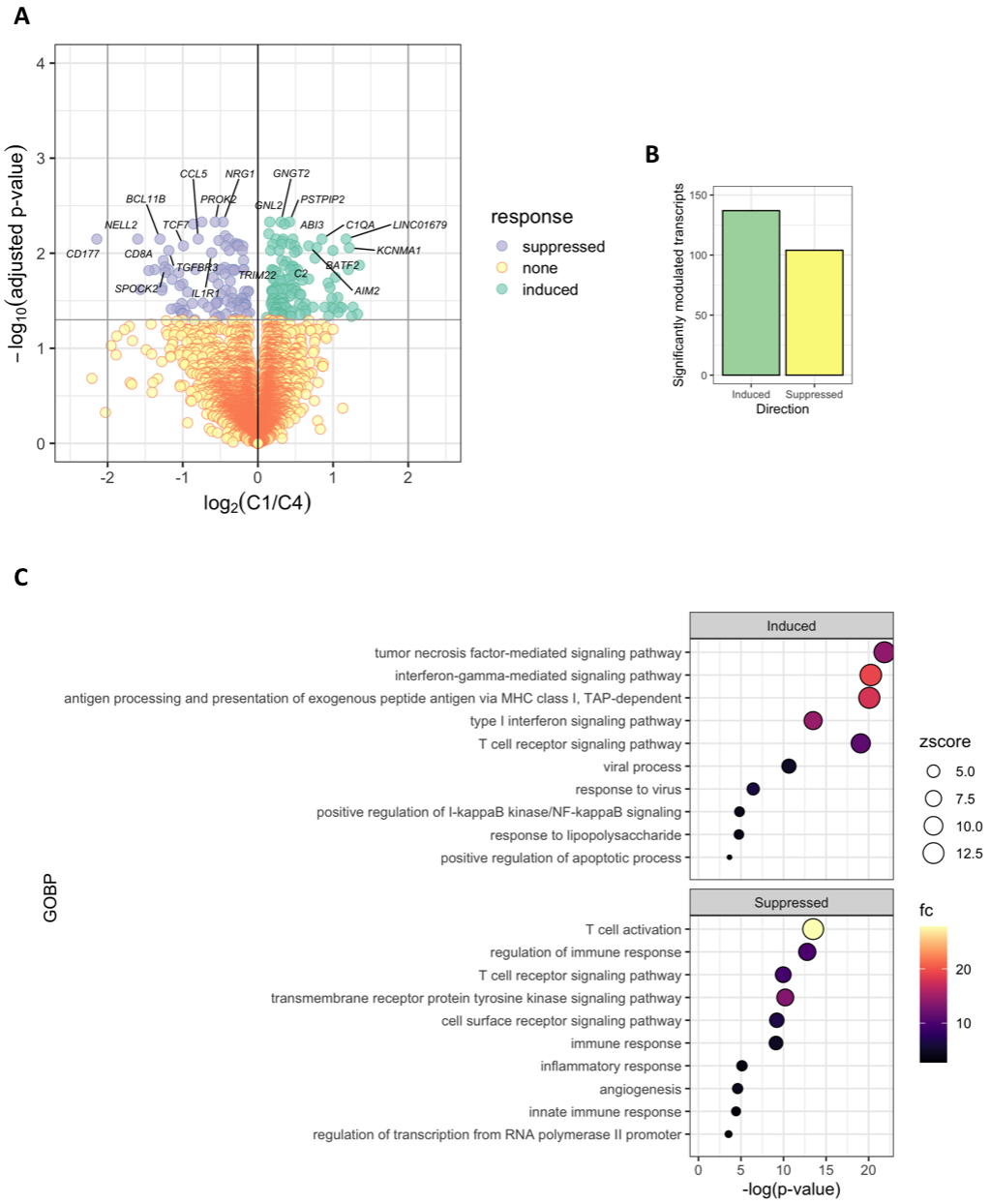


Figure 20: Transcriptomic profiles at Cycle 4 (C4)

A) Volcano plot to show most significantly differentially expressed transcripts at C4 (day 63) versus baseline (C1) (22 paired samples including 14 sICB patients and 7 cICB patients, P_{adj}<0.05). B) Plot to show number of significantly differentially induced (n=137) and suppressed (n=105) transcripts identified with pairwise differential expression analysis at baseline versus C4 (22 paired samples, P_{adj}<0.05). C) Dot plot to show all significantly differentially enriched GOBP pathways across all induced and suppressed transcripts with pairwise baseline versus C4 differential expression analysis (n= all 22 samples), P_{adj}<0.05.

3.3.6 Transcriptomic monocyte profiles are associated with clinical response to ICB

Differential expression analysis was then carried out to identify transcriptional correlates of response to ICB. Clinical outcome parameters including six-month

progression and the development of irAEs prior to the fifth cycle of treatment (C5) were used to dichotomise patients, and *DESeq2* analysis performed controlling for age, sex and where applicable, treatment status. I explored transcriptomic association across both C1 and C2 samples integrated together, and at baseline in untreated samples (C1) and post-treatment (C2) in turn.

A total of 61 transcripts were associated with six-month progression (97 C1 samples and 84 C2 samples from 99 patients, $P_{adj}<0.05$). Transcripts associated with progression included mitochondrial encoded cytochrome C oxidase gene *MTCO3P12* ($P_{adj}= 3.2x10^{-15}$), *MTUS1* ($P_{adj}= 9.2x10^{-5}$) and *MSLN* ($P_{adj}= 0.00010$). Expression of chemokine *CCL4* ($P_{adj}= 0.044$) and cytokine receptor *IL3RA* ($P_{adj}= 0.037$) were found to be negatively associated with six-month progression. At baseline, four transcripts were associated with six-month progression (97 C1 patients, $P_{adj}<0.05$), all of which were positively associated with progression, and which include proliferation index marker *MKI67* ($P_{adj}= 0.029$), *MTCO3P12* ($P_{adj}= 2.1x10^{-6}$) and endothelial gene *MMRN1* ($P_{adj}= 0.04$). Following ICB treatment, at C2 two transcripts were associated with six-month progression (84 samples): *MTCO3P12* ($P_{adj}= 6.7x10^{-6}$) and *IGHD* ($P_{adj}= 0.0042$), both of which showed positive association with disease progression.

There was greater association of transcriptomic profiles with the development of irAEs. irAEs were ascertained as per the Common Terminology Criteria for Adverse Events (CTCAE) version 4.03[317]. A total of 204 transcripts were significantly associated with the development of irAEs prior to cycle 5 ('AI_c5') across both treated and untreated samples (97 C1 samples, 82 C2 samples from 99 patients, $P_{adj} <0.05$). Interestingly, *CD274* expression was positively associated with associated with the development of irAEs ($P_{adj}= 0.037$), which encodes PD-L1. Other transcripts include genes associated with the response to IFN-γ including *GBP5* ($P_{adj}= 0.0077$) and *IFI27* ($P_{adj}= 0.00044$) as well as the chemokine *CCL3L3* ($P_{adj}= 0.010$). Transcripts negatively associated with irAEs include *IL1R2* ($P_{adj}= 0.00011$), *IL18R1* ($P_{adj}= 1.3x10^{-5}$), *MMP8* ($P_{adj}= 0.006$) and the proliferation index *MKI67* ($P_{adj}= 2.9x10^{-5}$).

Pathway enrichment analysis of transcripts negatively associated with irAEs identified 11 GOBP pathways including neutrophil degranulation (GO:0043312, $P_{adj}=3.5x10^{-9}$),

platelet activation (GO:0030168, P_{adj}= 0.00083), blood coagulation (GO:0007596, P_{adj}= 0.0054) and cell migration (GO:0030335, P_{adj}= 0.00017). These were notable as being many of the cancer-associated pathways enriched in monocytes in patients versus healthy controls. No pathways were enriched across the positivity correlated gene transcript set. This is in keeping with a switch from a healthy immune state which can facilitate an anti-tumourigenic immune response associated with irAEs, towards immune exhaustion which is associated with propensity to irAEs.

Limited to baseline (C1) transcriptomic profiles, 37 transcripts were significantly associated with irAEs prior to cycle 5 (37 samples), of which 35 showed negative association with irAEs prior to cycle 5. Notably, the proliferation marker *MKI67* was most significantly associated with irAEs (P_{adj}= $3.0x10^{-6}$), showing negative association with irAEs. Other significantly negatively correlated transcripts include *IL18R1* (P_{adj}= $1.2x10^{-2}$) and the IL-1 decoy receptor *IL1R2* (P_{adj}= 0.013).

Post-treatment differential expression analysis with irAEs prior to cycle 5 identified only six transcripts, including *HAPLN3* (P_{adj}= 0.0057) and *PTMAP2* (P_{adj}= 0.0057) which were positively associated; and *MTUS1* (P_{adj}= 0.00084), IGHD (P_{adj} = 0.00084) and *CPXM1* (P_{adj}= 0.0011), which were negatively associated (*DESeq2* analysis, 82 C2 samples).

It is notable that *IL1R2* expression was found to be associated with the development of irAEs. IL-1 signalling via IL-1R is key in facilitating pro-tumourigenic actions such as angiogenesis and invasion in RCC, with IL-1β expression positively associated with tumour stage[360]. The decoy receptor IL1R2 competitively binds IL-1 and thus has inhibitory effect on IL-1 signalling[406]. Negative association of IL1R2 expression in monocytes and irAEs suggests that the favourable prognostic and anti-tumourigenic effects associated with these irAEs may be in part mediated by the IL-1/IL-1R axis. Association of *MKI67* expression and clinical outcome was also of particular interest. Anti-correlation of *MKI67* expression with irAEs suggests that monocyte turnover and cycling have an immuno-inhibitory role. In addition, matrix metalloproteinase-8 (MMP8) is a marker of macrophage differentiation, and may be associated with pro-tumourigenic M2 macrophage phenotype[407]. Anticorrelation between *MMP8*

expression and irAEs provides a link between myeloid pro-tumourigenic profiles and irAEs.

Overall, these observations demonstrate that peripheral transcriptomic profiles may be potential predictive biomarkers in patients treated with ICB.

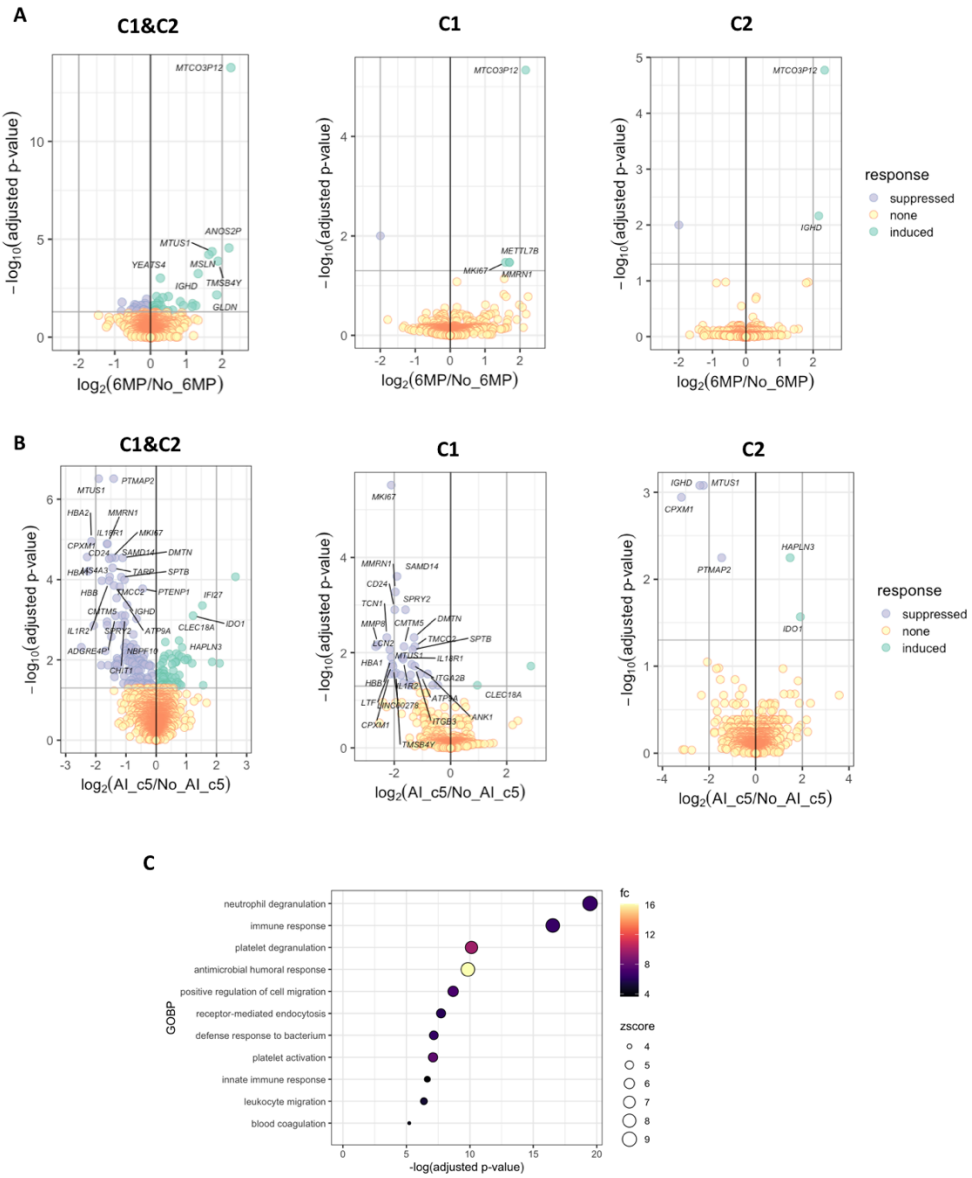


Figure 21: Transcriptional correlates of clinical response in peripheral monocytes
A) Volcano plot to show transcripts associated with disease progression at six months across treated and untreated samples ('C1&C2'), and separately in untreated ('C1') and treated samples ('C2') (DESeq2 analysis, controlling for age, sex and cycle (i.e. whether treated) where appropriate) B) Volcano plots to show transcripts associated with the development of irAEs prior to cycle 5 ('AI_c5') and those who do not ('No_AI_c5') (DESeq2 analysis, controlling for age, sex and cycle where appropriate). C) Dot plot to show pathways significantly enriched across transcripts negatively associated with developing irAEs prior to cycle 5 (P_{adj}<0.05).

3.3.7 Exploring peripheral monocyte subset proportions with flow cytometry and clinical outcome

In this chapter, many of the transcriptomic profiles described between conditions are suggestive of subset-wise responses and phenotypes. To explore whether

transcriptomic profiles in peripheral monocytes may be reflective of changes in peripheral monocyte subset populations, flow cytometry analysis of PBMCs was performed. As described in the Methods (section 2.4), the proportion of monocytes subsets (classical, intermediate and non-classical) was determined using flow cytometry, and the size of the monocyte group as a proportion of all PBMCs per sample calculated. Flow cytometry was carried out by Ms Chelsea Taylor.

No significant change in monocyte subset proportion was seen with ICB, either at C2 or C4, nor did the monocyte as a proportion of PBMCs change with treatment (46 C1 patients, 46 C2 patient, 18 C4 patients, Wilcoxon signed-rank test) (Figure 22A). There was also no change in monocyte subset size across treatment states when characterised by treatment type (see Appendix L).

To explore association between monocyte subset size and clinical outcome, Wilcoxon signed-rank test was used to determine association between monocyte subset size and either six-month progression or irAEs. Survival analysis was then performed using the *survminer*[339] *and survival* packages[340] in R and the variable of interest e.g. classical monocyte proportion, dichotomised according to values above or below the median value (OS analysis: n=46 C1 patients, n=46 C2 patients, n=18 C4 patients; PFS analysis: n=45 C1 patients, n=45 C2 patients, n=18 C4 patients).

Whereas most studies have demonstrated baseline monocyte count to be predictive in ICB[196,358], here it was following treatment that monocyte proportion was associated with less favourable response to ICB with a higher monocyte population as a proportion of all PBMCs associated with reduced PFS at following ICB treatment (at C2, 45 patients) (Figure 22D).There was no significant association between peripheral monocyte subset sizes with OS, PFS, six-month progression or the development of irAEs at baseline (C1) or C2 (day 21). However, a higher intermediate monocyte proportion (dichotomised into above/below the median subset proportion) was significantly associated with lower OS and reduced PFS at C4 (18 patients) (Figure 22B&C).

This demonstrated that the size of the circulating monocyte population may have predictive utility post-ICB treatment, is associated with an unfavourable response to ICB, and that this effect may be subset-dependent.

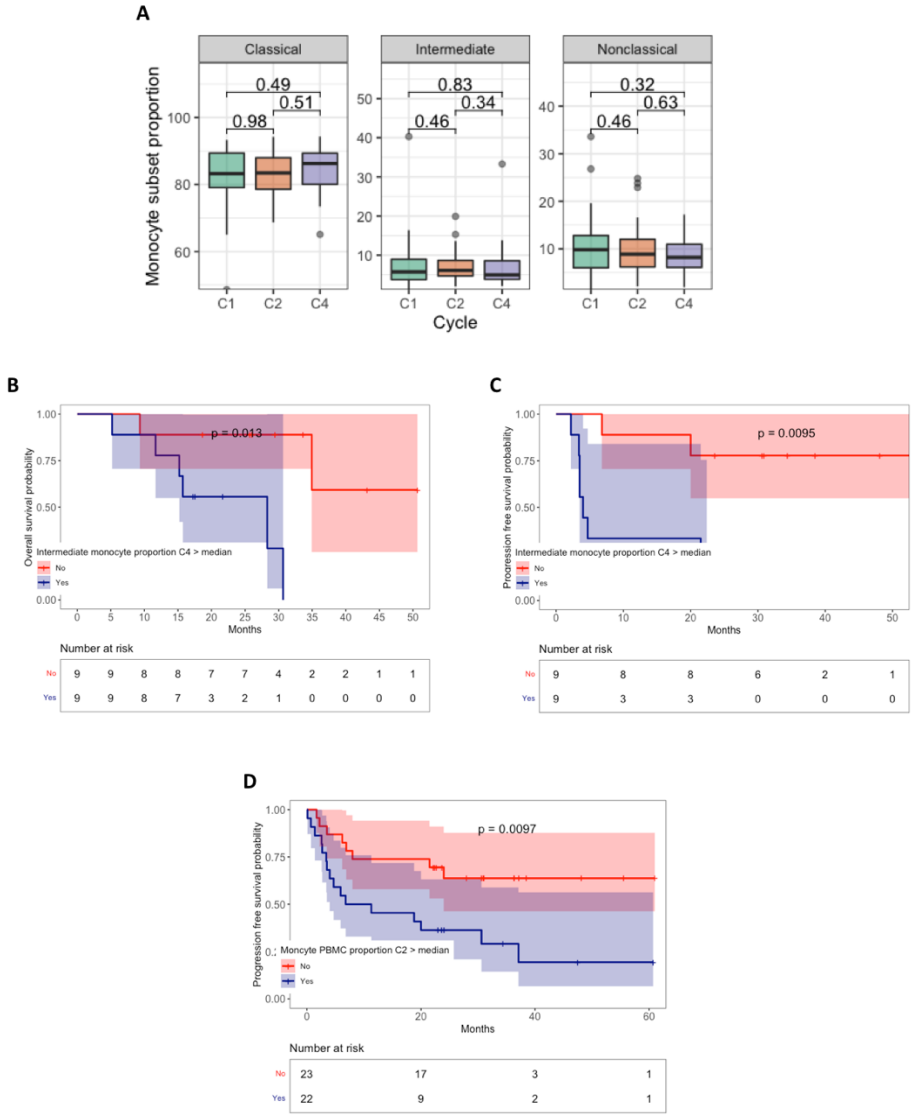


Figure 22: Monocyte subsets assessed with flow cytometry

A) Boxplots to show monocyte subset (classical, intermediate and non-classical) as a proportion of the total peripheral monocyte population as assessed by flow cytometry across C1 (n=46 patients), C2 (n=46 patients) and C4 (n=18 patients) (Wilcoxon signed-rank test). B) Kaplan-Meier curve to show intermediate monocyte subset size as a proportion of monocyte population above/below median) at C4 and overall survival (OS) (survival analysis across 18 patients). C) Kaplan-Meier curve to show intermediate monocyte subset size as a proportion of all monocytes (above/below median) at C4 and progression free survival (PFS) (survival analysis across 18 patients). D) Kaplan-Meier curve to show monocyte population as a proportion of PBMCs (above/below median) at C2 and progression free survival (PFS) (45 patients). Flow cytometry was carried out by Ms Chelsea Taylor at the MRC Weatherall Institute of Molecular Medicine, University of Oxford.

3.3.8 Peripheral monocyte subset size is associated with monocyte transcriptomic expression

To determine if subset-wise transcriptomic monocyte profiles could be identified within the bulk RNA-seq dataset, differential expression analysis was carried out using *DESeq2* to identify transcripts significantly associated with peripheral monocyte subset proportion defined by flow cytometry. The proportion of monocyte subsets as a total of the whole monocyte population was used as the variable of interest in DESeq2, controlling for the size of other monocyte groups.

A total of 93 transcripts were associated with non-classical monocyte subset proportion (controlling for cycle, age, sex and log Classical, intermediate and monocytes as a proportion of PBMCs). The non-classical monocyte subset defining marker *FCGR3A* was positively associated with non-classical subset size (P_{adj}= 0.00059). This is validatory of this approach. Other transcripts significantly differentially regulated according to non-classical subset proportion included interferon-induced genes *GBP4* (P_{adj}= 0.048), *IFIT1* (P_{adj}= 0.0038) and *IFIT2* (P_{adj}= 0.029). Non-classical monocytes have a pro-inflammatory type phenotype with enrichment of interferon-response pathways[85], so the transcriptional correlates observed here are in keeping with this profile.

A total of 111 transcripts were associated with intermediate monocyte proportion (using the same *DESeq2* parameters as previously described) (Figure 23). These included non-classical type gene transcripts, including *FCGR3A* (P_{adj}= $7.7x10^{-7}$), interferon induced genes including *IFITM3* (P_{adj}= 0.022), *IFI44L* (P_{adj}= 0.044). This demonstrates that there may be overlap between transcriptional profiles across non-classical and intermediate monocyte subsets, in keeping with previous studies[100]. It also demonstrates that the non-classical gene *FCGR3A* is a highly variably expressed gene within monocyte populations, changing as a function of monocyte subset proportion size.

No transcripts were associated with classical monocyte proportion size.

This demonstrates that peripheral monocyte subsets populations are associated with distinct profiles of gene expression within bulk RNA-seq data, however further elucidation of monocyte cluster-wise profiles and responses within a bulk RNA-seq dataset is limited by experimental parameters within flow cytometry data and

assessment of three conventional monocyte subsets. I examine this using scRNA-seq in Chapter Four.

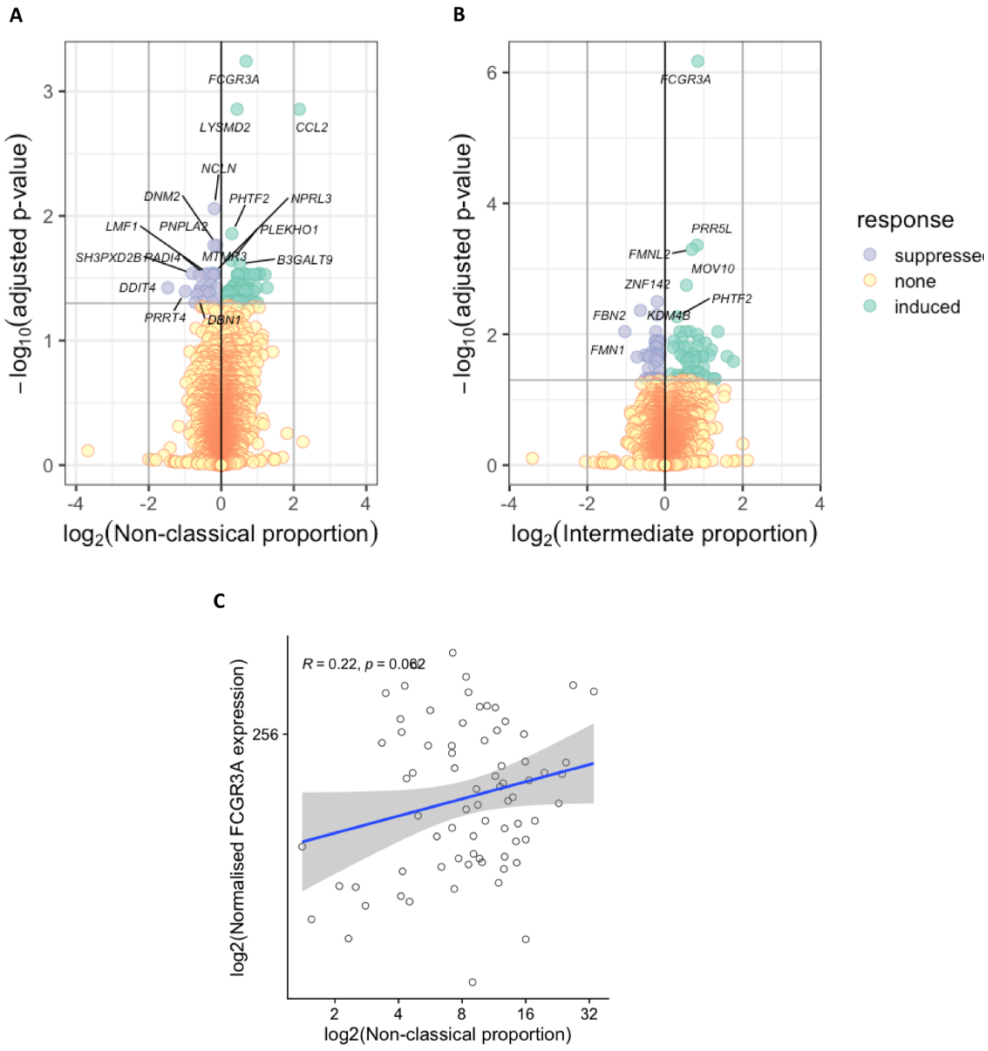


Figure 23: Transcriptomic profiles modulated according to peripheral monocyte subset proportions

Volcano plots to show most significantly differentially expressed transcripts induced and suppressed with increased size of the: A) non-classical and; B) intermediate, cluster size as a proportion of all monocytes (DESEq2 analysis, controlling for cycle, age, sex, monocytes a proportion of PBMCs and other monocyte subsets across 37 C1 patients, 36 C2 patients). C) Scatterplot to show correlation (Pearson's) between log_2 non-classical peripheral monocyte proportion with flow cytometry and normalised FCGR3A expression within monocytes.

3.4 Summary and Conclusions

A summary of key findings in this chapter are as follows:

- Peripheral monocytes exhibit distinct transcriptomic profiles in MM including enrichment of transcripts associated with chemotaxis, phagocytosis and response to LPS.
- Transcriptomic profiles of peripheral monocytes in patients with MM at baseline may have utility as a diagnostic classifier.
- Peripheral monocyte count is associated with cancer-associated MM expression profiles including cell migration and adhesion, chemotaxis and VEGF signalling. This may be a mechanism through which an elevated monocyte count is associated with poor prognosis in MM.
- ICB is associated with induction of transcripts associated with IFN-γ signalling, anti-viral responses, T cell activation and co-stimulation and type I interferon signalling in peripheral monocytes. Transcriptomic modulation by ICB in monocytes is more prominent in patients receiving cICB versus sICB.
- Transcriptomic responses to ICB in peripheral monocytes are sustained across multiple treatment cycles.
- Baseline and post-treatment transcriptomic profiles in monocytes are associated with clinical outcome in patients receiving ICB. In particular, peripheral monocyte profiles at baseline are associated with the development of irAEs.
- A higher proportion of monocytes compared to other PBMCs after treatment at cycle 2 was associated with reduced PFS, whilst a larger intermediate population at cycle 4 was associated with reduced OS and FPS.

In this chapter I have characterised transcriptomic expression profiling of monocytes within a large melanoma cohort. Transcriptomic profiles associated with a diverse range of biological pathways were differentially modulated in MM patients, including induction of pro-thrombotic and angiogenic pathways, and LPS signalling response. In keeping with studies in different cancer types[176,178], there was induction of chemotaxis

and cell migration-associated pathways. Downregulation of antigen presentation and TNF responses in MM were observed. Reduced expression of MHC class II on peripheral monocytes has previously been reported in melanoma[179], and may represent impaired physiological antigen presenting function and baseline myeloid dysfunction. The reprogramming of myeloid expression profiles in MM patients may reflect mechanisms by which peripheral monocytes have prognostic effects. In particular, the IL-8 receptor genes *CXCR1* and *CXCR2* were induced in monocytes in patients with MM. IL-8 has pro-tumourigenic effects and is associated with reduced response to ICB[371], and *CXCR1* expression by $CD8^+$ T cells has previously been reported to be associated with the absence of irAEs in MM[40]. It may therefore be partly via this mechanism that peripheral monocytes play a role in modulating tumorigenesis, with prognostic and predictive utility in patients treated with ICB.

In this chapter I describe on-treatment modulation of monocyte expression profiles with ICB, including prominent activation of anti-viral and interferon response pathways, complement and upregulation of antigen presenting via MHC class II genes. These pathways may be induced as a result of stimulation by pro-inflammatory cytokines such as IFN-γ, produced by activated T cells, but also demonstrate mechanisms by which circulating myeloid populations might have a co-stimulatory role on effector T cell populations. Recently findings from the Fairfax group show that $CD8^+$ T cell cytotoxicity expression profiles are a key determinant of response to ICB. Baseline expression of cytotoxic gene expression profiles in $CD8^+$ T cells are positively associated with PFS and OS in patients treated with ICB[408]. Furthermore, $CD8^+$ T cell clone size is also predictive of treatment response in this context. Whilst beyond the scope of this thesis, it would be informative to explore the association between $CD8^+$ T cell cytotoxicity and clonal diversity with circulating myeloid populations and expression profiles.

Downregulation of cancer-associated transcripts and pathways such as angiogenesis were observed with ICB, may reflect resolution towards the baseline towards health of cancer-associated transcriptional reprogramming. As previously reported in the $CD8^+$ T cell compartment[71], cICB and sICB induce a similar transcriptomic effect, with cICB associated with a greater magnitude of effect. Transcriptional effects are sustained in peripheral monocytes following ICB. However, these profiles might represent markers

of response, as within the C4 cohort rapidly progressing patients are no longer represented. Analysis across a larger cohort controlling for outcome effects, as well as over later treatment cycles, would provide useful insight into late- peripheral myeoid effects following ICB.

In this chapter I describe that transcriptomic profiles in circulating monocytes are predictive of clinical outcome. In particular, the expression of proliferation index marker *MIK67* at baseline (C1) was associated with the absence of irAEs, in keeping with proliferation of the peripheral monocyte compartment negatively predictive of treatment response. Monocytes play a prognostic and predictive role across multiple cancers including melanoma[165,181,196,197,358,359] and these findings provide novel insights into potential underlying mechanisms.

The distinct transcriptional profiles across conditions in MM and following ICB were suggestive of subset-wise responses. For example, many of the MM-enriched transcripts were those associated with a classical monocyte phenotype, such as high expression of chemotaxis and angiogenesis-associated transcripts. The ubiquitous non-classical gene *FCGR3B* was prominently induced with ICB. Furthermore, the induction of interferon signalling, complement and T cell costimulation pathways induced with ICB, are more in keeping with a non-classical monocyte phenotype. Whilst flow cytometry failed to demonstrate change in monocyte subpopulation changes in size with ICB, it may be that these transcriptomic profiles I have described in this chapter are reflective of subset-wise responses, in particular induction of a non-classical type transcriptomic profile.

Overall, this work provides novel insights into peripheral myeloid re-programming in a large melanoma patient cohort and novel on-treatment characterisation of transcriptomic modulation with ICB. This demonstrates that the peripheral monocyte compartment may have utility as both a diagnostic and predictive biomarker. In Chapter Four I use scRNA-seq to explore subset-wise MM monocyte profiles and on-treatment modulation.

4 Chapter 4: Characterising peripheral monocyte response to Immune Checkpoint Blockade with scRNA-seq

4.1 Background

Peripheral monocytes are a heterogeneous myeloid-derived cell population with three recognised subsets defined by the relative antigenic expression of CD14 and CD16[85]. Whilst numerous studies have sought to characterise monocyte subsets, these have typically been limited by sample size, or utilisation of low throughput methods with both few pre-selected gene panels and antigens[93,95,130,179,409]. Although some genic and antigenic markers, and association function, have been robustly described with consensus between studies, many report inconsistencies between transcriptomic and phenotypic profiles. scRNA-seq has allowed for unsupervised clustering of immune cell subsets and high-resolution transcriptomic phenotyping[104]. However, to date few studies have used scRNA-seq to explore monocyte transcriptomic phenotype at high resolution across a large number of cells, within multiple individuals, nor in the context of cancer and ICB.

In Chapter Three I demonstrated that peripheral monocytes exhibit distinct transcriptomic profiles in patients with metastatic melanoma (MM). Furthermore, significant on-treatment transcriptomic modulation following ICB was observed. Transcriptional profiles in peripheral monocytes were also found to be significant predictors of clinical response to ICB in patients with MM. These observations may reflect subset-wise responses. However, studies regarding monocyte subset-wise responses and function are limited, and results are conflicting. For example, Romano et al.[203] found a higher $CD16^{+}$ non-classical monocyte population at baseline was associated with response to ipilimumab in patients with melanoma, whilst Krieg et al.[73] found that baseline classical monocyte populations were associated with prolonged OS and PFS in patients with MM treated with anti-PD-1 therapy. The lack of consensus regarding the myeloid contribution to modulation of ICB responses may be in part due to delineation of monocyte responses to specific subsets amongst a transcriptionally and functionally heterogenous population.

In this Chapter, I outline how I used scRNA-seq to perform high-resolution characterisation of subset-wise monocytes transcriptomic profiles across multiple individuals with MM and healthy donors. Monocyte cluster-wise MM-associated profiles and responses to ICB are explored. Association of cluster-wise expression profiles with clinical outcome is also characterised.

4.2 Aims and Objectives

The aims and objectives of this chapter were as follows:

- To characterise transcriptomic profiles of monocyte subsets with scRNA-seq
- To characterise MM-associated peripheral monocytes transcriptional profiles in a cluster-wise fashion
- To characterise ICB responses across monocyte sub-populations
- To explore associations between monocyte expression profiles, monocyte subsets and clinical outcome

4.3 Results

4.3.1 Characterising monocyte subsets and phenotypes using scRNA-seq

I used scRNA-seq to characterise peripheral monocytes subsets in patients with MM receiving treatment with ICB. In total, eight paired pre- (C1) and post- (C2) treatment peripheral blood samples were obtained from MM patients. This included samples from four patients who received sICB (pembrolizumab) and four who received cICB (ipilimumab/nivolumab). This sample size balanced the requirement to provide insights into monocyte subset-wise inter-individual variation and response within the confines of clinical and experimental feasiblility of this project. A second batch of scRNA-seq was performed with samples from three healthy donors (see Appendix M for cohort details). Both $CD14^+$ and $CD8^+$ cells were isolated using magnetic-activated bead sorting (MACS) from PBMCs, and scRNA-seq library preparation and sequencing performed using the Chromium 10x platform. Following upstream processing as described in section 2.5.4, monocytes ($CD14^+$ cells) were selected for analysis by selecting cells expressing *CD14*, and removing those expressing *CD8A, CD3D, CD3E, CD3G, CD56, CD19* or *CD20.* The $CD8^+$

T cell response to ICB defined by these data is described by Fairfax et al.[71] and Watson et al.[408].

After filtering to select monocytes, cells expressing at least 300 genes and 500 reads were selected for downstream analysis. Cells with more than 20% of genes comprising mitochondrial genes were removed to remove low-quality cells, and those that may be dead or dying, which typically exhibit a high mitochondrial fraction. Genes expressed by fewer than five cells were also removed from downstream analysis. Following these QC and filtering steps, pre-treatment data ('C1') comprised a total of 9321 cells, and post-treatment data ('C2') 10989 cells. There was a median of 5504 post-normalisation read counts reads per cell and 1873 genes per cell. For the second batch comprising monocytes from healthy donors there were a total of 4147 cells with 5514 median reads per cell and 1824 median genes per cell.

4.3.2 Annotation of monocyte subsets identifies conventional monocyte clusters

Firstly, I performed unsupervised clustering with the *Seurat*[332] package in *R* to identify conventional monocyte subsets. Normalised data was then subject to variable feature selection, and control and patient datasets were integrated. Principal component analysis (PCA) was performed to reduce dataset dimensionality, visual inspection of the standard deviations of principal components (PCs) was performed using the *Seurat* function *Elbowplot* (see Appendix N for Elbowplot visualisation of PCs). The top 13 PCs were used for dimensionality reduction and clustering as this was the level at which there was minimal variation between successive PCs[410].

To determine an appropriate resolution for clustering to identify conventional monocyte subsets (i.e. classical, intermediate and non-classical), clustering at varying resolutions was performed across all monocytes (n=24,457 cells). Subsequent annotation was then performed by firstly characterising monocyte subsets according to key marker gene expression and secondly, with comparison to published gene expression datasets, as outlined as follows.

4.3.2.1 Visualising clustering trees

To aid selection of an appropriate level of clustering resolution, cells were clustered iteratively at varying resolutions and expression of a key marker genes visualised with the *clustree* package[333]. With increasing clustering resolution, greater numbers of monocyte subsets were identified (Figure 24). This approach identified a branch point delineating between monocytes with high expression of key non-classical and intermediate marker genes, such as *FCGR3A, C1QA*, *CDKN1C*, versus those that exhibited low expression of these genes (Subset 7, Figure 24/Appendix P). This subset (subset 7) exhibited low expression of classical marker genes *CD14* and *SELL* (Appendix O). This subset remained distinct and stable at increasing clustering resolutions (Figure 24). This approach suggested that subset 7 (at clustering resolution 0.25) was consistent with a non-classical/intermediate monocyte subset (Figure 24).

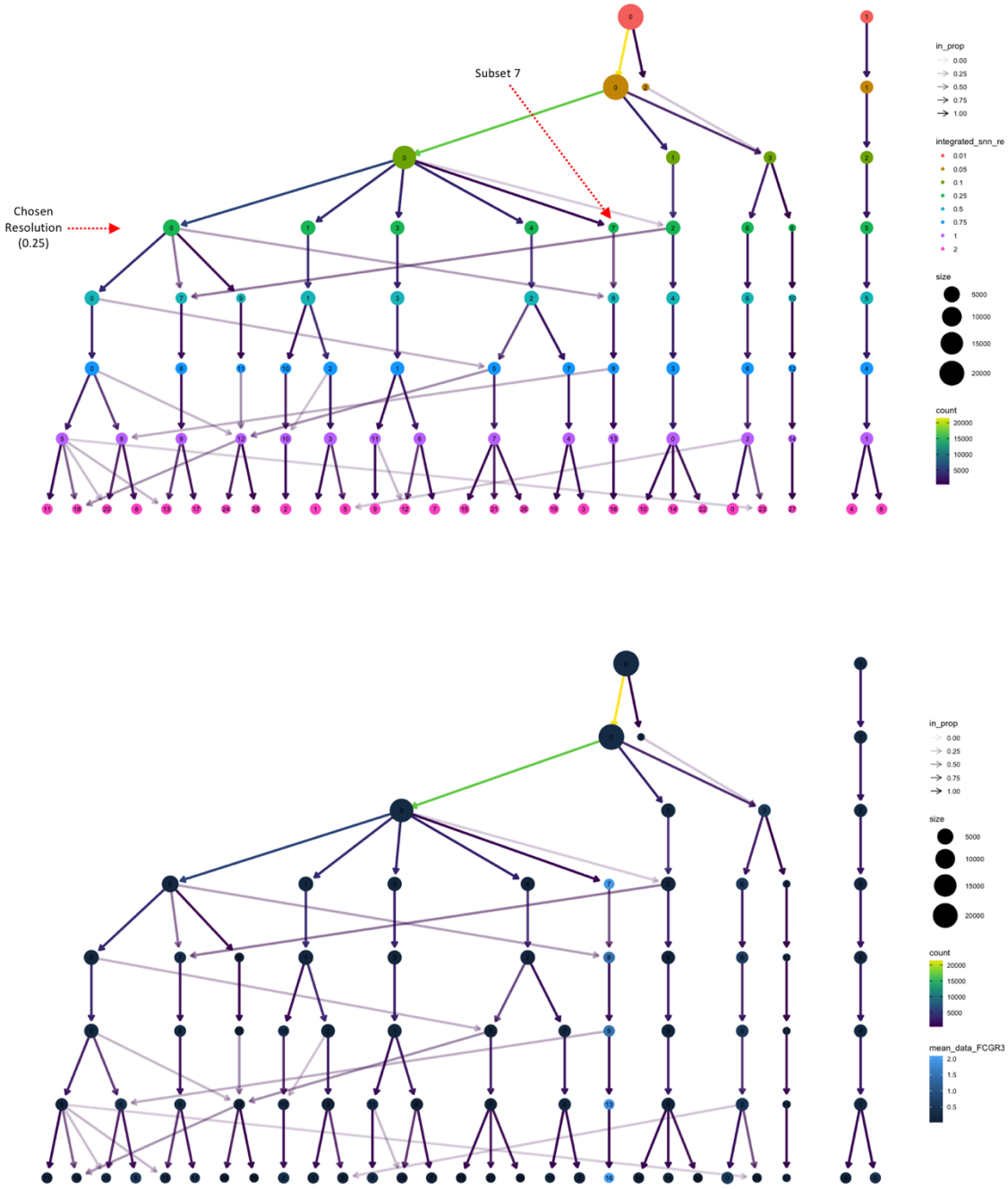


Figure 24: Clustree plots to show monocyte scRNA-seq clusters at varying resolution
Top plot: Clustree plot to show cluster tree with monocyte clustering at varying resolutions (0.01, 0.05, 0.1, 0.25, 0.5, 0.75, 1 and 2). Colour of dots represents clustering resolution. Size of dot represents size of cluster. Red dotted arrows indicate subset 7 and the chosen clustering rsolution of 0.25. Bottom plot: Clustree plot to show cluster tree and mean expression of FCGR3A across monocyte clusters at varying resolutions.

4.3.2.2 Characterising gene expression of key transcripts across the scRNA-seq dataset

At the clustering level at which the $CD16^+$ monocyte population ('Seurat subset 7') was distinct from the remaining monocyte population (clustering resolution 0.25), eight other monocyte subsets were identified (Seurat subsets 0-6 and subset 8). This level of clustering was chosen for further annotation. Please note that the terms cluster and subset were used largely interchangeably, as they describe group and sub-population of monocytes respectively.

To explore the expression of key marker genes across subsets at this level of clustering, the scRNA-seq dataset was visualised using uniform manifold approximation and projection (UMAP) plot (Figure 25A). On visual inspection, non-classical markers including *FCGR3A, C1QA*, and *CDKN1C* were most highly expressed in subset 7 (Figure 25). MHC Class II antigen presenting gene *HLA-DRA,* which is typically most highly expressed in intermediate monocytes, was also most highly expressed in subset 7. Markers of a classical phenotype such as *CD14, CCR2* and *SELL* showed low expression in subset 7 and higher expression across subsets 0-6 and 8 (Figure 25). However, it was noted that visualisation of expression levels of key transcripts across the scRNA-seq dataset demonstrated lack of specificity, with *HLA-DRA* and *CD14* expressed across the dataset (Figure 25B&C). Expression of these markers reflects the role of monocytes in the innate immune response via pattern recognition[113] and antigen presentation.

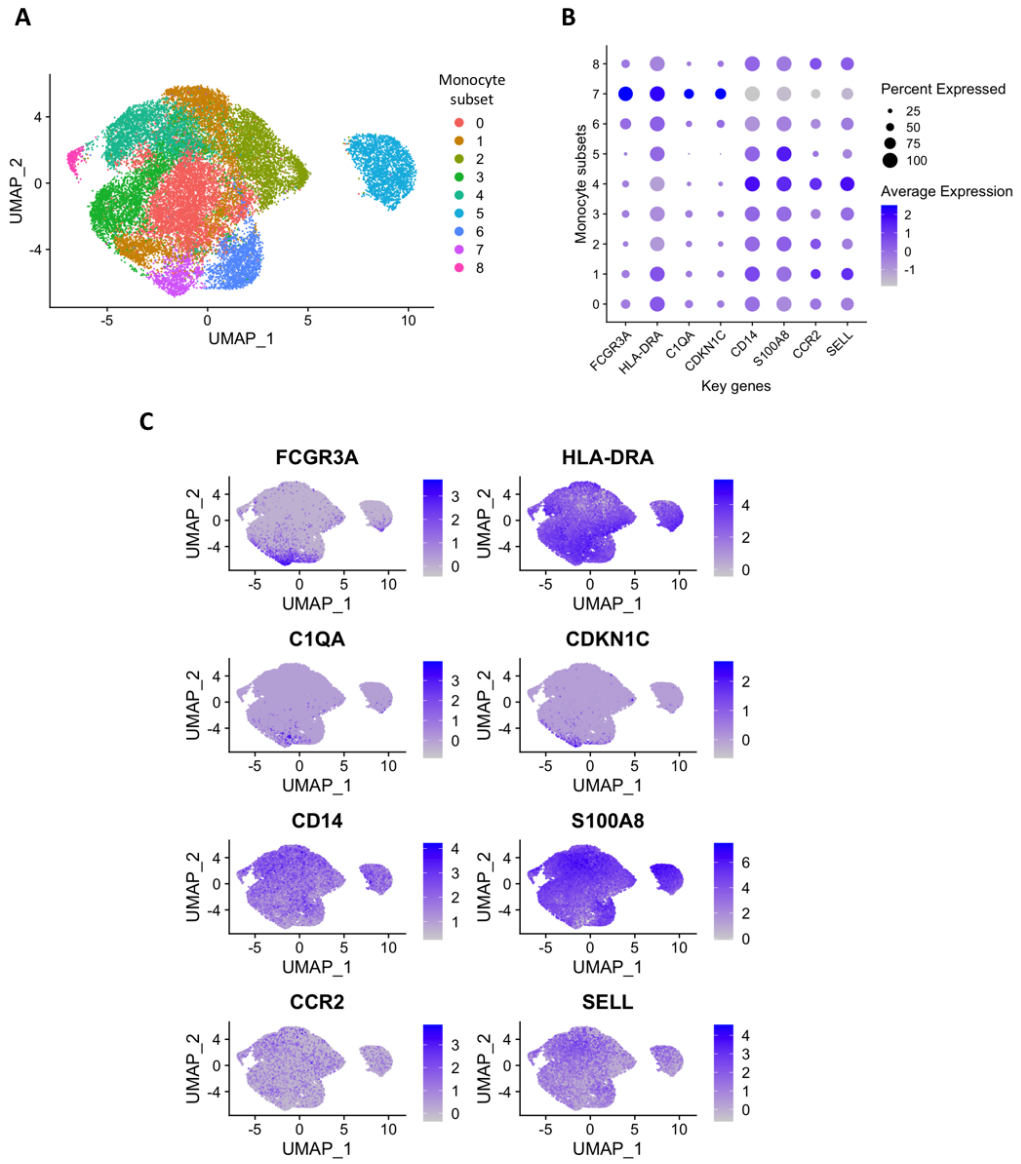


Figure 25: Expression of key gene markers
A) UMAP plot to show monocyte subsets at resolution 0.25 (subsets 0-8). Each dot represents a cell and colours distinguish each subset identity at this clustering level. B) Dotplot to show gene feature expression within each monocyte subset (0-8). Size of the dot indicates percentage of cells expressing a gene ('Percentage Expressed') and the 'Average Expression' denotes a scaled representation of expression across the dataset, with 0 indicating mean expression. C) Featureplot showing the relative expression of key transcripts across the scRNA-seq dataset. Scale indicates average expression.

4.3.2.3 Annotation with annotated gene expression sets

I also performed annotation of monocyte subsets using publicly available gene expression sets. The *SingleR*[411] package was used to compare the clusters defined within the scRNA-seq data to annotated expression datasets including the Human Primary Cell

Atlas dataset[335], Monaco Immune Data dataset[336], Immune Cell Expression Data[337] as well as data published by Villani et al.[104], one of the key papers which has characterised peripheral monocyte populations using scRNA-seq techniques, albeit across a small cell population. *SingleR* then allowed for assignation of cell type/identity based on similarity of expression profiles within the reference datasets.

Annotation with the Human Primary Cell Atlas dataset categorised subset 7 as a $CD16^+$ monocyte cluster, whilst subsets 0-6 and 8 were predominantly annotated as $CD14^+$/classical monocytes (Figure 26). On comparison to the Monaco Immune dataset, subset 7 was annotated as predominantly an intermediate subset, whilst comparison with the Immune Cell Expression data assigned subset 7 a $CD16^+$ monocyte cluster. Comparison with Villani et al.[104] scRNA-seq expression data annotated Seurat clusters 0-6 and 8 as a Mono1 cluster, which aligns to a classical monocyte group. Seurat cluster 7 most prominently exhibited a Mono2 signature, representative of a non-classical monocyte group.

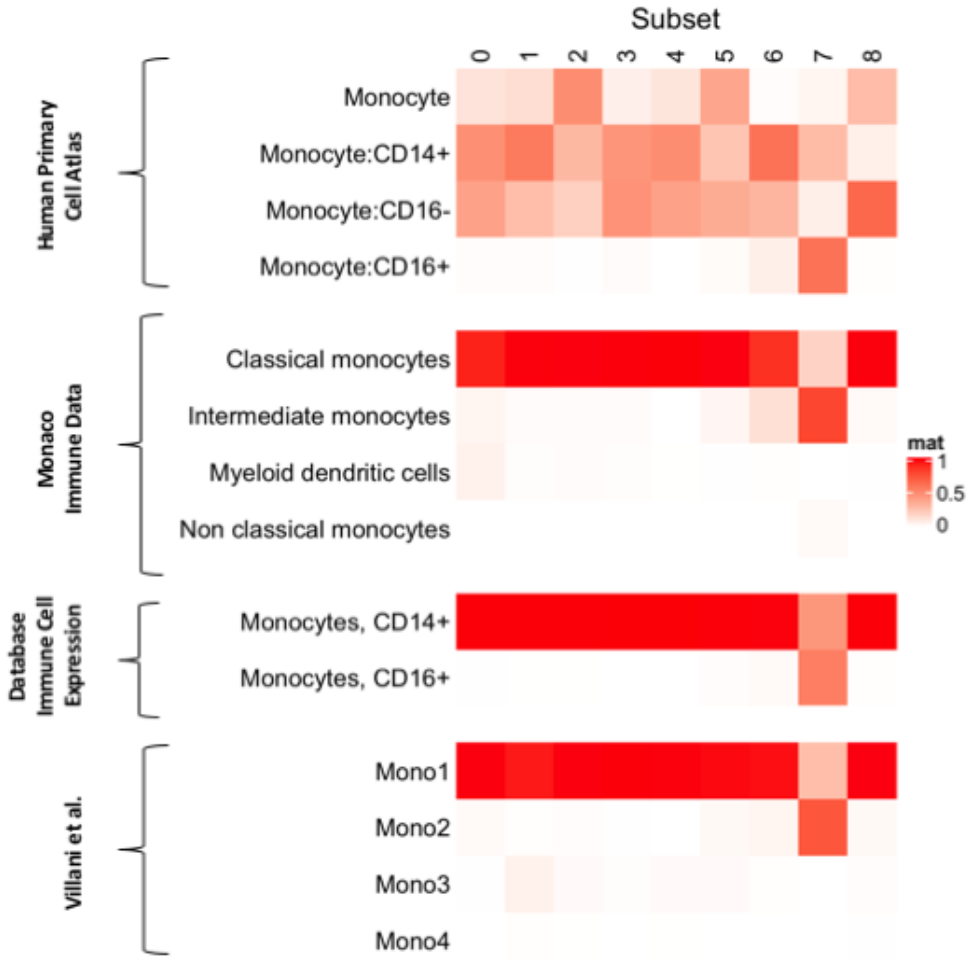


Figure 26: Annotation of scRNA-seq monocyte subset with annotated expression data
Heatmap to show proportion of assigned cell types according to SingleR annotation using expression sets including the Human Primary Cell Atlas, Monaco Immune, Immune Cell Expression and Villani datasets. SingleR was used to compare the expression of each single cell to that of an annotated expression set providing classification annotation. Only cell type annotations comprising at least 1% of the total scRNA-seq population were included. Scale ('mat') indicates proportion of cells in the subset assigned to that cell type with 1 indicating 100%.

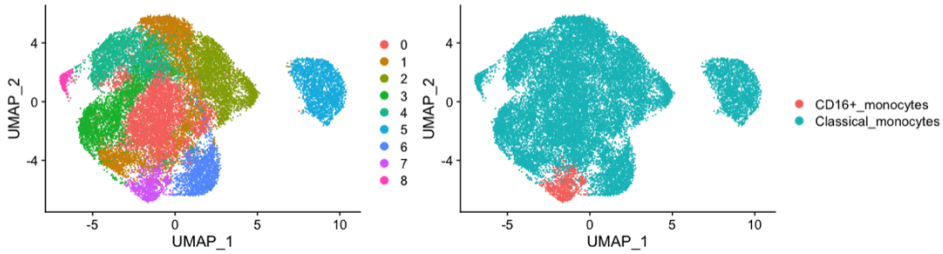
4.3.2.4 Summary of annotation to identify conventional clusters

Annotation approaches including identification of key defining marker genes and comparison with gene expression datasets, suggested that subset 7 was most in keeping with a $CD16^+$ monocyte subset, and represented both a non-classical and/or intermediate monocyte cluster (Figure 27). Seurat monocyte subsets 0-6 and subset 8 were most in keeping with a $CD14^+$/classical cluster. Notably, this included subset 5, which appeared distinct on the UMAP plot (Figure 27A). It should be noted that cluster assignment was performed according to the most prominent annotation signature

inferred by *SingleR* comparison, and with reference to the single key marker gene expression analysis. For example, subset 6 demonstrated a weak intermediate monocyte and $CD16^+$ signature following annotation with Monaco Immune data and Database Immune Cell Expression sets respectively. However, the most prominent signature identified in subset 6 was a classical monocyte signature. There was also considerable variability in the distribution of monocyte Seurat subsets (subsets 0-8) comprising each individual's monocyte compartment (Figure 28), with many samples composed of a single predominant classical subset. For example, the distinct subset 5 is only found in patient ID7 (Figure 28). This suggests that classical subsets may represent differential functional or activation states, rather than robust and distinct cell populations.

Therefore, it was most appropriate to initially merge those Seurat subsets annotated as classical to explore the transcriptional profiles of classical versus $CD16^+$ monocytes, in keeping with conventional categoriation. Subsets most in keeping with classical monocytes (subsets 0-6 and 8) were merged and referred to as a single 'classical monocyte' cluster (n=23600 cells) whilst subset 7 was referred to as a '$CD16^+$ monocyte cluster' (n=857 cells) (Figure 27). However, whilst firstly monocytes at this level of conventional functional categorisation will be explored, the annotation and clustering I present here demonstrates that there is likely to be further transcriptional complexity within the classical compartment. In this chapter I first explore transcriptomic profiles of conventionally-annotated monocyte subsets (classical versus $CD16^+$ monocyte clusters), characterising subset-wise profiles in MM patients and responses to ICB. I then explore in more detail the previously undescribed complexity within the classical and $CD16^+$ monocyte compartments afforded by this large-scale scRNA-seq approach.

A



B

Monocyte cluster	Annotation dataset				Seurat subset	Comment
	Human Primary Cell Atlas	Monaco Immune Data	Immune Cell Expression Data	Villani et al. (2017)		
Classical	Monocyte, Monocyte CD14+, Monocyte CD16-	Classical monocytes	Monocytes, CD14+	Mono1	0,1,2,3, 4,5,6,8	
Intermediate	Monocyte CD16+	Intermediate monocytes	Monocytes, CD16+	Mono3/4	7	No intermediate Villani signature
Non-classical	Monocyte CD16+	Non-classical monocytes	Monocytes, CD16+	Mono2	7	Minimal Monaco Immune Data non-classical signature

C

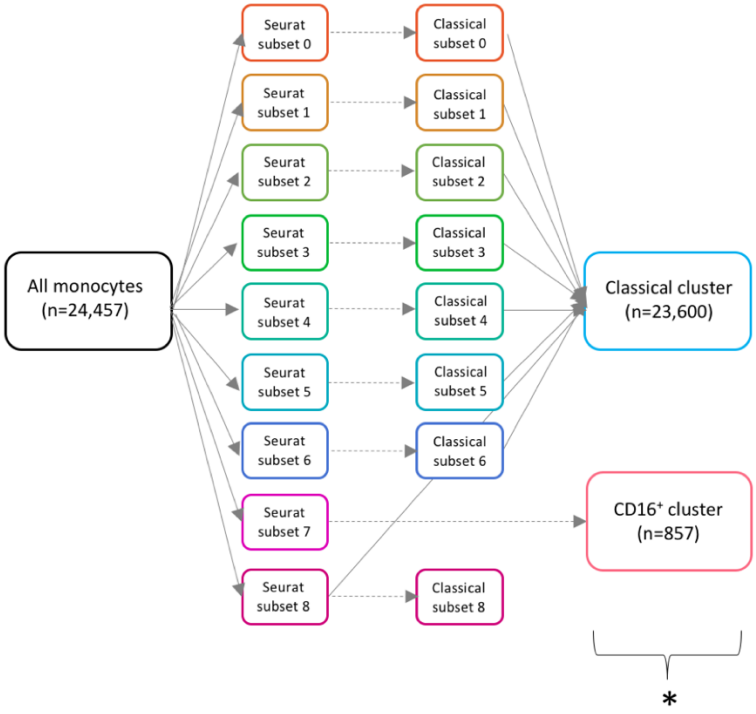


Figure 27: Identifying conventional monocyte clusters in scRNA-seq

A) (Left) UMAP plot to show all monocytes (from eight patients receiving ICB at baseline (C1) and day 21 post-treatment (C2), plus three healthy donor individuals) clustered into nine subsets indicated by different colours (total cells n= 24,457). (Right) UMAP plot to show monocytes coloured according to classical and CD16⁺ cluster annotation assignment. C) Table to show summary of annotation of Seurat monocytes clusters (Seurat clusters 0-8) with SingleR. C) Figure to show monocyte cluster assignment. Solid arrows indicate clustering or merging of subsets, dashed line indicates change of assignment but not cluster size/cell composition. The starred clustering level indicates the merged classical cluster and CD16⁺ cluster.

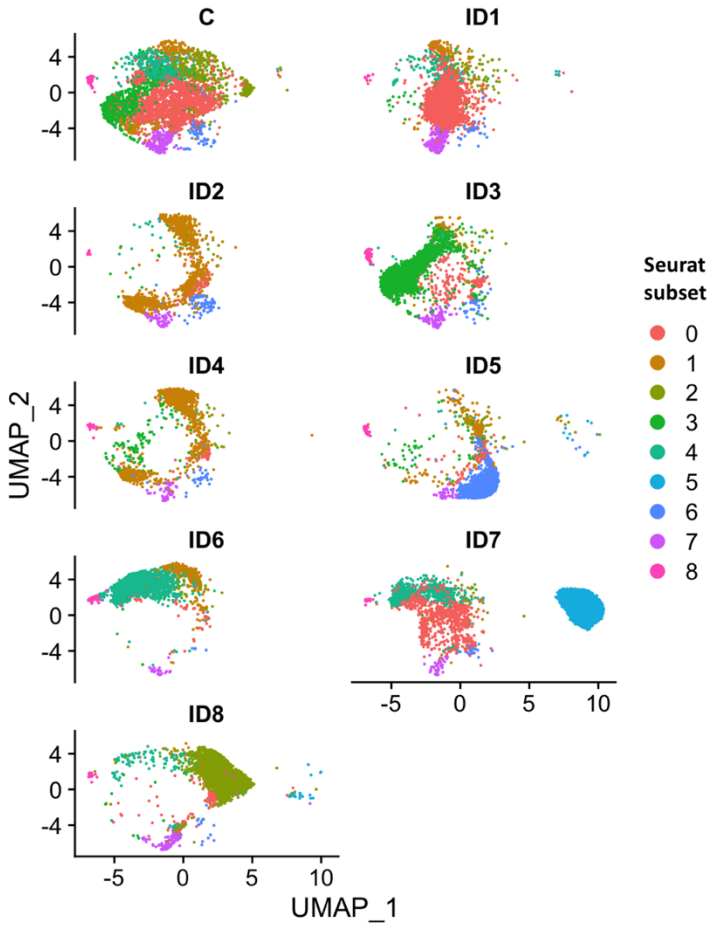


Figure 28: UMAP plot to show Seurat subset assignment by patient ID
UMAP plots to show Seurat subset assignment by patient ID (ID1-8). 'C' denotes control samples (comprises cells from three pooled healthy donor samples). Each point represents a cell and cells are coloured by Seurat subset assignment.

4.3.3 Characterising monocyte subsets

4.3.3.1 Expression profiling peripheral monocytes clusters

Following identification of conventional monocytes clusters: classical (Seurat subsets 0-6 and 8 merged) and $CD16^{+}$ monocytes (Seurat subset 7), I then characterised the transcriptional profiles across these subsets. Cluster-defining gene markers were identified using the Seurat function *FindAllMarkers.*

Classical monocytes were characterised by expression of 593 genes (Wilcoxon signed rank test, $P_{adj}<0.05$). The most significant of these include *S100A8* ($P_{adj}< 1x10^{-300}$), *S100A9* ($P_{adj}< 1x10^{-300}$) and *S100A12* ($P_{adj}< 1x10^{-300}$) and the ubiquitous monocyte marker gene and pattern recognition receptor[113], *CD14* ($P_{adj}= 2.0x10^{-244}$). *S100A8,*

S100A9 and *S100A12* are calcium-binding proteins which have a diverse range of functions including regulating cytoskeletal rearrangement , facilitating cell migration and modulating cytokine secretion[412] (Figure 29). Genes previously found to be highly expressed within classical monocytes and associated with cellular adhesion and migration were definitive of this cluster. These include *CLEC4E* (P_{adj}= 3.2x10^{-63}), which is involved in cell signalling and adhesion[413]; *SELL* (P_{adj}= 3.5x10^{-70}), which facilitates leukocyte-endothelial adhesion[115]; and *CCR2* (P_{adj}= 1.7x10^{-50}), a chemokine receptor shown to be required for migration of monocytes from the bone marrow compartment into the periphery, and to sites of inflammation within tissues[114]. *VCAN* was also highly upregulated within the classical compartment (P_{adj} < 1x10^{-300}). *VCAN* encodes versican, a proteoglycan expressed by stromal cells which facilitates leukocyte adhesion with pro-inflammatory effects[414], and which is also expressed by circulating monocytes, stimulating CCL2 effects[415]. Myeloid *VCAN* expression is associated with an immunosuppressive microenvironment[414]. Thought to facilitate cellular flexibility, the intermediate filament gene vimentin *VIM* was also highly expressed in this cell group (P_{adj}= 3.7x10^{-45}).

The antibacterial lysozyme-encoding gene *LYZ* was highly expressed within the classical compartment (P_{adj}< 1x10^{-300}) as was the anti-oxidant gene *GPX1* (P_{adj}= 4.0x10^{-212}). *GAPDH,* which is involved in glycolysis[416], was also definitive of the classical subset (P_{adj} = 2.7x10^{-147}), keeping with the previously described prominence of carbohydrate-based metabolism in classical monocytes[95]. *MSGT1*, which has a role in protection against oxidative stress [417], was also highly expressed (P_{adj}= 1.3x10^{-99}).

A high degree of transcriptional heterogeneity within the classical compartment was identified (Figure 29A), in keeping with a number of sub-clusters within this group not fully characterised with this high-level approach. This will be explored later in this chapter.

CD16^{+} monocytes were defined by expression of 734 canonical marker genes (Wilcoxon signed-rank test, P_{adj}<0.05), the most significant of which included *FCGR3A* (P_{adj} < 1x10^{-300}), *HLA-DPA1* (P_{adj} = 1.0x10^{-301}), and *IFITM3* (P_{adj}= 2.297251e^{-251}). Key gene markers of non-classical monocytes including *CX3CR1* (P_{adj}= 5.5x10^{-107}), which facilitates monocyte-

endothelial adhesion and promotes the patrolling behaviour of non-classical monocytes[120]; and *CDKN1C* (P_{adj} = 9.7x10^{-167}), which inhibits cellular proliferation via inhibition of cyclin/Cdk[118], were also definitive of this subset. The CD16^{+} monocyte cluster was also defined by cluster-wise expression of *TNF* (P_{adj}= 5.2x10^{-73}), MHC class II genes including *HLA-DRB1* (P_{adj}= 6.2x10^{-151}), *HLA-BPB1* (P_{adj}= 8.5x10^{-233}), and *HLA-DQA1* (P_{adj}= 1.9x10^{-205})). Classical complement genes including *C1QA* (P_{adj}= 9.1x10^{-98}), *C1QB* (P_{adj}= 6.5x10^{-91}), and *C1QC* (P_{adj}= 5.232569x10^{-70}) were highly expressed in CD16^{+} monocytes. Complement component C1q binds to antibody-antigen complexes and activate the classical complement cascade, which ultimately stimulates phagocytosis and cell lysis[418]. C1q also has a role in promoting self-tolerance and limiting auto-immunity[419]. Additionally, interferon-induced genes including *IFITM1* (P_{adj}= 4.3x10^{-137}), *IFIT2* (P_{adj}= 3.4x10^{-12}), and *IFIT3* (P_{adj}= 6.3x10^{-50}) were highly expressed in CD16^{+} monocytes.

To identify associated biological function across these monocyte clusters, pathway enrichment analysis of the GOBP database was performed. This demonstrated distinct biological profiles across monocyte clusters. Classical and CD16^{+} monocytes shared enrichment of 16 GOBP pathways demonstrating shared biological function, however the majority of pathways were selectively enriched (P_{adj}<0.05) (Figure 29). Classical monocytes were selectively enriched for 67 GOBP pathways including phagocytosis (GO:0006909), response to wounding and wound healing (GO:0009611, GO:0042060), cell-matrix adhesion (GO:0007160) and chemotaxis (including neutrophil (GO:0030593), monocyte (GO:0002548, GO:0090026) and dendritic cell chemotaxis (GO:0002407)). In addition, pathways such as chemokine signalling (GO:0070098) and antimicrobial response (GO:0019730), including defence response to Gram-positive and negative bacteria (GO:0050830 and GO:0050829 respectively) were enriched. This indicates that classical monocytes have a predominant role in modulating antibacterial innate immune responses, as well as chemotaxis, cell migration and wound healing. Pathways including platelet activation (GO:0030168) and degranulation (GO:0002576) were selectively enriched, in keeping with facilitation of a pro-thrombotic state.

CD16^{+} monocytes showed selective enrichment of 60 GOBP pathways (P_{adj}<0.05). These included IFN-γ-mediated signalling pathways (GO:0060333, GO:0034341), type I

interferon response (GO:0060337), T cell co-stimulation (GO:0031295) and TCR signalling (GO:0050852), TNF-mediated signalling (GO:0033209) and antigen processing and presentation via MHC-class II (GO:0019886) (Figure 29). Anti-viral response pathways were enriched in this cluster (GO:0051607, GO:0045071). Notably, this cluster was selectively enriched for positive regulation of telomerase activity (GO:0051973) in keeping with a more mature population[106]. Nitric oxide biosynthesis (GO:0045429) was also enriched, in keeping with previous studies which have shown prominent NO production and oxidative phosphorylation pathways in $CD16^+$ monocytes[87]. This demonstrates that $CD16^+$ monocytes have a pro-inflammatory expression profile, which may represent an activated immune phenotype with prominent co-stimulatory effect across both innate and adaptive immune responses.

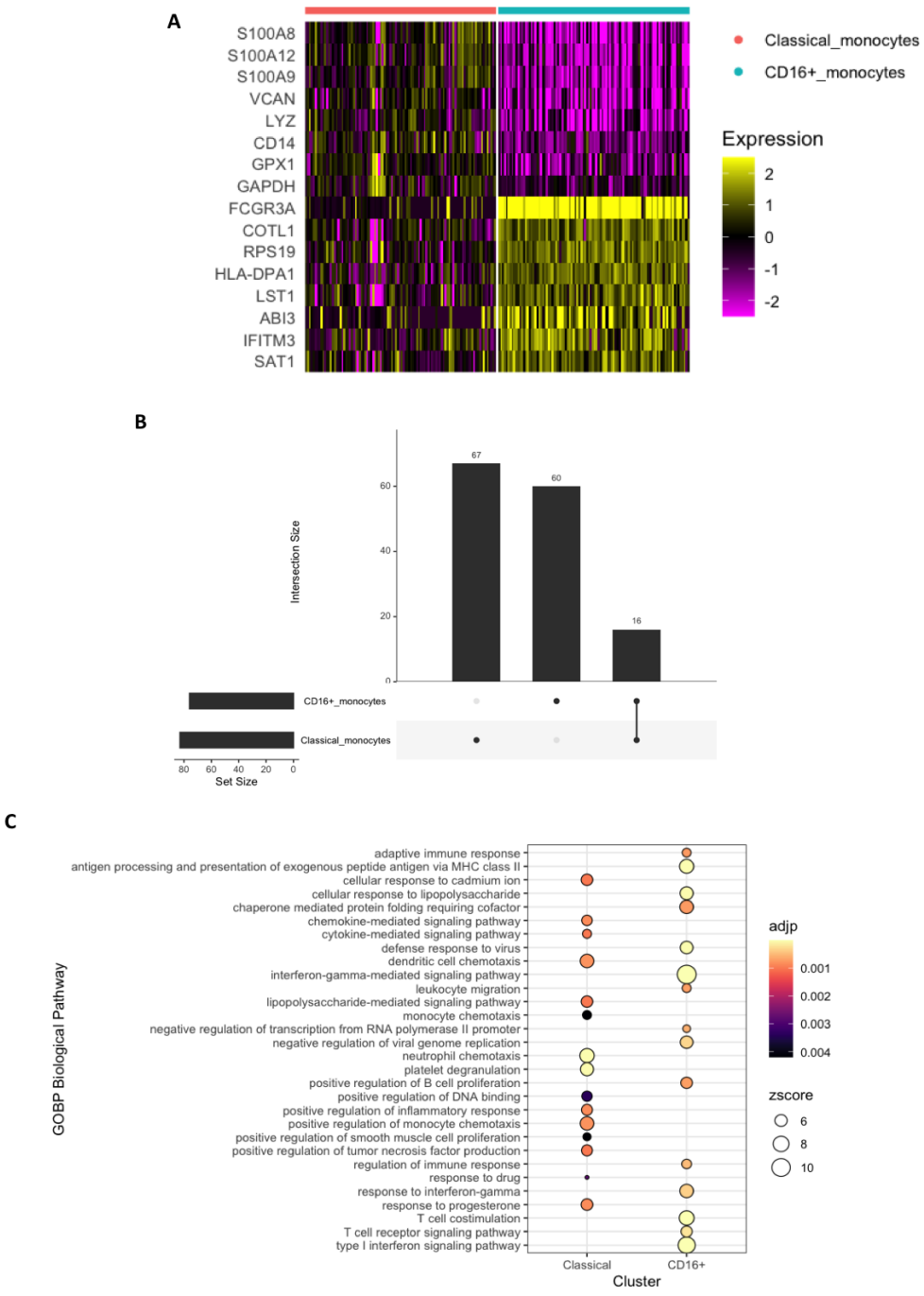


Figure 29: Characterising monocyte clusters

A) Expression heatmap showing gene expression (top 8 most significant marker genes per cluster) for classical and CD16⁺ monocytes clusters, with each cluster sub-sampled to 100 cells. B) Upset plot to show shared and cluster-wise enrichment of GOBP pathways across classical and CD16⁺ monocytes. C) Dotplot to show GOBP pathways most significantly enriched (top 15 most significant pathways, P_{adj}<0.05) in classical and CD16⁺ monocyte clusters.

4.3.3.2 Characterising monocyte clusters with gene expression-correlated module scoring

To further characterise expression profiles across monocyte clusters, expression of gene sets associated with key myeloid functions such as chemotaxis, immaturity, antigen presentation via MHC class II, complement production, macrophage differentiation and response to IFN-γ, were identified (Figure 30). Gene sets were defined by identifying the top 50 genes, the expression of which across the scRNA-seq dataset (all cells n=24,457) were most significantly positively correlated with expression of a candidate 'hub gene' (e.g. *C1QA*). Genes recognised to be definitive of, or particularly highly expressed, in a particular monocyte subset, or associated with biological process or function were chosen as 'Hub genes' (see Figure 30 and Appendix U, Appendix V, Appendix W, Appendix X and Appendix Y for gene sets). The average expression of these gene sets were calculated per cell using the *Seurat* function *AddModuleScore.*

Key hub gene	Expression score name	Associated phenotype/function
CD14	Classical score	Classical monocytes
FCGR3A	Non-classical score	Non-classical monocytes
CCR2	Chemotaxis score	Chemotaxis
CD36	Immaturity score	Monocyte immaturity
HLA-DRA	Antigen-presentation score	MHC class II antigen presentation.
C1QA	Complement score	Complement signalling
STAT1	IFN-γ signalling score	Response to IFN-γ
CD38	Pro-inflammatory score	Pro-inflammatory marker
CSF1R	Macrophage differentiation score	Macrophage differentiation
IL10	Immunosuppression score	Immunosuppression/anti-inflammatory cytokine

Figure 30: Key hub genes and expression score modules
Table to show key genes ('hub genes'), associated phenotype and/or function and the expression score name. Hub genes were selected based on their association with a key myeloid-associated phenotype or biological function.

The immaturity score (derived from the top 50 transcripts across all scRNA-seq cells most significantly correlated with expression key marker of monocyte immaturity *CD36*)

included genes definitive of classical monocytes, including *S100A8, S100A9, S100A12* and *SELL* (see Appendix V). Similarly, the chemotaxis score-correlated genes (hub gene *CCR2*) included transcripts predominantly definitive of a classical phenotype including *S100A9* and *S100A10*, as well as the immaturity marker *CD36* (see Appendix U). Both antigen-presentation and complement scores were also significantly correlated with expression of genes characteristic of non-classical transcriptional profiles (Appendix W).

Expression of phenotype and/or function-associated gene module expression scores in classical and $CD16^+$ monocyte clusters showed higher expression of gene modules associated with antigen-presentation, complement production, IFN-γ signalling and inflammation in $CD16^+$ monocytes, whilst classical monocytes exhibited higher expression of immaturity and chemotaxis-associated gene modules (Figure 31). $CD16^+$ monocytes exhibited higher expression of the macrophage marker scoring (hub gene *CSF1R*) than classical monocytes (Figure 31). *CSF1R* is a marker of macrophage differentiation and is required for migration of macrophages into tissue spaces[152]. The *CSF1R*-associated gene set included predominantly non-classical/intermediate-associated genes, particularly MHC class II antigen presenting genes (Appendix Y). Whilst it is recognised that fetal-derived tissue resident macrophages exist variably according to tissue type[103,145,147,148], macrophages derived from the circulating pool of peripheral monocytes are thought to be derived from classical monocytes[103] with a non-classical monocytes representing an alternate cell fate. However, this suggests that macrophages may be derived from a more mature, $CD16^+$ monocyte population.

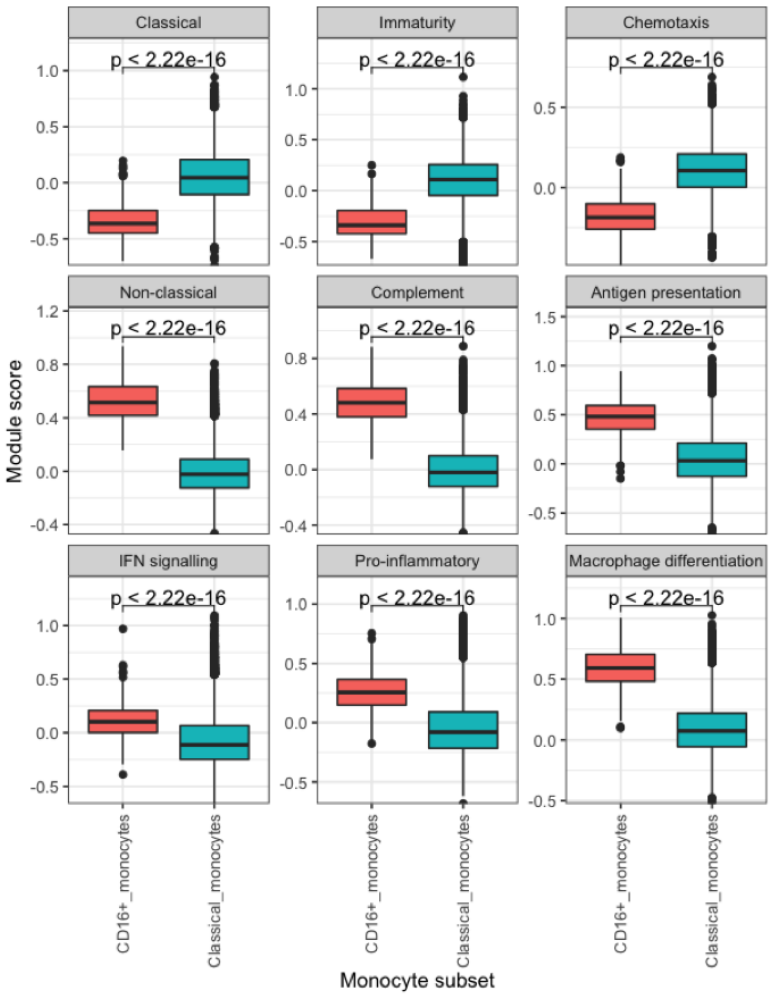


Figure 31: Characterising monocyte clusters
Boxplot to show expression of module expression scores across classical (n=23600 cells) and CD16^{+} clusters (n=857 cells). Module expression scores compared between classical and CD16^{+} monocyte subsets using a Wilcoxon signed-rank test).

4.3.3.3 Cell cycle annotation and activation status of monocyte clusters

To characterise activation states, cells were assigned to cell cycle phases dependent on the expression of cell cycle-associated genes. Genes associated with cell cycle phases as defined by Tirosh et al.[420] available within the *Seurat* package were used for annotation. Expression of these genes across each cell within the monocyte scRNA-seq data was annotated with the *CellCycleScoring* function (Figure 32A). A greater proportion of cells were found to be in S phase and G1 phase in the CD16^{+} monocyte compartment compared to classical monocytes, whilst a greater proportion of cells were within G2M phase in classical monocytes, in keeping with prominent cycling activity and higher cell

turnover (Figure 32A). Expression of proliferation index marker *MKI67* was present almost exclusively in classical monocytes, although expression of this gene was found in a limited number of cells (Figure 32B).

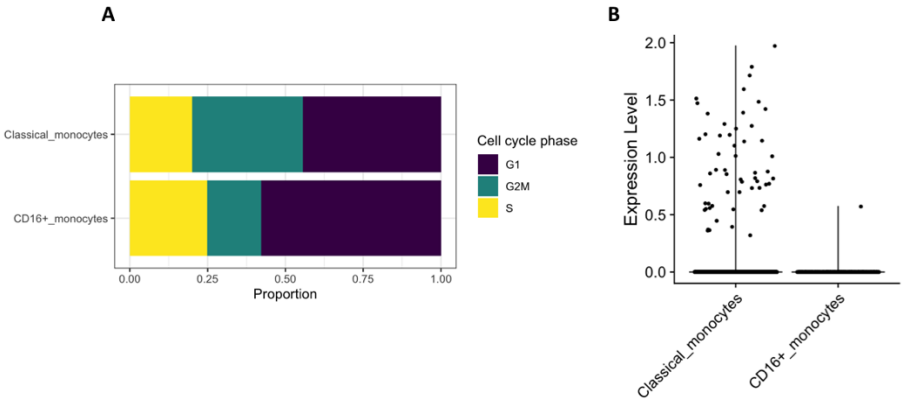


Figure 32: Proportion of monocyte cluster cells assigned to cell cycle phase
A) Proportion of cells in the classical and CD16⁺ subset assigned to cell cycle phase according to the CellCycleScoring function in Seurat. B) Violinplot to show MKI67 expression across classical and CD16⁺ monocyte subsets (n=23600 classical cells, n=857 CD16⁺ cells).

Overall, these findings are broadly in keeping with the previous literature, which describe CD16⁺ non-classical monocytes as a more mature, senescent population, compared to their classical counterparts which have a more prominent cycling activity[91].

4.3.4 Monocyte differentiation and fate explored with trajectory analysis

It is thought that monocytes exist across a linear differentiation continuum, with immature classical monocytes differentiating into intermediate and then non-classical monocytes[149]. However to date, studies exploring monocyte differentiation are limited and have typically utilised radioactive labelling in vivo with subsequent analysis of subsets with flow cytometry. To explore differentiation states in classical versus CD16⁺ monocytes across this scRNA-seq dataset, trajectory analysis was performed using the *Monocle* package[338].

Monocytes from across all conditions were randomly subsampled (n=2000 cells). The pseudotime map demonstrated a population of classical monocytes present at the origin point, with a trajectory towards a termination point enriched with CD16⁺ monocytes (Figure 33). Non-classical gene marker *FCGR3A,* as well as antigen-presenting

gene *HLA-DRA*, showed increased expression with pseudotime (Figure 33C-D). One main bifurcation point was seen, populated predominantly with classical monocytes (branch point 5). Pseudotime is a measure of the stage at which a cell is at across a particular biological process, so a higher pseudotime score infers a greater degree of differentiation[338]. $CD16^{+}$ monocytes exhibited a significantly higher pseudotime score than classical monocytes, in keeping with a more mature phenotype (Figure 33E). To explore association between pseudotime score and changes in expression profiles and function, correlation between pseudotime score and gene module expression scores as described in section 4.3.3.2 were characterised. Pseudotime score was significantly positively correlated with non-classical, antigen-presenting, complement, pro-inflammatory and macrophage differentiation gene module scores and negatively correlated with classical, chemokine, immaturity gene module scores (Figure 34).

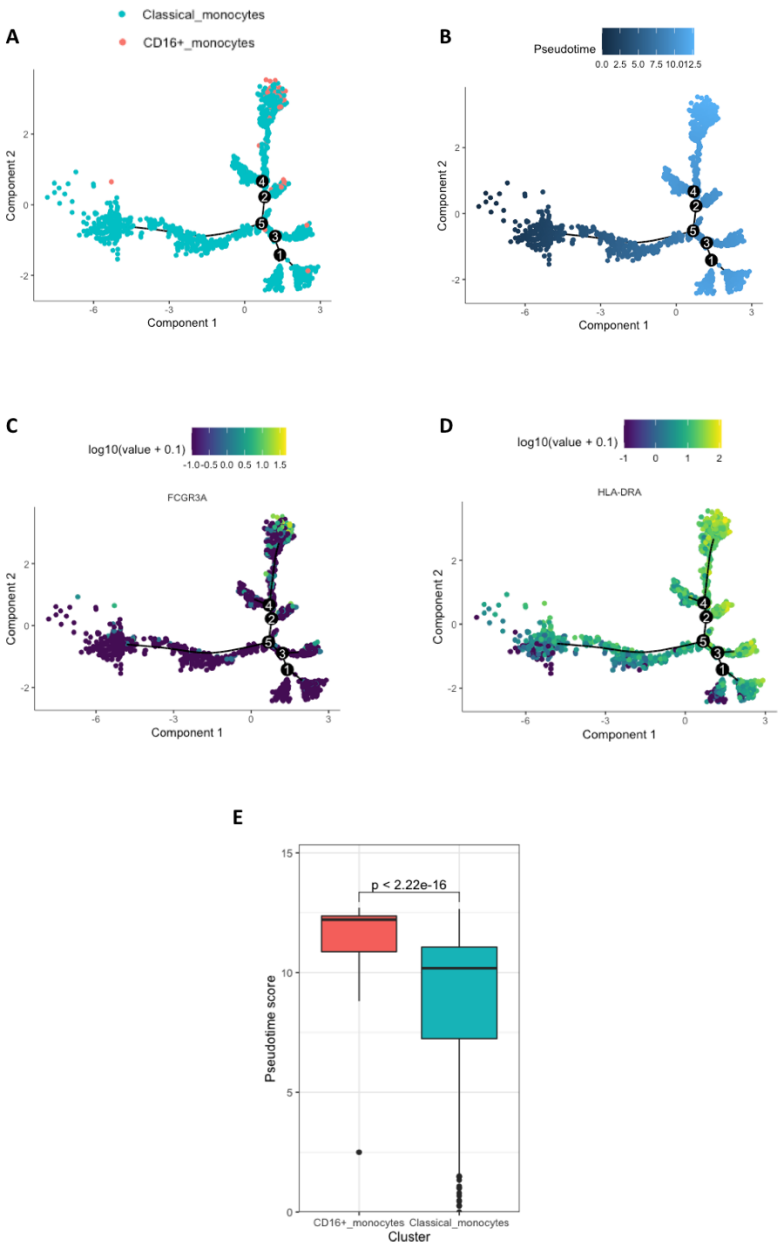


Figure 33: Trajectory analysis of monocyte clusters

A) Pseudotime plot to show pseudotime scores of cells coloured according to classical (1923 cells) and CD16^{+} monocyte (77 cells) clusters assignment. B) Pseudotime trajectory plot to show pseudotime scores. C-D) Expression of (C) FCGR3A and (D) HLA-DRA across the pseudotime trajectory plot. E) Boxplot to show pseudotime score across classical and CD16^{+} monocyte clusters (1923 classical cells, 77 CD16^{+} cells, Wilcoxon signed-rank test).

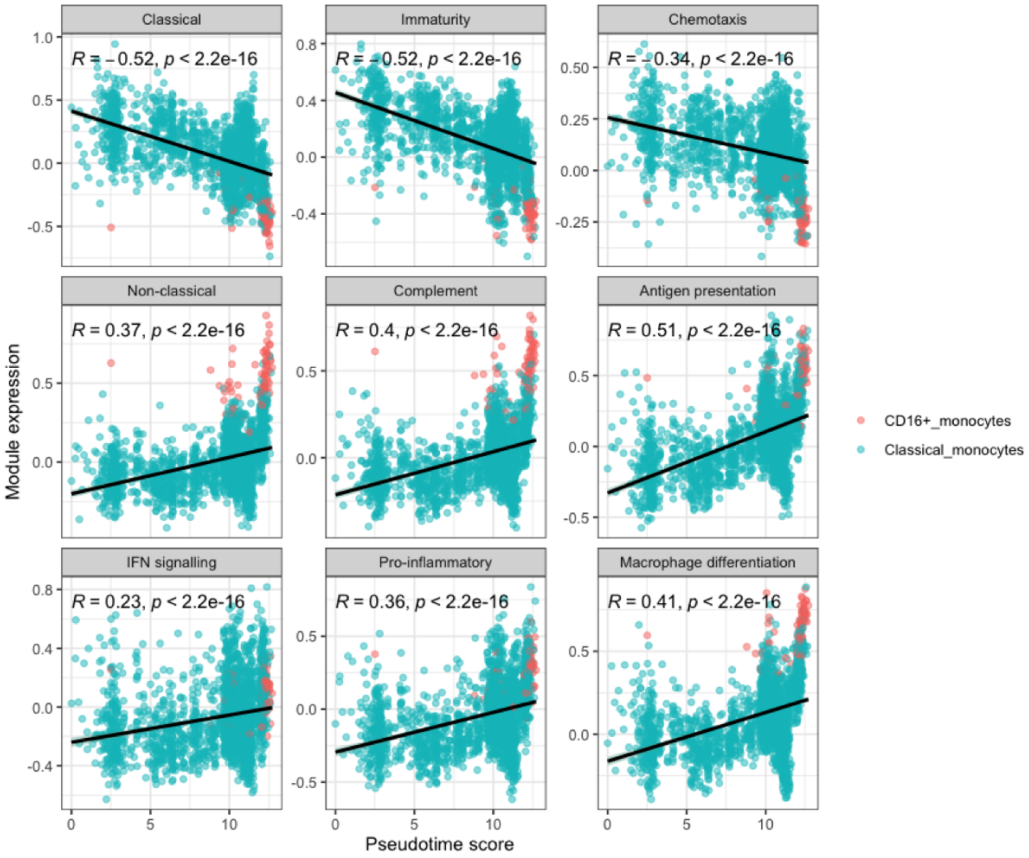


Figure 34: Correlation between pseudotime and gene module score expression
Pearson's correlation analysis between pseudotime scores and gene module expression score (denoted by key hub gene) across 2000 monocytes across controls (n=328 cells), C1 (n=753 cells) and C2 (n=919 cells). Cells are coloured according to classical (blue) or CD16^{+} (red) assignment.

Genes whose expression changed as a function of pseudotime were then identified using the Monocle function *differentialGeneTest* (p-value <0.05). Genes associated with classical monocytes transcriptomic profiles, including *CD14, SELL, FOS, VCAN*, *S100A8, S100A9* and *S100A12,* showed reduced expression with pseudotime (Figure 35). Expression of the definitive marker of non-classical phenotype, *FCGR3A,* increased as a function of pseudotime. Numerous ribosomal genes showed in increased expression with pseudotime, with highest expression just prior to the terminal point. Similarly, the expression of numerous MHC class II antigen presenting genes were modulated as a function of pseudotime, with increasing expression with increasing pseudotime. These observations are in keeping with a classical to non-classical monocyte phenotype differentiation continuum, with an activated population representing a mature but non-terminally differentiated cell group.

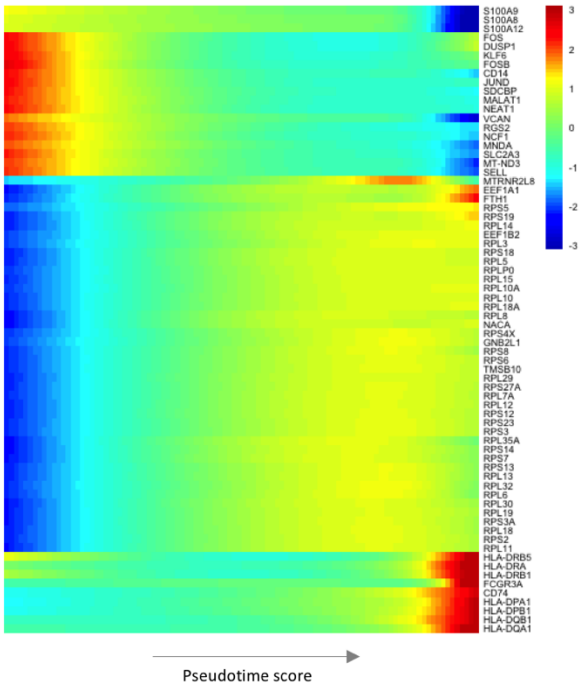


Figure 35: Genes expression modulation as a function of pseudotime
Relative expression heatmap to show to top 70 genes that change as a function of pseudotime (left side of heatmap is origin point, right side is termination point of pseudotime trajectory).

4.3.5 Dissecting monocyte subsets in MM patients versus healthy controls

To characterise monocyte cluster-dependent profiles in melanoma patients versus healthy controls, cluster-wise differential expression analysis was performed using the Seurat function *FindMarkers*. Cells were subsampled to control for subset size (n = 217 cells as this was the size of the smallest condition-wise cluster). Across classical monocytes a total of 363 genes (Wilcoxon signed-rank test, $P_{adj}<0.05$) were significantly differentially expressed in patients versus healthy controls. This included 236 genes induced in MM patients, and 127 genes suppressed (Figure 36). In $CD16^+$ monocytes there were 570 significantly differentially expressed transcripts in $CD16^+$ monocytes between patients versus healthy controls ($P_{adj}<0.05$), including 305 induced in patients, and 265 downregulated (Figure 36). It should be noted that the fold change of expression was low.

Some transcripts showed shared modulatory behaviours across both monocyte clusters, with 72 transcripts induced in MM in both clusters (Figure 36B). Both classical and $CD16^+$ monocytes showed upregulation of the pro-inflammatory cytokine *IL1B*, as well as

chemokines including *CCL4* and *CCL5.* Other genes induced in both clusters included *FCGR3B*. Transcripts associated with interferon signalling and cytotoxicity were also found to be induced across clusters in MM patients, including *JAK3*, which is a type I tyrosine kinase receptor, *GNLY* and *KLRF1.* Both classical and $CD16^+$ monocytes exhibited increased expression of numerous ribosomal genes.

The majority of transcripts in patient with MM versus healthy controls showed cluster-wise modulation (Figure 36). Classical monocytes showed selective induction of complement gene *C1QB* (P_{adj}= 0.0099) and *C2* (P_{adj}= $4.4x10^{-7}$) and pro-inflammatory chemokine *CXCL11* (P_{adj}= $2.8x10^{-10}$). Some transcripts associated with IFN signalling were induced, such as *JAK2,* a type II IFN receptor (P_{adj}= 0.04). The costimulatory receptor *CD40* was also selectively induced in classical monocytes (P_{adj}= 0.0055).

Genes selectively modulated in $CD16^+$ monocytes included the transcription factor *STAT4* (P_{adj}= $5.2x10^{-14}$), the pro-inflammatory chemokine receptor *IL18R1* (p_{adj}= 0.00017) and *CD38* (P_{adj}= 0.037). The pro-inflammatory cytokine *CXCL10* was also selectively induced in $CD16^+$ monocytes (P_{adj}= $5.0x10^{-17}$). Interferon-associated transcripts *IFIT2* (P_{adj}= 0.0058) and *IFIT3* (P_{adj}= $5.9x10^{-6}$) were induced selectively, as well as some transcripts associated with mitogenic signalling such as *MAP3K7CL* (P_{adj}= $5.1x10^{-13}$) and protein kinase C gene *PRKCH* (P_{adj}= 0.0037).

It should be noted that whilst the overlapping gene set was not large (Figure 36), modulation across both subsets showed similar functions and processes, with induction of transcripts associated with pro-inflammatory and chemotaxis-associated activity.

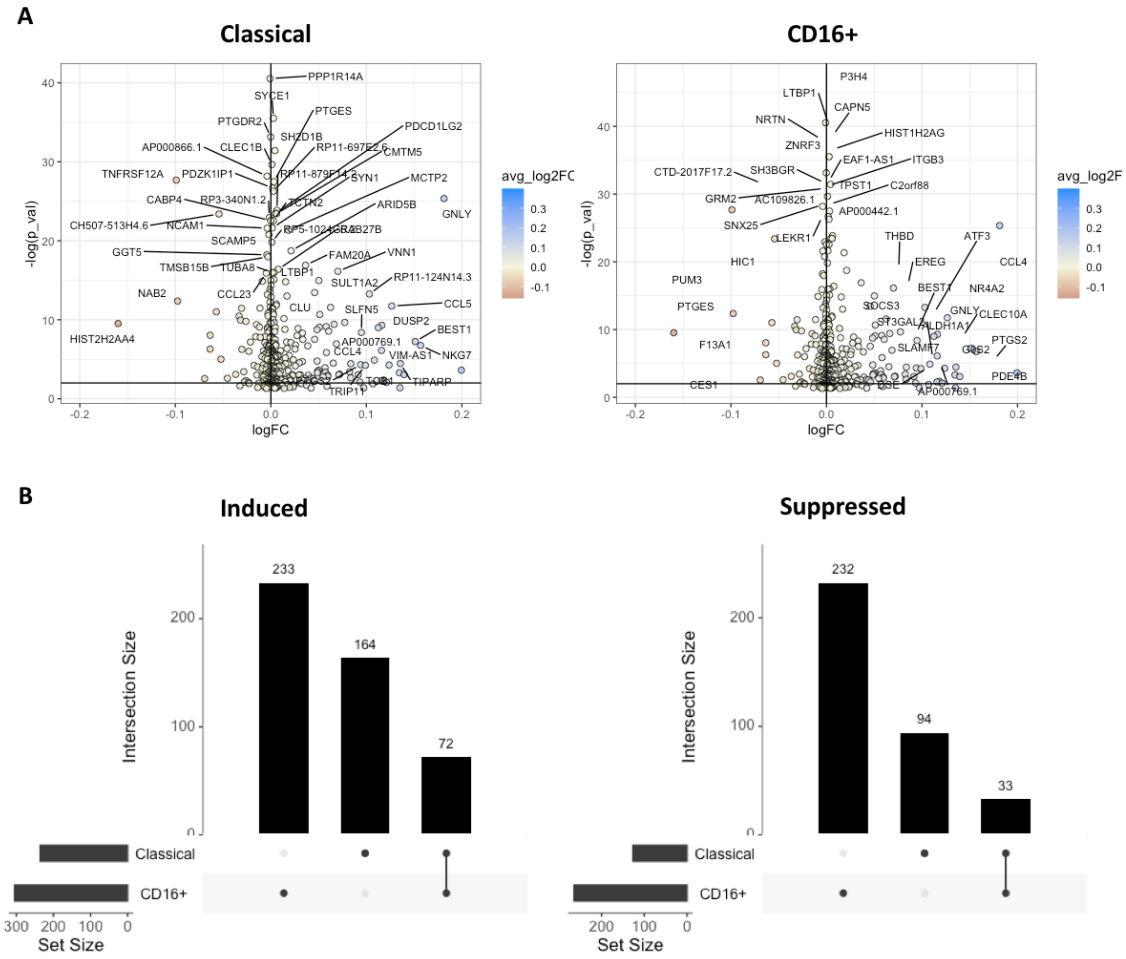


Figure 36: Cluster-wise monocyte differential expression in melanoma patients versus healthy controls

A) Volcano plots to show significantly differentially expressed genes in patients versus healthy controls in classical monocytes and CD16$^+$ monocytes (P_{adj} <0.05). B) Upset plots to show overlap of differentially expressed induced or suppressed transcripts in classical or CD16$^+$ monocytes in patients with MM (8 patients) versus healthy controls (three healthy donors).

There were notably many similarities in the pathways modulated in classical and CD16$^+$ monocytes in MM patients. Pathways enriched in patients across both monocyte clusters included those involved in response to virus (GO:0009615) and response to LPS (GO:0032496). Notably, pathways associated with leukocyte migration were induced in MM across both subsets, (leukocyte migration GO:0050900 in both subsets; positive regulation of cell migration in classical monocytes GO:0030335). Pathways associated with chemokine signalling or chemotaxis were induced in both custers (classical monocytes showing induction of the chemotaxis pathway (GO:0006935), and CD16$^+$

monocytes induction of the chemokine signalling pathway (GO:0070098)). Pathways including platelet degranulation (GO:0002576 and GO:0030158) and neutrophil chemotaxis (GO:0030593) were also enriched in both monocyte clusters in MM (Figure 37A).

There was cluster-wise enrichment of some GOBP pathways (Figure 37B). Classical monocytes showed selective enrichment of pathways involved in cell adhesion (GO:0007155), protein kinase B signalling (GO:0043491), signal transduction (GO:0007165), and protein phosphorylation (GO:0001934). $CD16^{+}$ monocytes showed selective induction of extracellular matrix organisation (GO:0030198) and wound healing (GO:0042060), positive regulation of calcium ion transport (GO:0051928) and GTPase (Figure 37B).

There was concordance between MM-associated transcriptomic profiles in the bulk RNA-seq data described in Chapter Three, and those described here in this scRNA-seq dataset. Both datasets demonstrate induction of chemotaxis, cell migration, response to LPS, and activation of MAPK signatures in monocytes in MM. Induction of platelet activation was seen in both, consistent with induction of a pro-thrombotic state in MM. However, whilst most of these pathways were modulated across both classical and $CD16^{+}$ monocytes, only classical monocytes demonstrated prominent cell adhesion and chemotaxis pathways, whilst $CD16^{+}$ monocytes demonstrate induction of extracellular matrix and wound healing pathways (Figure 37B). There were some differences between MM-associated monocyte transcriptional profiles in bulk RNA-seq and scRNA-seq datasets. In particular, enrichment of viral-response pathways across both monocyte subsets were identified within the scRNA-seq dataset, however this signature was not identified within the bulk RNA-seq dataset in chapter three.

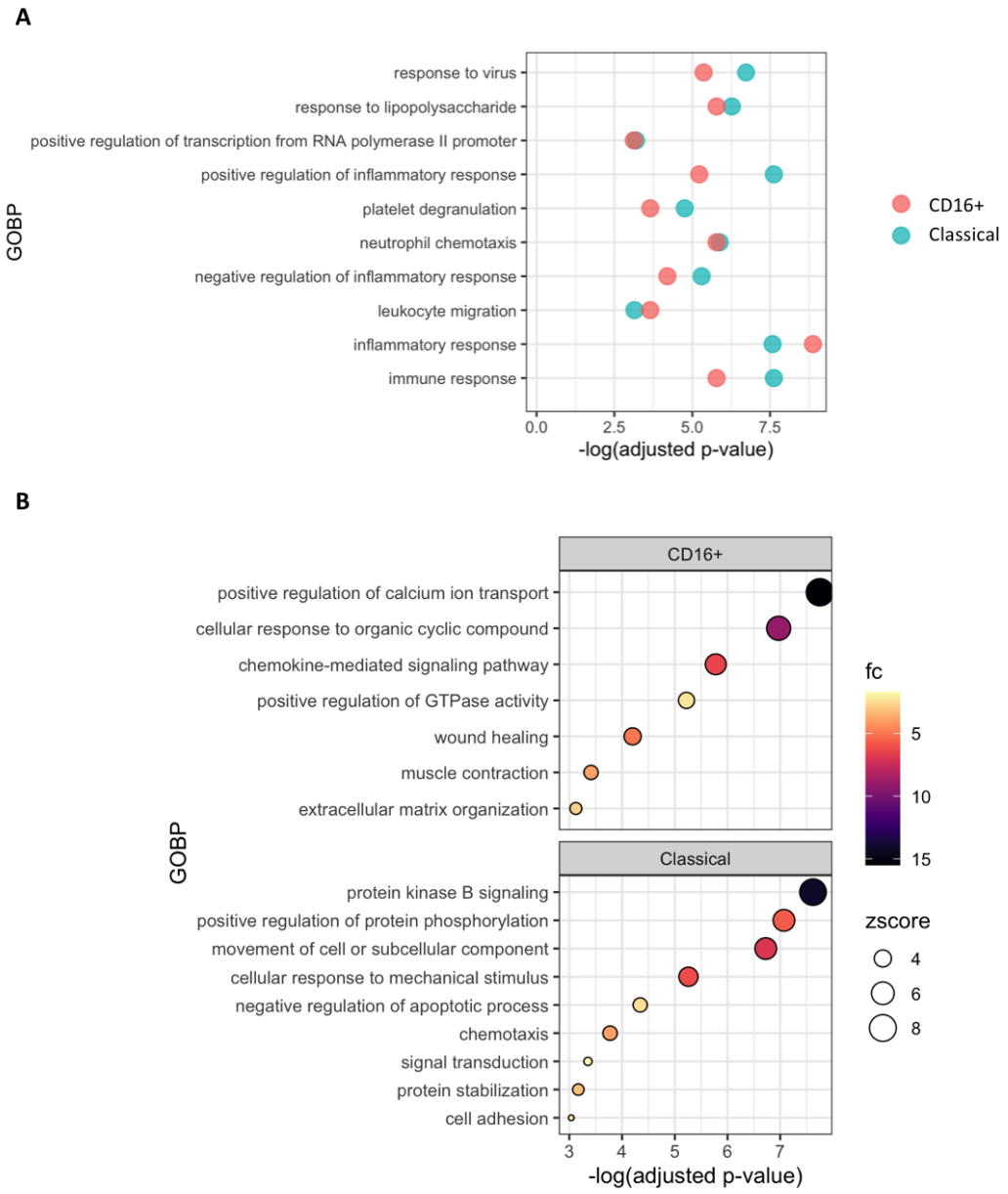


Figure 37: GOBP pathways enriched in melanoma patients
A) Dotplot to show all GOBP pathways that are upregulated in MM patients versus healthy donors (P_{adj}<0.05). Pathway enrichment analysis was performed using XGR interrogation of the GOBP database across all transcripts induced in classical and CD16^{+} monocytes in MM patient versus healthy donors. B) Key GOBP pathways with cluster-wise enrichment in MM patients versus healthy donors (P_{adj} <0.05).

4.3.5.1 Module scoring characterises MM-associated peripheral monocyte profiles

Monocyte cluster-wise MM-associated profiles were characterised further with gene expression module scoring as previously described in section 4.3.3.2. Classical

monocytes exhibited higher immaturity, chemotaxis, classical and IFN-signalling scores in MM patients pre-treatment compared to healthy controls (Figure 38A). $CD16^+$ monocytes exhibited lower immaturity score and increased antigen presentation, complement and macrophage differentiation expression scores in untreated (C1) MM patients versus healthy donors (Figure 38B).

This suggests that the polarisation towards an immature myeloid population and phenotype previously observed in cancer[184] may be due to modulation within classical monocytes. This also suggests that MM-associated profiles may be associated with variation in monocyte differentiation/maturation.

4.3.5.2 Trajectory analysis demonstrates differences in maturation states in MM and healthy controls

To explore variation in monocyte differentiation states in patients with MM, the trajectory analysis performed as described in section 4.3.4 was explored across conditions. Classical monocytes exhibited a lower pseudotime score in patients at baseline (C1) versus healthy controls (Wilcoxon signed-rank test, $P_{adj}<0.05$), in keeping with a more immature phenotype (Figure 39). This was consistent with the reduction in immaturity module scoring observed in in section 4.3.5.1. Conversely, $CD16^+$ monocytes were found to have significantly higher pseudotime score in MM patients versus healthy controls, in keeping with a more mature phenotype (Figure 39). This is in keeping with the more mature expression profile identified with gene expression module scoring in section 4.3.5.1.

This suggests that whilst MM is associated with induction of a phenotype characterised by chemotaxis and cell migration across the monocyte compartment, there is variable shift in differentiation within the heterogeneous monocyte population.

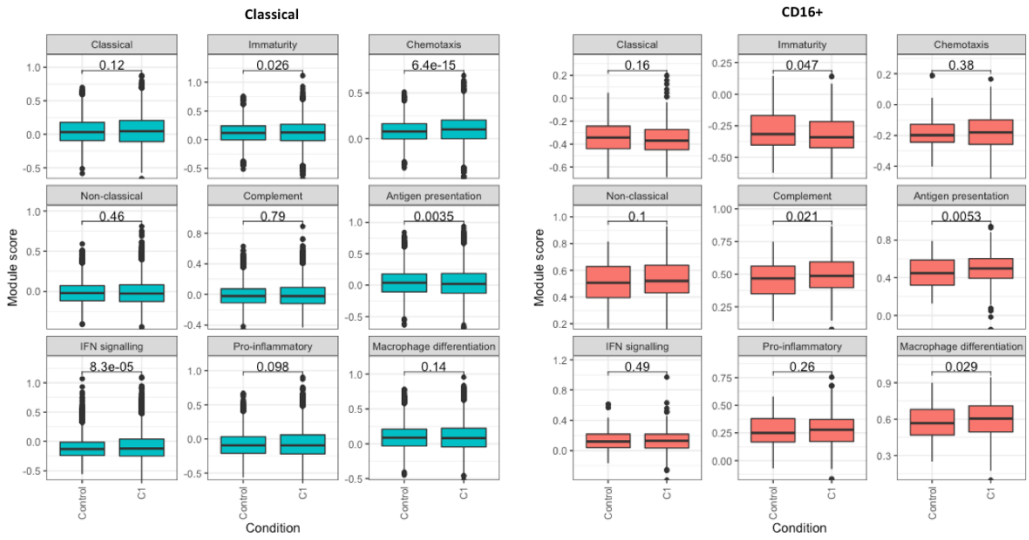


Figure 38: Module scoring by monocyte cluster in melanoma patients and healthy controls
Gene module expression scores in classical and CD16^{+} monocyte clusters in healthy controls versus MM patients (Classical monocytes: C1 n=8979, Control n=3930 cells; CD16^{+} monocytes: C1=342, Control nn=217) (Wilcoxon signed-rank test).

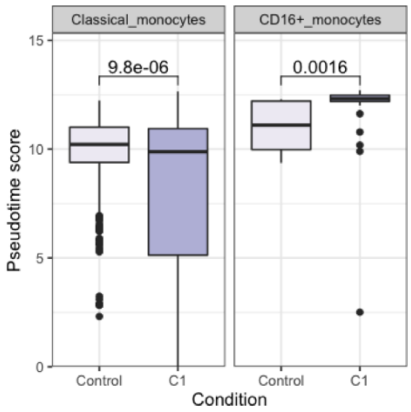


Figure 39: Trajectory analysis of monocyte clusters in MM and healthy controls
Boxplot to show pseudotime score of monocytes in Control and C1 samples (1929 classical monocytes, 71 CD16^{+} monocytes, Wilcoxon signed-rank test).

4.3.6 Dissecting differential monocyte subset responses to ICB

4.3.6.1 Differential gene expression in monocyte subsets with ICB

To explore monocyte cluster-wise on-treatment modulation of transcriptomic profiles, differential expression analysis was performed using the Seurat function *FindMarkers.* Random subsampling of classical monocytes was performed to select 500 cells. Within the CD16^{+} compartment, the differential expression analysis was run across 298 cells

(the size of the pre-treatment population). This however means that magnitude of response cannot be assessed by comparing size of the altered gene set across the two subsets.

Overall there was prominent downregulation of gene expression with ICB across both monocyte clusters. In classical monocytes, a total of 967 transcripts were significantly differentially expressed with ICB (Wilxocon signed-rank test, $P_{adj}<0.05$), with 111 genes induced, and 856 genes suppressed (Figure 40). 781 genes were modulated in $CD16^{+}$ monocytes with ICB, including 288 induced, and 494 downregulated transcripts ($P_{adj}<0.05$).

There were 296 transcripts downregulated in both classical and $CD16^{+}$ monocytes following ICB ($P_{adj} < 0.05$). These include immunosuppressive cytokine *IL10*, ribosomal transcripts, as well pro-inflammatory cytokine *CD38.* Transcripts with shared induction across monocyte compartments included IFN-induced transcript *IFI27,* complement gene *C2* and transcripts associated with a regulatory function such as *CBX2*, which has a role in chromatin remodelling and transcriptional repression; and *MCM2,* which is involved in DNA replication ($P_{adj}<0.05$). Transcripts showing classical-wise modulation include non-classical marker gene *FCGR3A* ($P_{adj}= 0.00029$), *IFITM1* ($P_{adj} = 8.6x10^{-12}$), as well as complement transcripts *C1QA* ($P_{adj} = 2.6x10^{-27}$)*, C1QB* ($P_{adj} = 1.7x10^{-16}$)-and *C1QC* ($P_{adj} = 3.9x10^{-5}$), all of which were induced. Selectively downregulated transcripts (n=560 genes) in classical monocytes included those typically associated with coagulation and platelet aggregation such as *PPBP* ($P_{adj} = 1.3x10^{-5}$) and *GP9* ($P_{adj}= 5.3x10^{-8}$). Several cyclin-dependent kinases were downregulated such as *CDK6* ($P_{adj}= 5.2x10^{-7}$) and *CDKL5* ($P_{adj}= 2.6x10^{-8}$). Transcripts selectively induced in $CD16^{+}$ monocytes include chemokines *CXCL3* ($P_{adj} = 1.6x10^{-41}$) and *CXCL10* ($P_{adj}= 9.0x10^{-22}$), numerous ribosomal transcripts, *IGF1R* ($P_{adj}= 0.00017$) and *IGFBP2* ($P_{adj} = 5.5x10-18$), as well as chemokines *CCL3* ($P_{adj} = 0.0053$) and *CCL28* ($P_{adj} = 1.2x10^{-31}$).

Pathway enrichment analysis of induced and suppressed gene sets was performed using *XGR* to explore modulation of biological function with ICB. Shared responses were observed across both monocyte clusters (Figure 41). There was downregulation of pathways including response to lipopolysaccharide (GO:0032496) and negative

regulation of T cell proliferation (GO:0042130) across both clusters. Cluster-wise responses were more prominent. In classical monocytes there was induction of classical complement activation (GO:0006958), response to hypoxia (GO:001666) and ageing (GO0007568) pathways. Pathways associated with cell motility and migration and adhesion (GO:0007155), were selectively downregulated with ICB (GO:0048870, GO:0030334, GO:2000145). Pathways including, platelet activation (GO:0030168) and degranulation (GO:0002576), and phagocytosis (GO:0050766) were also downregulated (Figure 41). In $CD16^+$ monocytes, several pathways associated with mitogenic signalling such as the MAPK cascade (GO:0000165), ERBB2 signalling (GO:0038128), PI3K activity (GO:0051897) and protein kinase B (GO:0051897) pathways were downregulated. Induced pathways in $CD16^+$ monocytes included chemotaxis (GO:0006935), chemokine signalling (GO:0070098), the innate immune response (GO:0045087) and neutrophil degranulation (GO:0043312). The response to virus pathways (GO:0009615) was both induced and suppressed within the $CD16^+$ monocyte compartment.

Overall, these findings demonstrate distinct modulation of transcriptional profiles in classical and $CD16^+$ monocyte groups with ICB. Classical monocytes broadly showed downregulation of pathways typically associated with a classical phenotype such as cell motility, cell migration pathways, cytokine signalling, and phagocytosis, with induction of transcripts associated with a non-classical phenotype such as *FCGR3A* and complement genes. Enrichment of the ageing GOBP pathway in classical monocytes is in keeping with this transition to a more mature non-classical transcriptional phenotype. $CD16^+$ monocytes demonstrated acquisition of classical-associated transcriptomic profiles, such as chemotaxis and chemokine pathways. This group also showed prominent downregulation of mitogenic signalling pathways such as MAPK signalling, PI3K and ERBB2 signalling. This is particularly interesting given the association with immature cycling myeloid cells and poor clinical outcomes in the context of infections such as COVID-19[421]. This shows that ICB is associated with dynamic changes in maturation and differentiation across monocyte subsets.

Common patterns of response to ICB were seen within this scRNA-seq dataset and those within the bulk RNA-seq data presented in Chapter Three. These include induction of complement, with this response limited to classical monocytes. Downregulation of

pathways associated with cancer-associated transcriptional myeloid phenotype such as platelet activation and degranulation were noted across both scRNA-seq and bulk RNA-seq datasets, and found to be a classical-wise response. Whilst the induction of IFN-γ pathways was not seen within the cluster-wise differentiation expression analysis, downregulation of the negative regulation of T cell proliferation pathway was enriched across both classical and CD16^{+} monocytes, in keeping with T cell activation response.

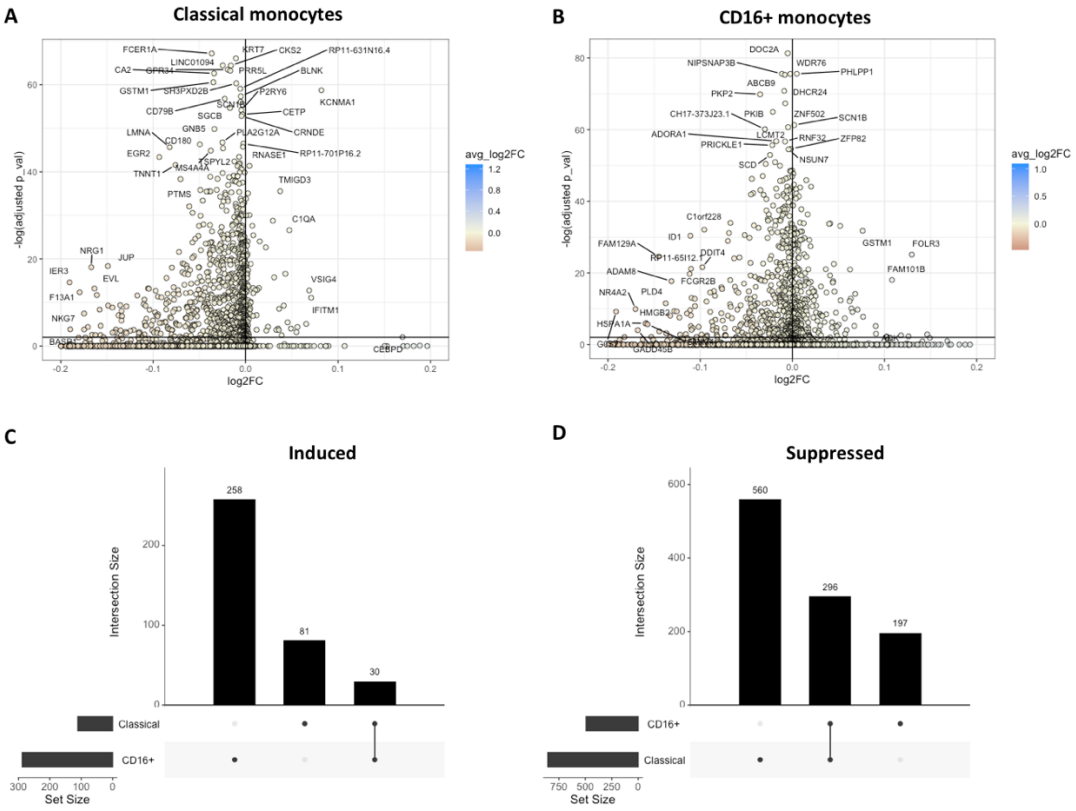


Figure 40: Modulation of transcriptomic profiles in monocyte clusters with ICB
Volcano plots to show the significantly differentially regulated genes with ICB by (A) classical and (B) CD16^{+} monocyte clusters. C) Upset plots to show intersection between induced trancripts in classical versus CD16^{+} monocytes with ICB. D) Upset plots to show intersection in suppressed trancripts in classical versus CD16^{+} monocytes with ICB.

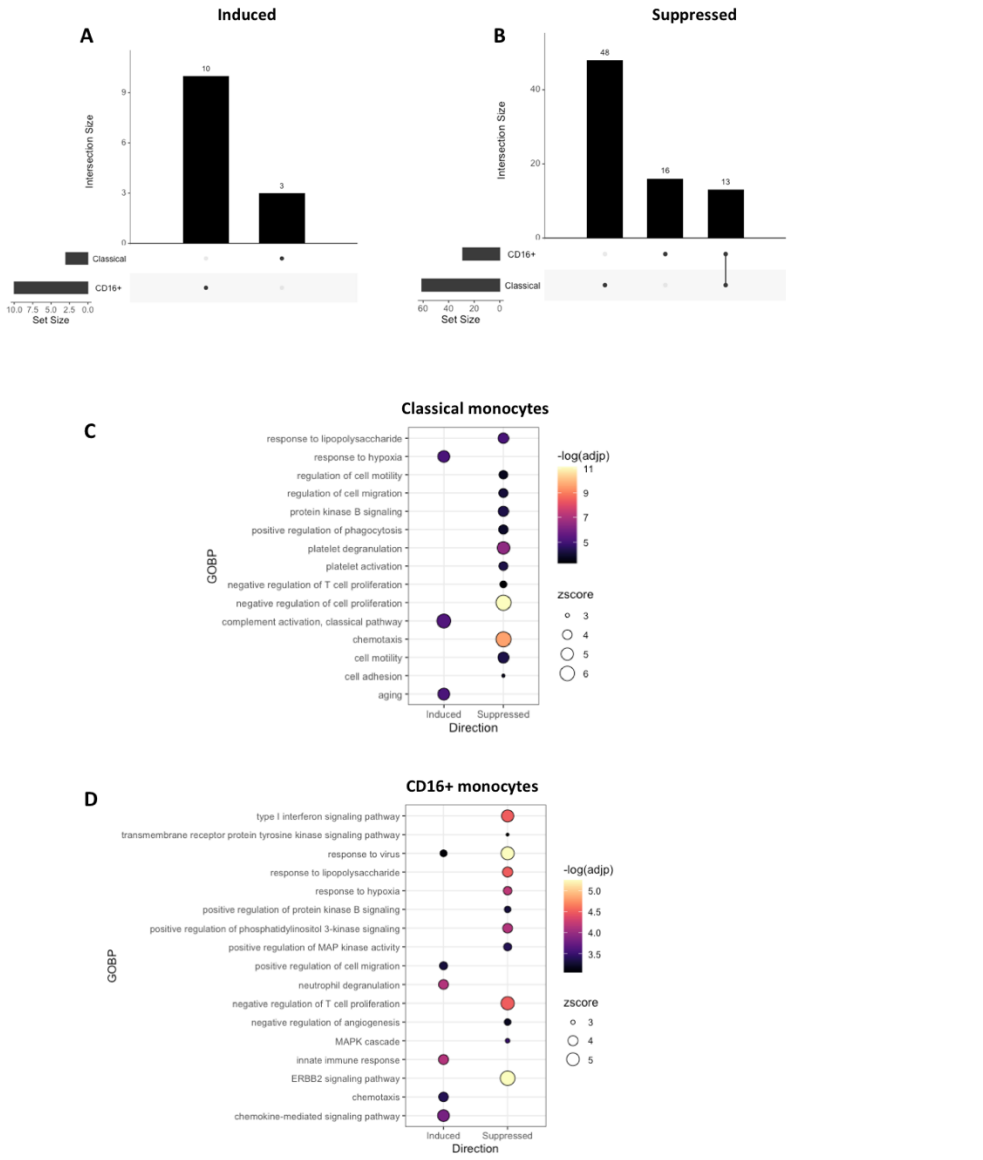


Figure 41: GOBP pathway enrichment in monocytes with ICB treatment
A) Upsetplot to show intersection in enriched GOBP pathways between classical and CD16^{+} monocytes across enriched transcripts following ICB. B) Upsetplot to show intersecrion between enirched GOBP pathways between classical and CD16 monocytes across suppressed transcripts with ICB. C) Dotplots to show key induced and suppressed GOBP pathways in classical monocytes with ICB (P_{adj}<0.05). D) Dotplots to show key induced and suppressed GOBP pathways in CD16^{+} monocytes with ICB (P_{adj}<0.05).

4.3.6.2 Module expression scoring characterises ICB responses in monocyte subsets

To further explore changes in gene expression profiles, function and activation status in monocyte clusters following ICB, modulation of phenotype and activation scores as previously defined in section 4.3.3.2 were explored across treatment states (Figure 42).

Classical monocytes showed reduction in immaturity module expression score with ICB (Figure 42) whilst phenotype/function scores associated with a non-classical phenotype including antigen presenting and complement scores were induced (Figure 42A). The non-classical gene module expression score was also significantly increased in classical monocytes with ICB (Figure 42). In $CD16^+$ monocytes there was no significant difference in complement, antigen presentation, and non-classical score (Figure 42). Interestingly, the IFN-signalling score was reduced with ICB in $CD16^+$ monocytes, but significantly increased in classical monocytes.

Overall, this suggests that ICB therapy induces maturation within the peripheral classical monocyte compartment, with polarisation towards a $CD16^+$ monocyte/non-classical phenotype with associated acquisition of complement and antigen-presenting-type expression profiles. These observations suggest that the induction of IFN signalling pathways observed within the bulk RNA-seq data in Chapter Three may in part be due to shift of the monocyte compartment towards a non-classical, pro-inflammatory phenotype.

Myeloid populations are associated with immunosuppressive effects in cancer. An expression module score was calculated according to expression of the 50 genes most significantly associated with immunosuppressive cytokine *IL10* expression using the same method as described in section 4.3.3.2 (See Appendix Y for gene set). The immunosuppressive module score was reduced across both classical and $CD16^+$ monocytes following ICB. This suggests that ICB may have effects via modulating immunosuppressive myeloid signatures. This may be attributable to monocyte maturation towards a T cell costimulatory, pro-inflammatory phenotype, or may be due to effects on a M-MDSC population. This will be explored later in this chapter.

Interestingly, the macrophage differentiation module expression score was reduced in $CD16^+$ monocytes following ICB, but not in classical monocytes. Tissue resident macrophages are in part derived from differentiation from a peripheral monocyte reservoir[75]. This reduction in macrophage module expression might reflect a migratory shift of these macrophage-destined myeloid cells from the peripheral immune

compartment into tissues, or polarisation to a more immature phenotype following ICB in conjunction with a right-shift in the classical compartment.

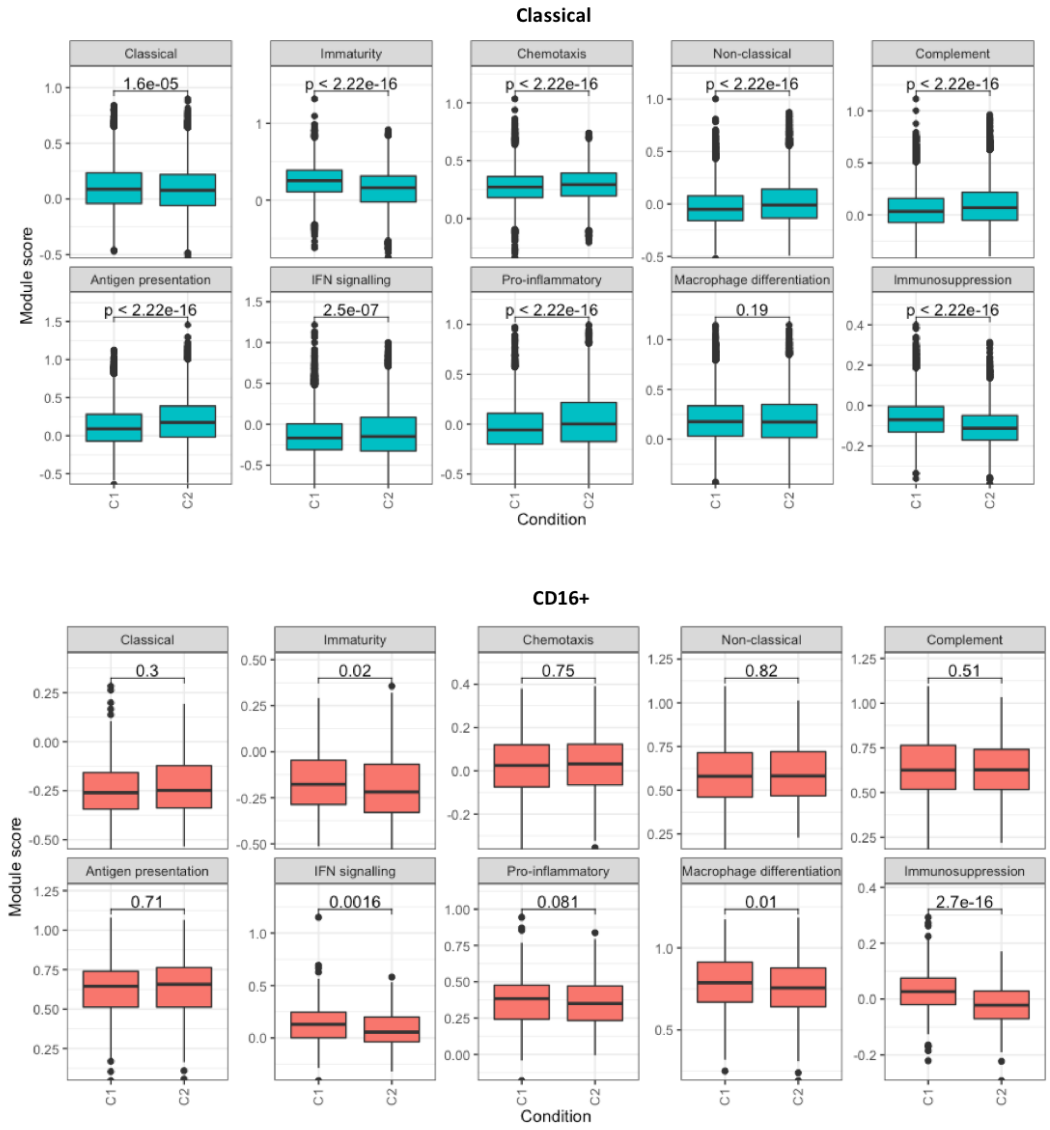


Figure 42: Transcriptomic response to ICB characterised by module scoring
Module expression score in MM patients at baseline ('C1') and following ICB treatment ('C2') (Wilxocon signed-rank test) in classical monocytes (8979 cells C1, 10691 cells C2) abd CD16^{+} monocytes (342 cells C1, 298 cells C2). Expression of gene set modules per single cell were calculated with Seurat function AddModuleScore.

4.3.6.3 Trajectory analysis explores maturation in classical monocytes with ICB

Trajectory analysis was used to explore changes in differentiation states in monocytes following ICB. Using the trajectory analysis performed in section 4.3.4, changes in pseudotime were explored across untreated and treated samples. No significant change in pseudotime was observed in classical monocytes or CD16^{+} monocytes following ICB

(comparing median pseudotime score per sample, paired Wilcoxon signed rank test) (Figure 43). Treatment-wise responses also showed no significant change in pseudotime score across both classical or CD16^{+} monocytes (data not shown). However, inter-individual variation in maturation response with ICB was seen in the classical compartment, with a significant increase in pseudotime in classical monocytes in the majority of patients (patient IDs 1,4,5,6,8) (Figure 44). Whilst not definitive, this suggests that ICB may be associated with maturation response which is variable across individuals.

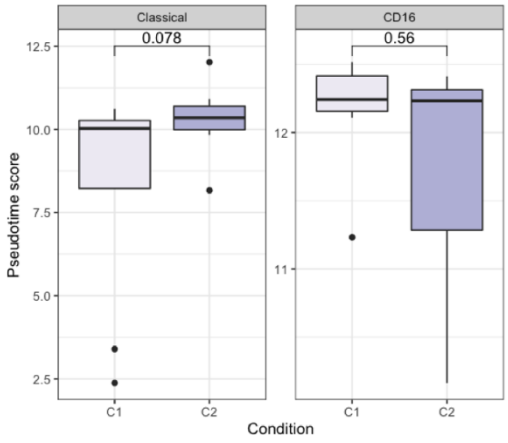


Figure 43: Trajectory analysis characterising ICB responses across monocyte clusters
A) Boxplot to show pseudotime score pre- and post- ICB in MM patients in classical (n=1929 cells) and CD16^{+} monocytes (n = 71 cells). Median pseudotime score per patient was compared across condition (paired Wilcoxon signed-rank test).

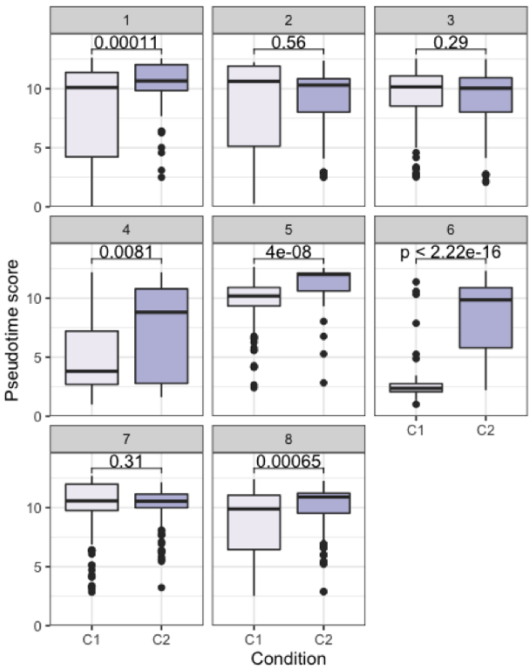


Figure 44: Inter-individual variation in pseudotime following ICB
Boxplot to show pseudotime score across the eight patients that comprise the scRNA-seq cohort (ID 1-8). Pseudotime scores for each single cell were calculated as described in section 4.3.4. Comparison between C1 and C2 performed with Wilxocon signed-rank test.

4.3.7 Monocyte subset proportions

Proportions of classical and CD16^{+} monocytes following ICB were determined to characterise whether differential transcriptional profiles may be accompanied by shift in monocyte subset size. Classical and CD16^{+} monocyte populations were present in all patients at relatively even proportions, which is supportive of exploring monocyte phenotype initially across two conventional subsets. There was however no significant change in size of either monocyte cluster with ICB (Figure 45), nor when this was explored by treatment type (data not shown). This is in keeping with the stable monocyte subset proportions identified with flow cytomtery in chapter three (3.3.7).

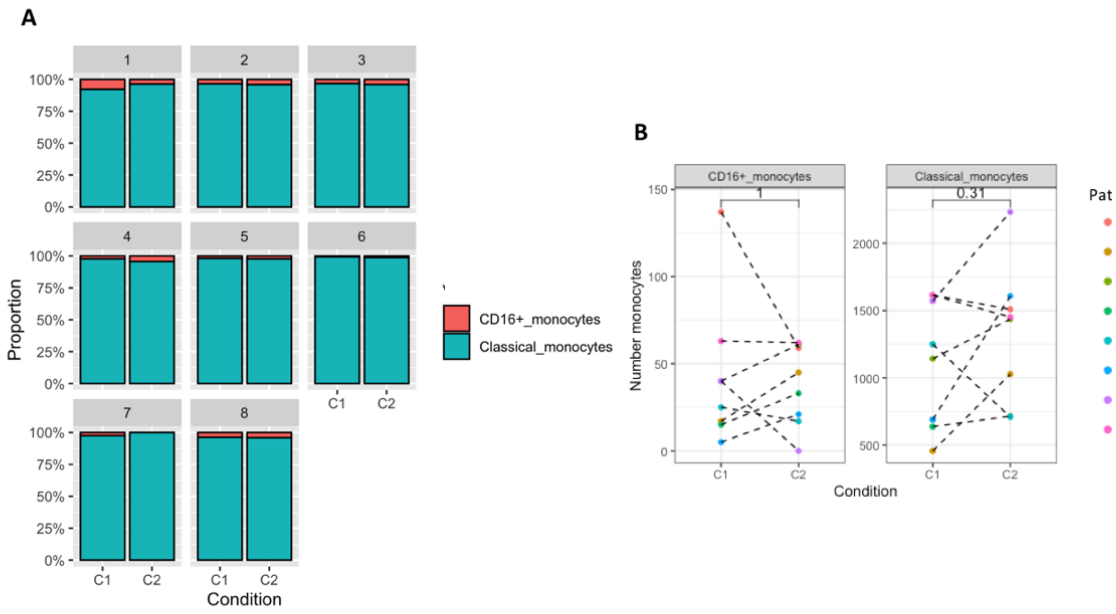


Figure 45: Classical and CD16$^+$ monocyte subset size
A) Proportion of monocyte cluster (classical and CD16$^+$ monocytes) comprising the monocyte population per patient (patient ID 1- 8). B) Plot to show the number of cells per patient ('Patient ID') by condition and monocyte cluster (paired Wilcoxon signed-rank test).

To explore changes in monocyte cluster size across a larger MM patient cohort, the top 40 most significant highly differentially expressed genes defining classical and CD16$^+$ monocytes as defined in section 4.3.3.1 were selected (see Appendix Q for gene sets). The geometric mean expression of these 40 genes was calculated per sample within the normalised bulk RNA-seq dataset described in Chapter Three (across 115 C1 samples, 98 C2 samples and 45 healthy donors). No significant change in classical or CD16$^+$ monocyte score was seen with ICB within the bulk RNA-seq dataset, and no difference was seen when cICB and sICB was characterised separately (see Appendix Z). This is in keeping with the findings with flow cytometry, which showed no significant change in monocyte subset proportion following ICB (section 3.3.3, Figure 22: Monocyte subsets assessed with flow cytometry).

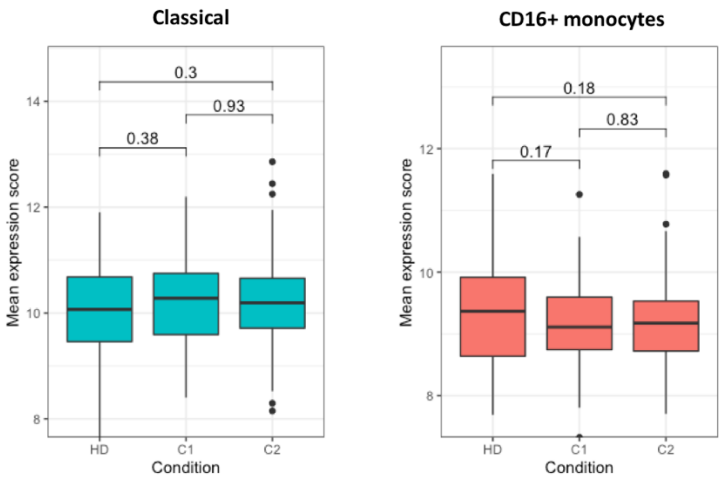


Figure 46: Mean expression scores in bulk RNA-seq data
Normalised expression scores of classical and CD16^{+} monocyte signatures. These were calculated by identifying the top 40 most significantly definitive gene markers of cluster identity and then calculatin the geometric mean expression of that gene set per bulk RNA-seq sample (114 C1 samples, 98 C2 samples and 45 samples from healthy donor samples).

4.3.8 Analysis of classical monocytes demonstrates transcriptional and phenotypic heterogeneity

In section 4.3.2 monocyte subsets 0-6 and 8 were annotated as classical subsets and hence merged to allow exploration of transcriptomic profiles across a dual-categorisation of conventional monocyte subsets (classical versus CD16^{+} monocytes). Gene markers highly expressed within this group were concordant with classical assignment, demonstrating the validity of this approach. However, the classical compartment is large (n=23,600 cells) and shows considerable transcriptional heterogeneity not previously characterised in the literature (Figure 29). Therefore, I next explored the classical population in further detail.

The level of unsupervised clustering resolution (13 dimensions, resolution 0.25) that identified a distinct CD16^{+} population (*Seurat* subset 7), also identified eight classical subsets (subsets 0-6 and 8) (Figure 27). This level was used to explore classical monocyte heterogeneity (see Figure 47).

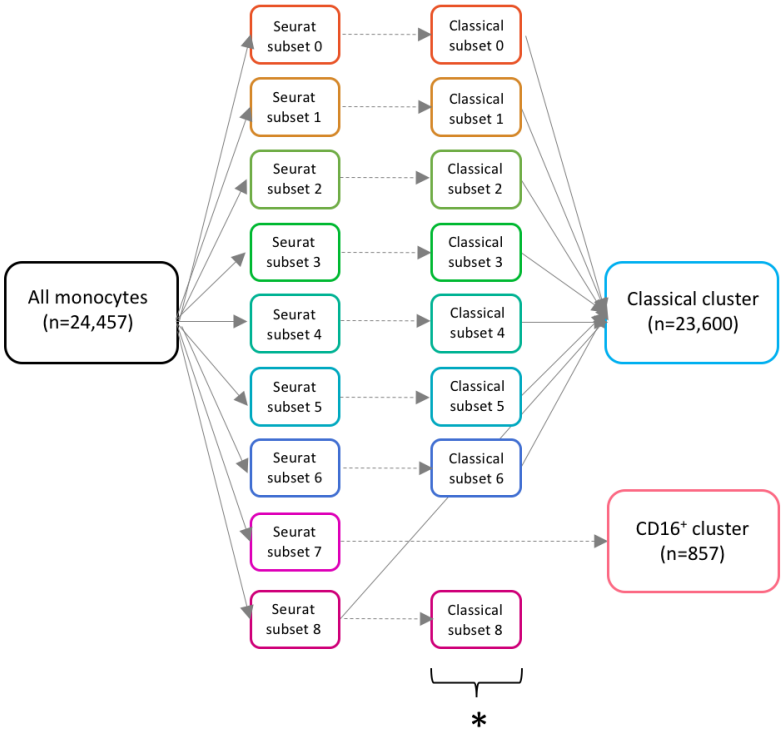


Figure 47: Clustering levels and assignment across the scRNA-seq dataset
Figure to show clusters and cell assignment. Solid arrows indicate clustering or merging of subsets, dashed line indicates change of assignment but not cluster size/cell composition. The starred clustering level indicates the merged classical subsets.

4.3.8.1 Charactering heterogeneity within the classical monocyte compartment

Firstly, the distribution of classical subsets across the patient cohort and conditions was explored. Many patient samples comprised monocyte populations with a single predominant classical subset, and typically a limited diversity of classical subsets (Figure 48). For example, Seurat subset 0 (a classical subset) comprised the majority of the monocyte population in patients 1 (ID1), whilst subset 6 comprised the majority monocytes both before (C1) and after treatment (C2) in patient 5 (ID5) (Figure 48). In monocytes from the healthy donor sample (comprising cells from three HD individuals pooled prior to library preparation), all classical subsets except subset 5 were represented (Figure 48). The classical subsets described here are likely to represent variable activation or maturation within the classical monocyte compartment. It can be hypothesised that circulating monocyte populations in cancer are susceptible to polarisation towards a particular activation or maturation state. The inter-individual variation across the cohort was also marked.

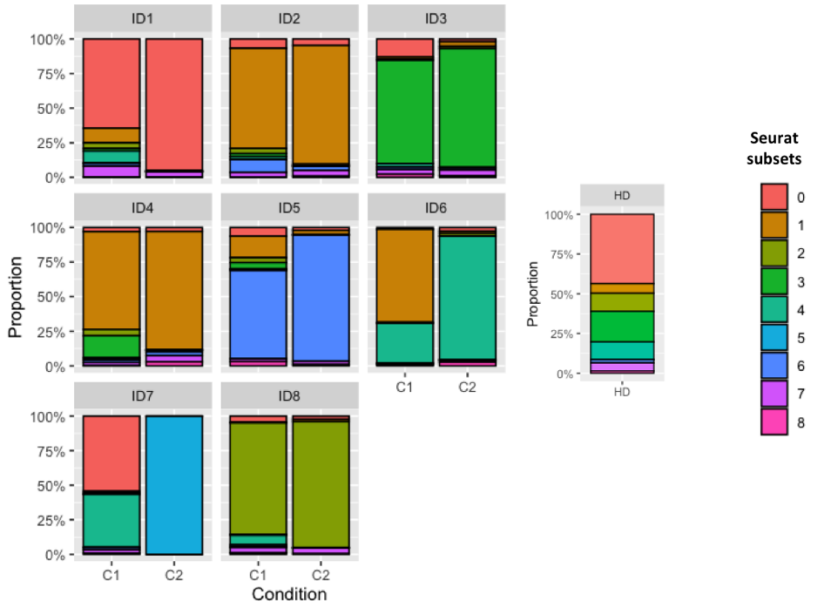


Figure 48: Monocyte subsets by sample
Bar plots to show proportion of Seurat subset comprising each patient sample (ID1-8) and healthy donor sample. Seurat subsets 0-8 are those identified by clustering across all cells (n=24,457 cells) as defined in section 4.3.2.2 with subsets 0-6 and 8 denoted as classical and subset 7 as CD16^{+} monocytes.

Canonical marker genes of these classical subsets were identified using the Seurat function *FindAllMarkers*. This differential expression analysis compares each cluster in turn to all other remaining cells, so was performed without the CD16^{+} monocyte subset (subset 7) to maximise dissection of inter-classical transcriptional variation, rather than classical versus CD16^{+} variation (see heatmap in Figure 49A). Pathway enrichment analysis across subset-defining genes and gene expression module scoring was then performed, to further characterise these classical monocyte subsets.

Canonical marker genes highly expressed in classical subset 0 included antigen presenting genes such as *HLA-DRA* (P_{adj}= 4.8x10^{-97}), *HLA-DQB1* (P_{adj}= 6.0x10^{-235}) and *HLA-DPB1* (P_{adj}= 1.4x10^{-202}), as well as classical complement genes *C1QA* (P_{adj}= 3.1x10^{-83}), *C1QB* (P_{adj}= 1.5x10^{-153}) and *C1QC* (P_{adj}= 1x10^{-300}). Transcripts associated with response to IFN including *GBP1* (P_{adj}= 5.9x10^{-105}), *GBP2* ($P_{ad\ j}$= 1.2x10^{-26}) and *GBP5* (P_{adj}= 1.5x10^{-163}), as well as *STAT1* (P_{adj}= 9.4x10^{-32}), were highly expressed in this subset. Other highly significant canonical gene markers definitive of this subset include *EEF1A1* (P_{adj} = 1x10^{-}

^{300}), which has a role in cytoskeletal regulation and apoptosis, *LYZ* (P_{adj}= $1x10^{-300}$), *TMEM176A* (P_{adj}=$1x10^{-300}$) and *TMEM176B* (P_{adj}= $1x10^{-300}$). *TMEM176B* is an innate immune checkpoint, which acts to inhibit activation of the inflammasome inhibiting $CD8^+$ T cell responses[422]. Blockade of TMEM176B has been shown to enhance the activity of anti-PD-1 and anti-CTLA-4 therapy in mice[422]. Subset 0 exhibited high non-classical, antigen-presenting, complement, inflammation and response to IFN-γ gene expression module scoring and low levels of immaturity and chemokine module expression (Figure 49).

Subsets 1, and particularly 4, demonstrated high expression of classical, chemokine and immaturity module expression scores (Figure 49). In keeping with this, subset 4 was defined by many key classical-type gene markers, such as *CD14* (P_{adj}= $1.5x10^{-232}$), immaturity marker *CD36* (P_{adj} = $5.4x10^{-95}$), chemokine receptor *CCR2* (P_{adj}= $2.1x10^{-97}$) and *S100A9* (P_{adj}=$1x10^{-300}$), *S100A9* (P_{adj}= $1x10^{-300}$) genes. Subset 1 also demonstrated a prominently classical-type transcriptional profiles, with transcripts including immaturity marker *CD36* (P_{adj} = $8.6x10^{-25}$) and *CD14* (P_{adj} = $4.2x10^{-46}$), *FOS* (P_{adj}= $2.3x10^{-92}$) and *VCAN* (P_{adj} = $5.3x10^{-137}$) highly expressed. Subset 1 exhibited the highest percentage of mitochondrial genes amongst the classical subsets, in keeping with a proliferating, mitotically active cell group (Appendix BB).

Subset 2, although present in patients 1, 2, 4, 6 and 8, was found predominantly in patient 8, comprising the vast majority of this patient's classical monocyte population both pre- and post-treatment (Figure 48). Subset 2 was characterised by high expression of *MTRNR2L8* (P_{adj}= $1x10^{-300}$) and its paralogue *MTRNR2L12* (P_{adj}= $8.1x10^{-230}$), which are mitochondrial-derived peptides reported to be neuroprotective, cytoprotective against stress and which have antiapoptotic effects[423,424]. These genes are expressed predominantly in granulocytes, but also by monocytes[425] and recent work demonstrated that differential methylation of this gene is associated with the diagnosis of large-artery atherosclerotic stroke[426]. *MTRNR2L8* and *MTRNR2L12* were also amongst the most highly expressed genes in subset 5.

Classical subset 5 was found only in patient 7 following ICB treatment, where it comprised the entire monocyte population. It was characterised by expression of IFN-

induced, complement (*C1QA, C1QB, C1QC*) and antigen-presenting genes. This cluster exhibited particularly high antigen-presenting, complement, interferon-response, and pro-inflammatory module scores, and interestingly the lowest immaturity score, despite high chemokine signalling score, and moderate classical expression score (Figure 49). This cluster demonstrated prominent expression of ribosomal (but not mitochondrial) genes (Appendix BB). It may be that this cluster represents an activated and mature classical population showing evidence of stress and/or degeneration. Interestingly, patient 7 received cICB and experienced autoimmune colitis refractory to steroid treatment and which required infliximab therapy. This patient progressed rapidly and died within a few months of starting ICB. It was noted that this subset was the only one not present within the healthy donor samples (Figure 48).

Subset 6 was characterised by prominent expression of IFN-induced genes, including mpst significantly *IFITM1* ($P_{adj}<1x10^{-300}$), *IFI44L* ($P_{adj}<1x10^{-300}$) and *IFITM3* ($P_{adj}<1.0x10^{-300}$). This cluster was also characterised by high expression of non-classical markers *FCGR3A* ($P_{adj}= 9.6x10^{-188}$) and *CDKN1C* ($P_{adj}= 1.5x10^{-58}$). Module expression scoring revealed that this cluster exhibited the highest expression of non-classical, antigen-presenting, complement, IFN and pro-inflammatory module scores (Figure 49). Conversely, this subset demonstrated low expression of chemokine, immaturity and CD14 gene expression modules. Classical subset 6 is likely therefore to represent a mature classical monocyte subset.

Classical subset 8 was characterised by high expression of platelet-associated genes including *PPBP* [427] ($P_{adj}= 4.3x10^{-144}$), which encodes CXCL7 which is produced by activated platelets and which has an anti-thrombotic role; *GP9* ($P_{adj} = 3.3x10^{-62}$) , a von Willebrand factor receptor which is required for clotting[428], and *TUBB1* ($P_{adj} = 1.4x10^{-84}$), which encodes tubulin which is specifically expressed in platelets and megakaryocytes[429]. This cluster may therefore reflect a group of monocytes characterised by monocyte platelet aggregates (MPA). MPA are well recognised to be associated with the development of cardiovascular events[430], and are predictive markers of poor prognosis in infections such as COVID-19[431]. Both bulk RNA-seq and scRNA-seq dataset have demonstrated induction of a pro-thrombotic type signature in

MM patients versus healthy donors. These effects may be facilitated via this group of classical monocytes.

Pathway analysis of the significant canonical marker genes across the classical subsets was performed and revealed varied enrichment of GOBP biological processes across the classical compartment (Figure 50). Subset 5 and 6 were enriched for response to type 1 IFN (GO:0060337) and viral responses (GO:0051607, GO:009615), whilst chemotaxis pathways (GO:0006935), typically associated with a classical phenotype, were enriched in subsets 1 and 4. Subset 1 and 4 demonstrated shared enrichment of a high number of GOBP pathways only enriched in subsets 1 and 4, demonstrating shared expression profiles and possible funtion (Figure 50A). Subsets 0,1 and 4 showed enrichment of response to LPS (GO:0032496). Subset 2 showed a distinct profile of pathway enrichment, with pathways including ribosomal subunit assembly GO:0000028), cell morphogenesis (GO:0000902) and fat cell differentiation (GO:0045444).

This demonstrates the distinct and diverse range of function and activation states within the classical monocyte compartment.

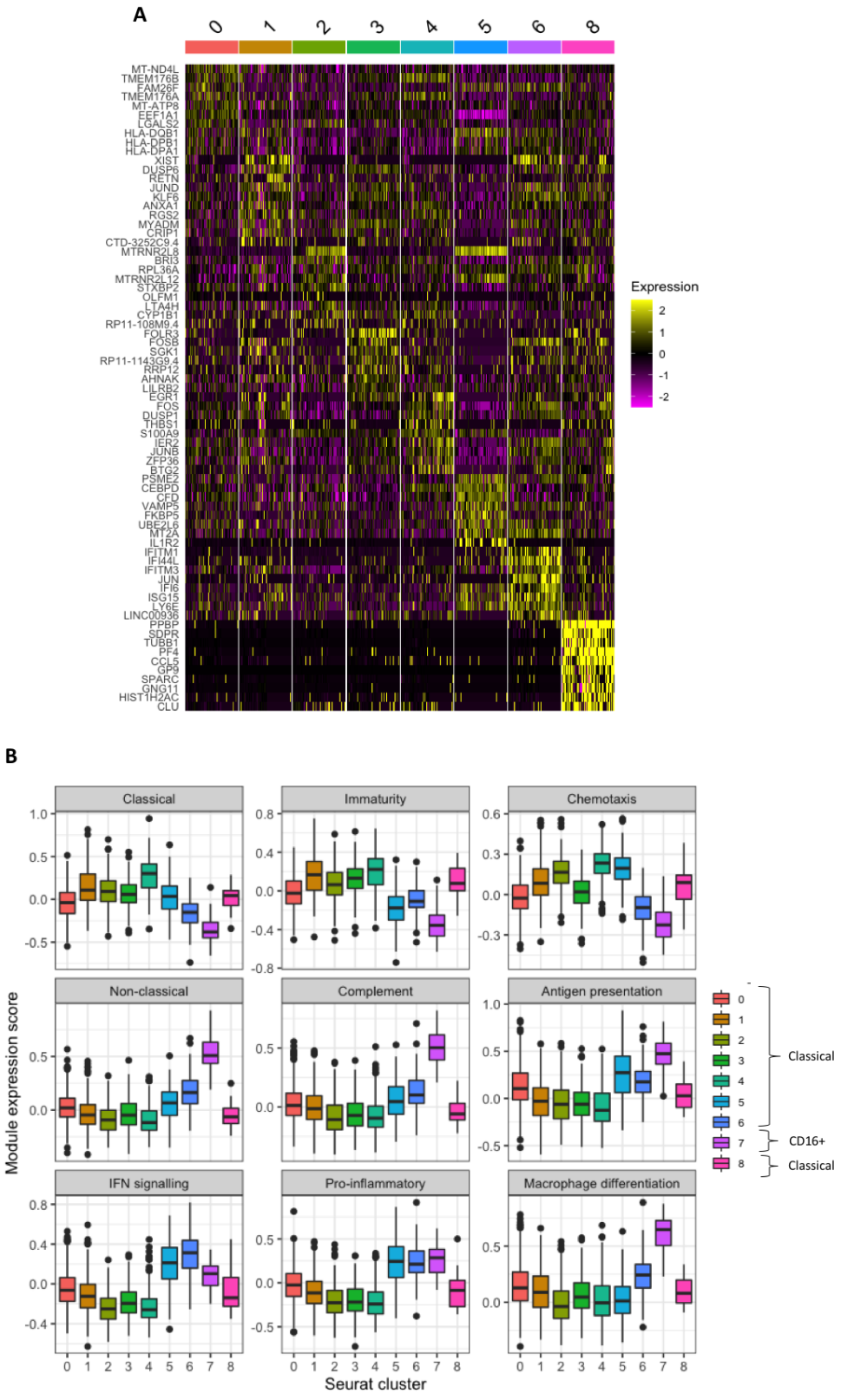


Figure 49: Characterising transcriptional profiles of classical monocyte subsets

A) Heatmap to show top 10 most significant marker genes in classical monocyte subsets 0-6 and 8 (Wilcoxon signed-rank test, P_{adj} <0.05). Key canonical marker genes were identified with FindAllMarkers, with subset 7 (the CD16^{+} monocyte subset) removed from the analysis to facilitate

identification of intra-classical variation rather than classical versus CD16⁺ monocyte diferential expression. B) Module scoring to show module expression score across classical monocyte subsets (cluster 0-6 and 8). Subset 7 is the CD16⁺ monocyte subset, plotted to allow for comparison with the classical subsets (n=24457).

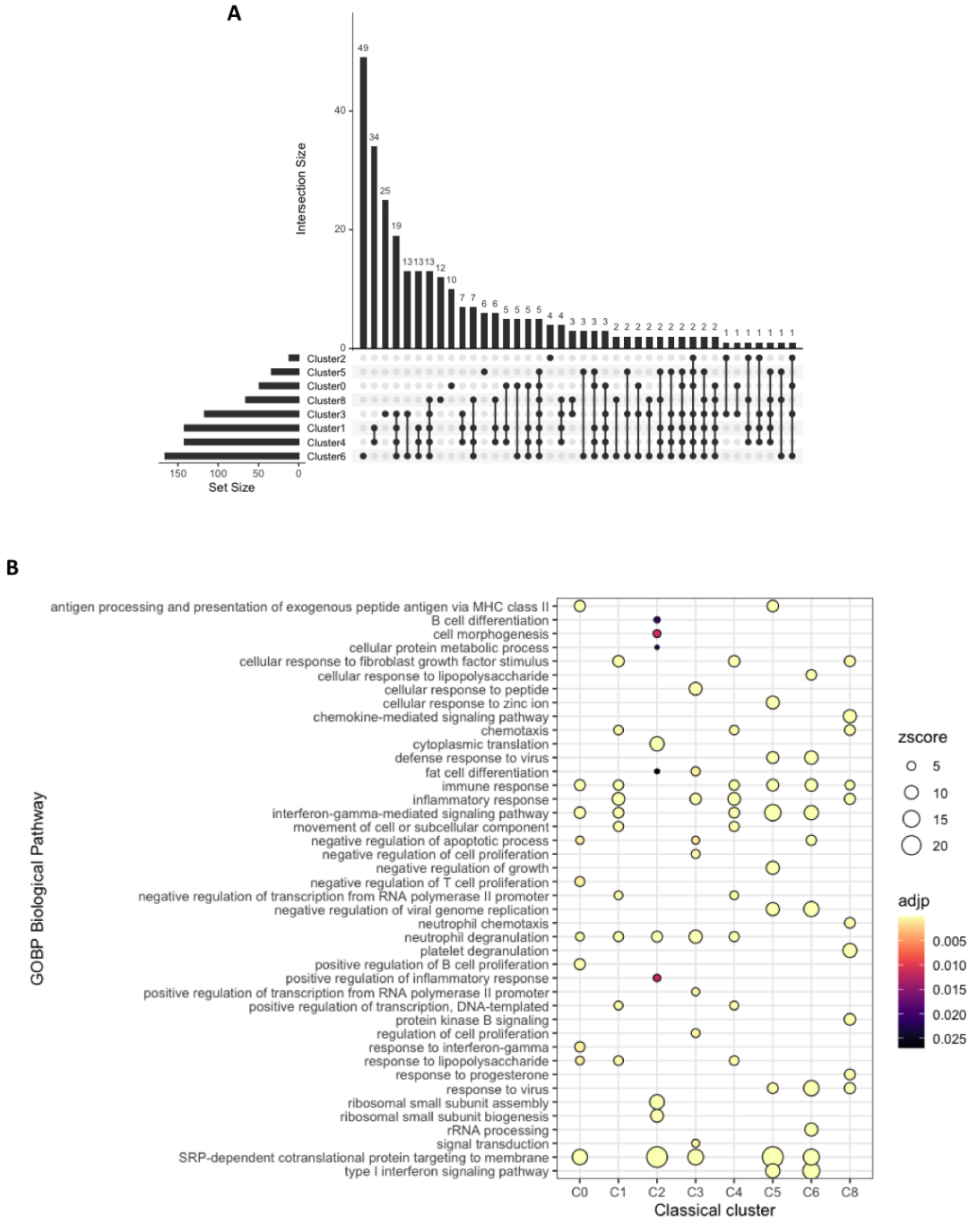


Figure 50: Pathway enrichment of classical monocytes subsets

A) Upsetplot to show intersection of significantly enriched GOBP pathways across classical monocyte subsets (P_{adj}<0.05). B) Dotplot to show top 10 most significantly enriched GOBP pathways per classical monocyte subset (cluster0-6 and cluster 8 ('C0-6, 8') (P_{adj} <0.05).

4.3.8.2 Trajectory analysis demonstrates varying maturation states across the peripheral classical monocyte compartment

To explore differentiation states across the classical monocyte compartment, the trajectory analysis described in section 4.3.4 was visualised and explored according to Seurat clusters assignment, including classical subsets 0-6 and 8 and the CD16^{+} cluster (subset 7). Classical subsets demonstrated marked overlap along the pseudotime trajectory, however there were varying pseudotime scores between cluster types. Visual inspection of the trajectory plot demonstrated that subset 1, and 4 were mostly present towards the origin point, in keeping with a more immature monocyte clusters (Figure 51A). Cells of subset 0 and 6 were seen more towards the termination point, in keeping with a more mature population. The CD16^{+} monocyte subset 7 was located near the termination point, with the highest median pseudotime score, in keeping with a more mature population as previously described. One key branch point was seen (branch point 5, Figure 51A), comprising the majority of cells of subset 5. This suggests that transcriptional heterogeneity observed in monocyte subsets may represent different maturation states and stages of differentiation.

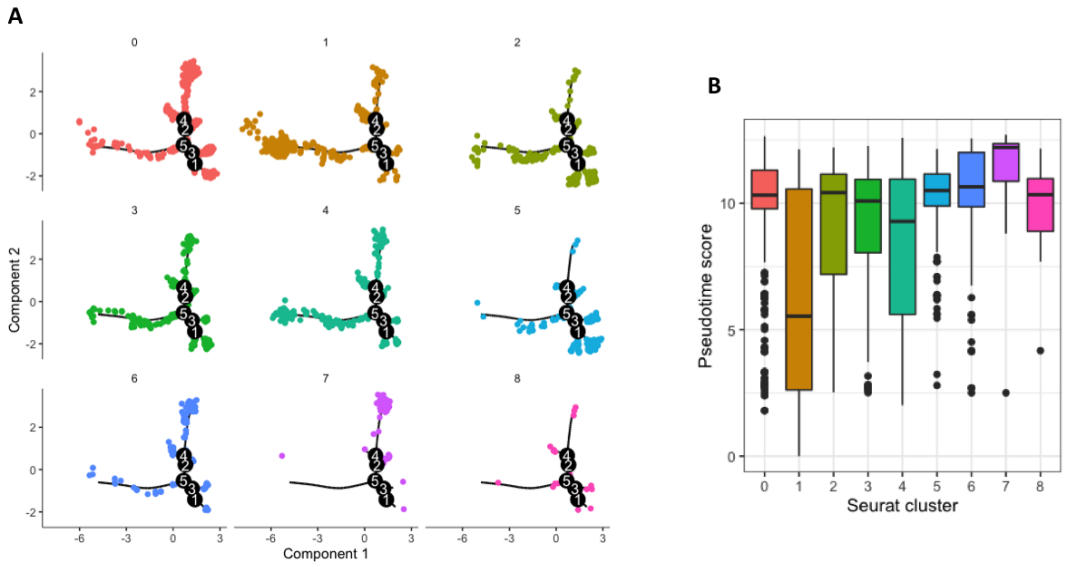


Figure 51: Trajectory analysis showing classical monocyte subsets
A) Pseudotime trajectory plot according to classical monocytes subsets (0-6 and 8) and CD16^{+} subset (cluster 7). B) Pseudotime scores per single cell by classical monocytes subsets (0-6 and 8) and CD16^{+} subset (cluster 7).

To explore the main trajectory branch point (branch point 5) seen on visual inspection of the trajectory plot, genes that showed branch-dependent expression were identified using the *BEAM* function from *Monocle* (Figure 52). At branch point 5 I noted two different branches – branch A and branch B (Figure 52A). At branch point 5, genes including numerous ribosomal genes, as well as *MTRNR2L12, MTRNR2L8* and genes associated with metabolism such as *GAPDH*, showed increased expression along branch A (Figure 52). Transcripts associated with classical monocytes profiles such as *S100A8, S100A9*, *S100A12 and VCAN* also showed increased expression along this branch. This is in keeping with a branch comprising activated and metabolically active classical monocytes. The branch-dependent gene profile varied across classical subsets.

Transcripts showing branch-dependent expression with increased expression along branch B included non-classical marker genes such as *FCGR3A, IFITM2 and IFITM3* (Figure 52). Expression of transcripts associated with classical monocytes such as *VCAN,* and *S100A8, S100A9* and *S100A12* were reduced along this branch from the branch origin point (Figure 52). This is in keeping with acquisition of a mature, non-classical expression profile along this branch point. Interestingly, whilst expression of MHC class II antigen presenting genes were branch-dependent and increased along branch B, there was also an increased expression in one classical subset along branch point A, in keeping with activation of this subset.

This suggests that branch A, populated mostly by subset 5, represents an activated and possibly degenerating, exhausted cell population. Along branch B, maturation was associated with acquisition of a non-classical transcriptomic profile.

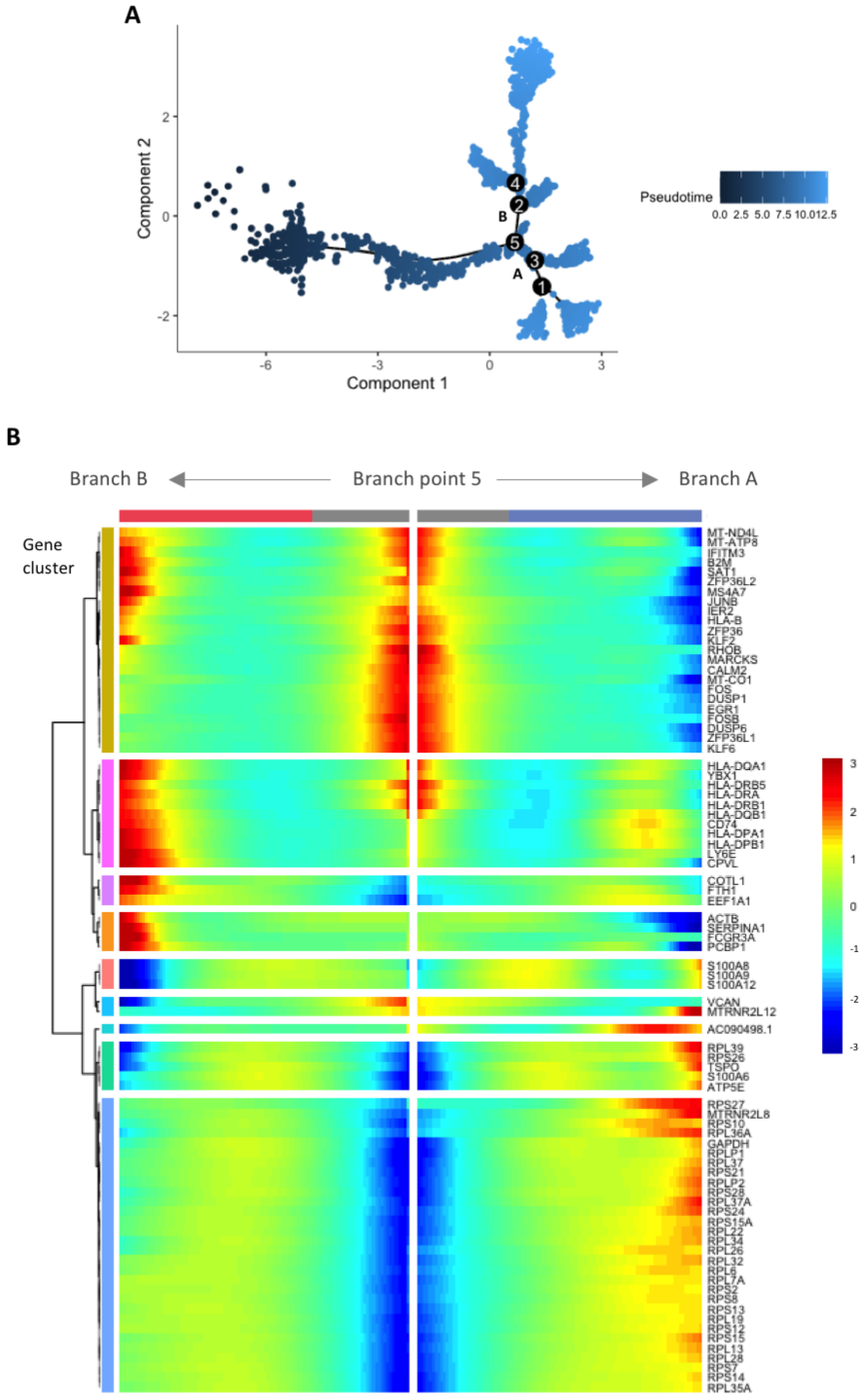


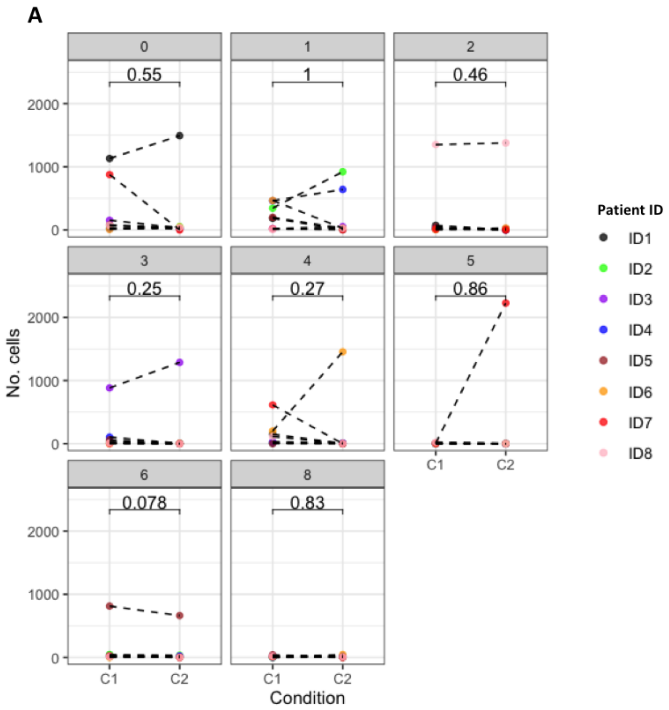
Figure 52: Analysis of trajectory branch points

A) Trajectory plot across subsampled monocytes (n=2000 cells) showing pseudotime score. Branch point 5 is indicated and alternative branch fates at branch point 5 (branch A and branch B) are labelled. B) Heatmap to show relative expression of branch-dependent genes at branch point 1 (plot left) and branch point 2 (plot right) with increasing pseudotime (qval < 1e-34). The scale indicates relative gene expression from 3 (high expression) to -3 (low expression). Branch dependent genes were identified with the BEAM function with Monocle.

4.3.8.3 Classical monocyte subsets demonstrate varying responses to ICB

To explore whether ICBs is associated with a change in the classical subset populations, classical subset size was explored across conditions. No significant change in any of the classical subsets was observed following treatment (Figure 53A).

Gene expression module scoring was then used to explore transcriptional changes in response to ICB across classical monocyte subsets (as described in section 4.3.3.2). There were distinct patterns of transcriptional response across the subsets (Figure 53B). Broadly, subset (subset 1 and 6) exhibited induction of non-classical, complement, IFN-γ, inflammation and macrophage differentiation module expression scores and downregulation of classical, and immaturity module scores (Figure 53B). Subset 0 and 5 also demonstrated induction of antigen presentation expression scoring (Figure 53B). Conversely, subset 4 and 8 demonstrated increase of classical and chemotaxis expression scores, in keeping with acquisition of a classical-type expression profile (Figure 53B). This might reflect generation of immature circulating monocytes from myeloid precursors. This suggests that whilst overall there is maturation of the classical compartment following ICB with acquisition of a non-classical pro-inflammatory phenotype, response to ICB within this compartment is dynamic and limited to only some classical monocyte populations. Interestingly, ICB was associated with reduction in immunosuppression expression score across all classical subsets except 3 and 8.



B

Classical Subset	Classical	Chemotaxis	Immaturity	Non-classical	Complement	Antigen presentation	IFN signalling	Pro-inflammatory	Macrophage differentiation	Immunosuppression
0										
1										
2										
3										
4										
5										
6										
8										

Increased expression score C1vC2 (p-value <0.05)
Reduced expression score C1vC2 (p-value <0.05)
No change in expression score C1vC2 (p-value>0.05)

Figure 53: Classical monocyte subset response to ICB
A) Plot to show the classical subset size (number of cells) by condition (bar plots labelled 1-6 and 8 indicate classical subsets). Dots are coloured by patient ID (1-8) (paired Wilcoxon signed-rank test). B) Table to show the response to ICB ('C1vC2') across the classical monocyte subsets as per module expression score. As per the key, yellow indicates significant induction of module score with treatment (Wilcoxon signed-rank test, p<0.05) and blue indicates significant reduction in module score (Wilcoxon signed rank test, p<0.05), with blank squares indicating no significant change with ICB.

To explore classical subset-wise muration following ICB, pseudotime scores in healthy controls, untreated and treated samples were compared across classical monocyte subsets. Whilst across the whole compartment, trajectory analysis demonstrated that classical monocytes were more immature in MM patients than healthy donors (Figure 39), this was limited to classical subsets 1 and 2 (Figure 54A). Pseudotime score was

increased with ICB only in subset 0, indicating maturation of this subset, potentially in keeping with a more intermediate/non-classical phenotype (Figure 54B). This might be due to ICB effecting maturation in only some classical subsets, or that monocytes with a more prominent classical signature are shorter lived, and treatment effects cannot be identified in this immature population. It might also reflect plasticity within the compartment, with cells maturing into other activation/functional phenotypes with immature progenitors maintaining a steady state across the cell population.

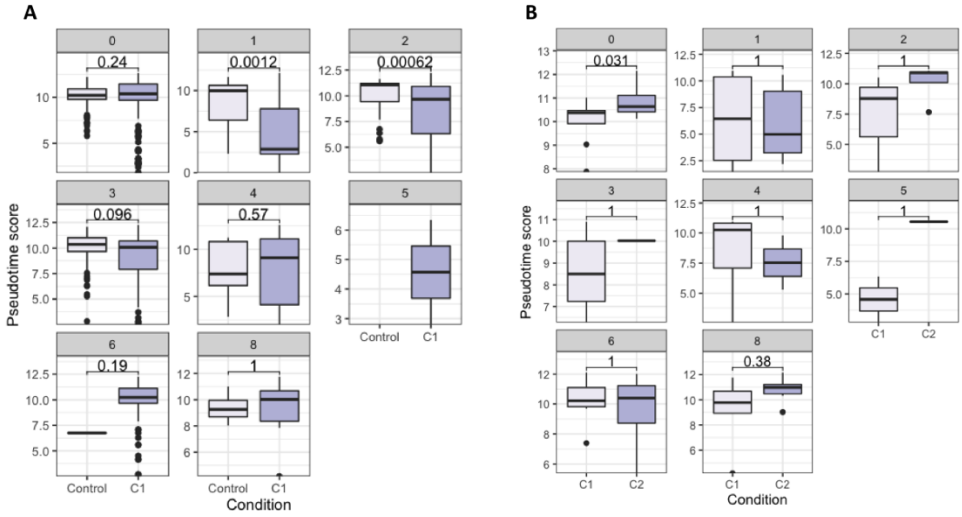


Figure 54: Pseudotime across classical subsets
Boxplots to show pseudotime scores across classical subsets 0-6 and 8 by condition including Control (healthy donors) and C1 (Wilcoxon signed-rank test). B) Boxplots to show median pseudotime scores per patient ID across classical subsets 0-6 and 8 by condition at C1 and C2 (pairwise Wilcoxon signed-rank test). Cells were sub-sampled by randomly selecting 2000 cells from all Seurat clusters across all patients and healthy donor samples.

4.3.8.4 Classical monocyte subset expression profiles and association with clinical outcome

To explore whether classical monocyte subset gene expression profiles were associated with clinical outcome, the top 40 most significant canonical gene markers highly expressed in each classical subset were used to calculate a geometric module expression score per sample in the bulk RNA-seq data (see Appendix R and Appendix S for gene sets). Association between normalised expressed score and OS and PFS was explored using Kaplan Meier analysis and log-rank tests. Post-treatment (C2) both classical subsets 5 and 6 (the subset associated with an activated classical and a more mature,

non-classical type phenotype respectively) normalised expression scores were associated with reduced OS (92 patients, p-value <0.05) (Figure 55). There was no association with subset 0-4 and 8 normalised expression scores and OS (data not shown).

Whilst classical subset 6 appears to be a more mature classical subset with increased expression of non-classical-type markers, exploration of the gene set for module analysis (the top most significant highly expressed genes) revealed numerous interferon-induced genes (*IFITM1, IFI44l, IFITM3, IFIT3* and *IFI27,* as well as several classical-type genes such as *JUN, FOSB, LY6E* and *FOS*) (Appendix S). Therefore, the module scoring may best be regarded as a classical-type score with prominent interferon gene expression.

Association between normalised expression scores across all classical subsets and irAEs (any grade prior to treatment cycle 5) at baseline and following ICB treatment were also explored, with no significant association identified (data not shown).

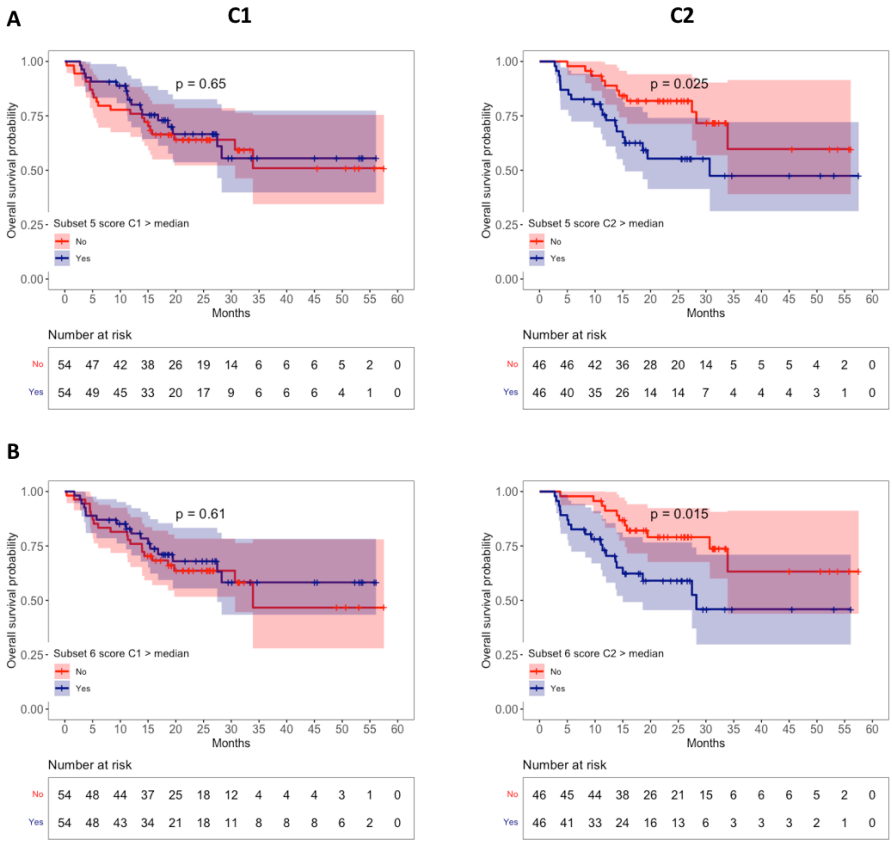


Figure 55: Kaplan-Meier curves of overall survival (OS) by classical subset expression score in bulk RNA-seq data

Survival curves to show classical subset geometric mean score and association with OS. The top 40 most highly differentially expressed transcripts in each classical subset were taken to create the gene set and a geometric mean expression score across these 40 transcripts calculated per sample in bulk RNA-seq data (108 C1 samples, 92 C2 samples). Survival analysis was run with the survival package. Survival curves were plot according to expression score per subset above or below the median value. Shaded areas indicate 96% confidence interval. Survival curves to show: A) Classical subset 5 expression, B) classical subset 6 expression score.

4.3.9 Analysis of CD16^{+} monocytes identifies an intermediate monocyte cluster

Historically, initial studies identified two main monocyte subsets: a CD14^{hi}CD16^{low} population, and a CD14^{low}CD16^{hi} population. Subsequently, a third intermediate monocyte population, characterised by intermediate CD16 expression and prominent MHC class II antigen presentation, was identified[432]. In the initial annotation of the scRNA-seq as described in section 4.3.2.4, no distinct intermediate population was identified. However, the CD16^{+} compartment ('Seurat subset 7') exhibited the most prominent antigen presenting expression profile, and annotation with other expression

sets using *SingleR* demonstrated an intermediate signature within this subset (Figure 26, Figure 27C). Notably, the identification of a robust intermediate population using scRNA-seq has to date been problematic. The few studies that have characterised monocyte populations with scRNA-seq have either identified two populations ($CD14^{+}CD16^{-}$ classical population and $CD14^{low}CD16^{++}$ populations)[135], or three/four populations, but with these additional groups likely to represent additional classical subsets, or doublet contaminants[104,133,134].

To explore transcriptional and functional heterogeneity within the $CD16^{+}$ monocyte compartment and to identify an intermediate-like subset, additional clustering analysis of the $CD16^{+}$ monocyte compartment was performed (n=857 cells). The number of principle components (PCs) for dimensionality reduction were chosen based on visual inspection of the data on the Seurat function *Elbowplot.* Unsupervised clustering (with 13 principal components for dimensionality reduction, clustering resolution 0.3) of the $CD16^{+}$ monocyte group was performed. Three distinct CD16+ clusters were identified (see Figure 56).

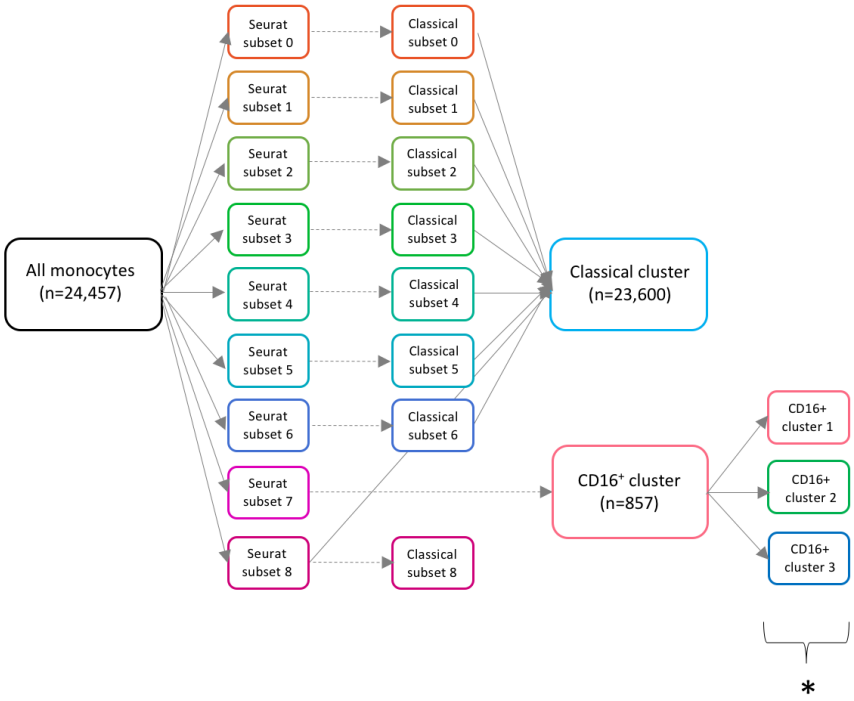


Figure 56: Clustering levels and assignment across the scRNA-seq dataset
Clustering levels and assignments across the scRNA-seq dataset. Solid arrows indicate clustering or merging of subsets, dashed line indicates change of assignment but not cluster size/cell composition. The starred group of clusters infers the $CD16^{+}$ subpopulations (CD16+ cluster 1, CD16+ cluster 2 and CD16+ cluster 3).

Key marker genes highly expressed within these clusters were identified with the *FindAllMarkers* function. This identified a subset characterised by prominent MHC class II expression and low *FCGR3A* expression ('CD16$^+$ cluster 1') (Figure 57B) in keeping with an intermediate-like monocyte subset. Two further CD16$^+$ monocytes subsets were identified at this clustering resolution. The first, 'CD16$^+$ cluster 2', was characterised by relative high expression of classical monocyte marker such as *S100A8*, *LYZ* and *FOLR3*, as well as the immaturity marker *CD36* and chemokine receptor *CCR2* (Figure 57B). The second, CD16$^+$ monocyte cluster 3, was characterised by relative high expression of *FCGR3A* and other non-classical markers such as *CDK1NC,* with lower relative expression of classical gene markers (Figure 57). These three CD16$^+$ subsets were present across patient samples and in the merged heathy donor samples, with a few exceptions. For example, at baseline patient ID6's CD16$^+$ population comprised solely CD16$^+$ cluster 1 (Figure 58).

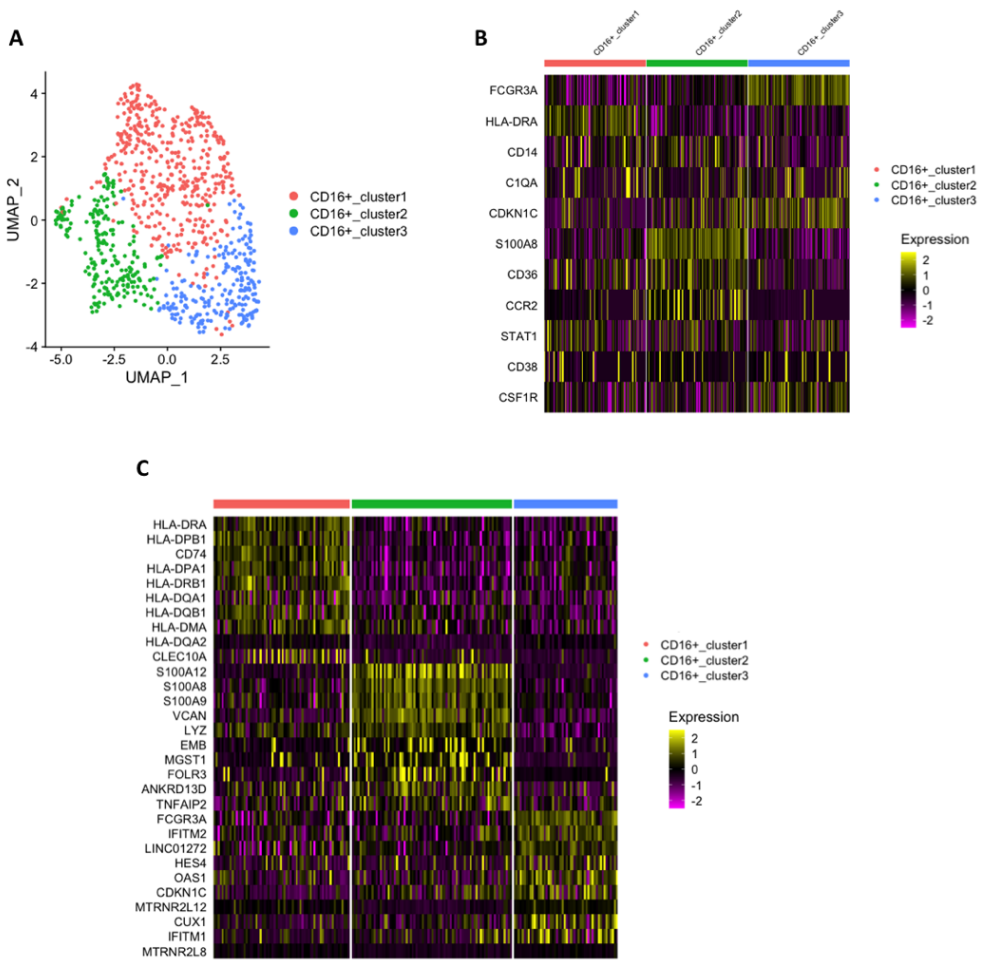


Figure 57: Diverse transcriptional profiles within the CD16^{+} monocyte cluster
A) UMAP plot to show clustering of CD16^{+} monocytes from Controls, C1 and C2 patients (857 cells) with each dot representing a single cell and clusters represented by colours (clustering dimension 13, resolution 0.3). B) Heatmap to show relative expression across CD16^{+} monocyte clusters 1-3 (subsampled for plotting to n=100 per cluster) of key phenotype and function marker genes including FCGR3A, HLA-DRA and CD14. C) Heatmap to show relative expression of top 10 genes by log2FC significantly enriched gene marker per CD16^{+} subset (200 cells visualised per cluster).

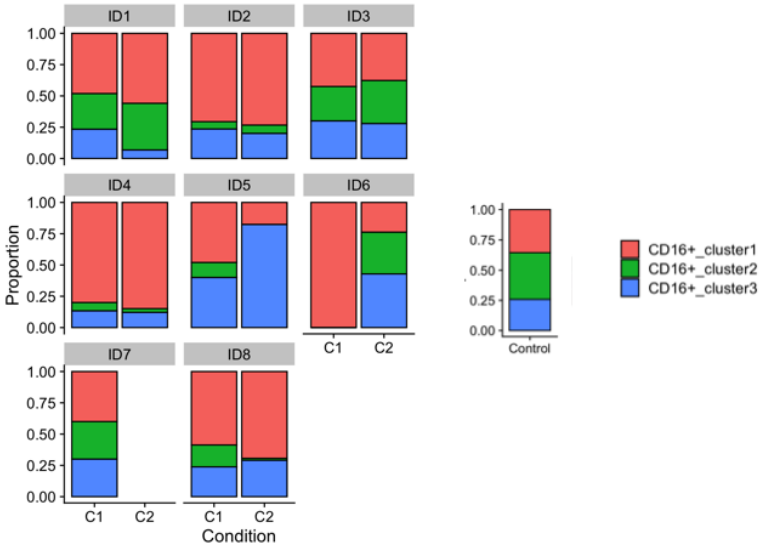


Figure 58: CD16^{+} monocyte cluster proportions
Barplot to show proportion of CD16^{+} clusters that comprise the CD16^{+} monocyte compartment by patient (ID1-ID8) and across the three merged healthy donor samples ('Control').

Pathway enrichment analysis across all significantly highly expressed genes across the three CD16^{+} monocyte clusters was performed. Most enriched GOBP pathways demonstrated cluster-wise enrichment, with only four pathways showing shared enrichment (Figure 59A). The *HLA-DRA^{hi} FCGR3A^{low}* monocyte cluster ('CD16^{+} cluster1') demonstrated selective enrichment of antigen presentation via MHC class II (GO:0019886), the T cell receptor signalling pathway (GO:0050852) and IFN-γ production (GO:0032729) (Figure 59). CD16^{+} cluster 2 demonstrated selective enrichment of innate immune response (GO:0045087), cellular response to calcium ion (GO:0071277), blood coagulation (GO:0007596) and chemokine signalling (GO:0070098) pathways (Figure 59B and Appendix AA). CD16^{+} cluster 3, which demonstrated the highest *FCGR3A* expression across the CD16^{+} monocyte compartment, demonstrated enrichment of response to virus, MAPK signalling (GO:0000165), and ageing (GO:0007568) pathways (Figure 59B and Appendix AA). Expression profiles are best considered as markers of relative expression within the CD16^{+} compartment. This demonstrates that although CD16^{+} monocytes make up a small proportion of the peripheral monocyte compartment, there is even transcriptional

heterogeneity and functional variation within this population, which may be reflective of varying maturation and functional states.

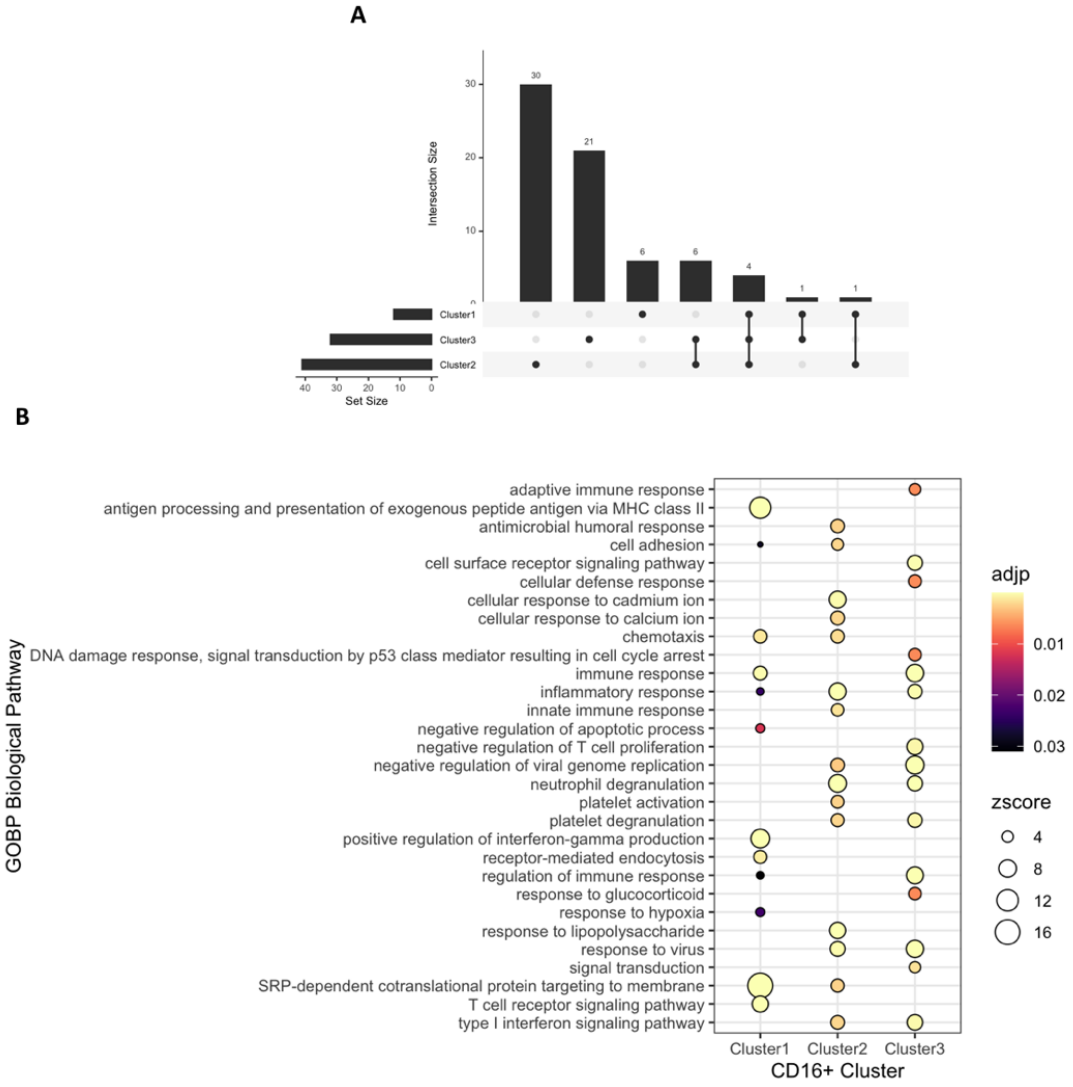


Figure 59: Pathway enrichment analysis of CD16$^+$ clusters
A) Upset plot to show intersection of significantly enriched GOBP pathways between CD16+ monocyte clusters, CD16$^+$ cluster 1, CD16$^+$ cluster 2 and CD16$^+$ cluster 3. B) Top 15 most significantly enriched pathways across CD16+ monocytes clusters 1, 2 and 3 (P_{adj} <0.05).

To explore association with the CD16$^+$ clusters and clinical outcome, the top 40 most significant canonical gene markers defining each of the CD16$^+$ clusters were used to calculate a geometric mean expression score per sample in the bulk RNA-seq CD14$^+$ dataset (Appendix T for gene sets). Association between the geometric mean expression values with the development of irAEs was determined (Wilcoxon signed-rank test). Survival analysis and visualisation (using the *survival* and *survminer* packages

respectively) were also used to explore association between the geometric mean expression score within the bulk RNA-seq dataset with OS and PFS both at baseline (C1) and post-treatment (C2). Most strikingly, CD16+ cluster 3 expression score at baseline (C1) was associated with improved OS and extended PFS (Figure 60, Figure 61). There was no association between the $CD16^+$ cluster 1 (the 'intermediate-like' cluster) expression score with clinical outcome at baseline, however following ICB this expression score was associated with reduced OS (92 C2 patients). This is in keeping with the observations made in chapter three (3.3.8), where the intermediate population size with flow cytometry was found to be associated with reduced OS (although within the flow cytometry data, this association was seen at cycle 4).

The $CD16^+$ cluster 2 geometric mean score post-treatment (C2) was inversely associated with irAEs (Figure 62). Whilst clearly annotated as a $CD16^+$ subset, the $CD16^+$ cluster 2 demonstrated highest expression of classical-associated transcripts amongst the three CD16+ clusters. The $CD16^+$ cluster 2 expression score is comprised of classical-type genes, such as *S100A8, S100A9,* S100A12, *VCAN* and *LYZ*, as well as many ribosomal genes (Appendix T). Association with this signature and irAEs is in keeping with findings in chapter 3, demonstrating inverse association with an immature/classical type signature and irAEs. The prominent ribosomal component comprising the $CD16^+$ cluster 2 signature may however represent monocyte exhaustion or degeneration.

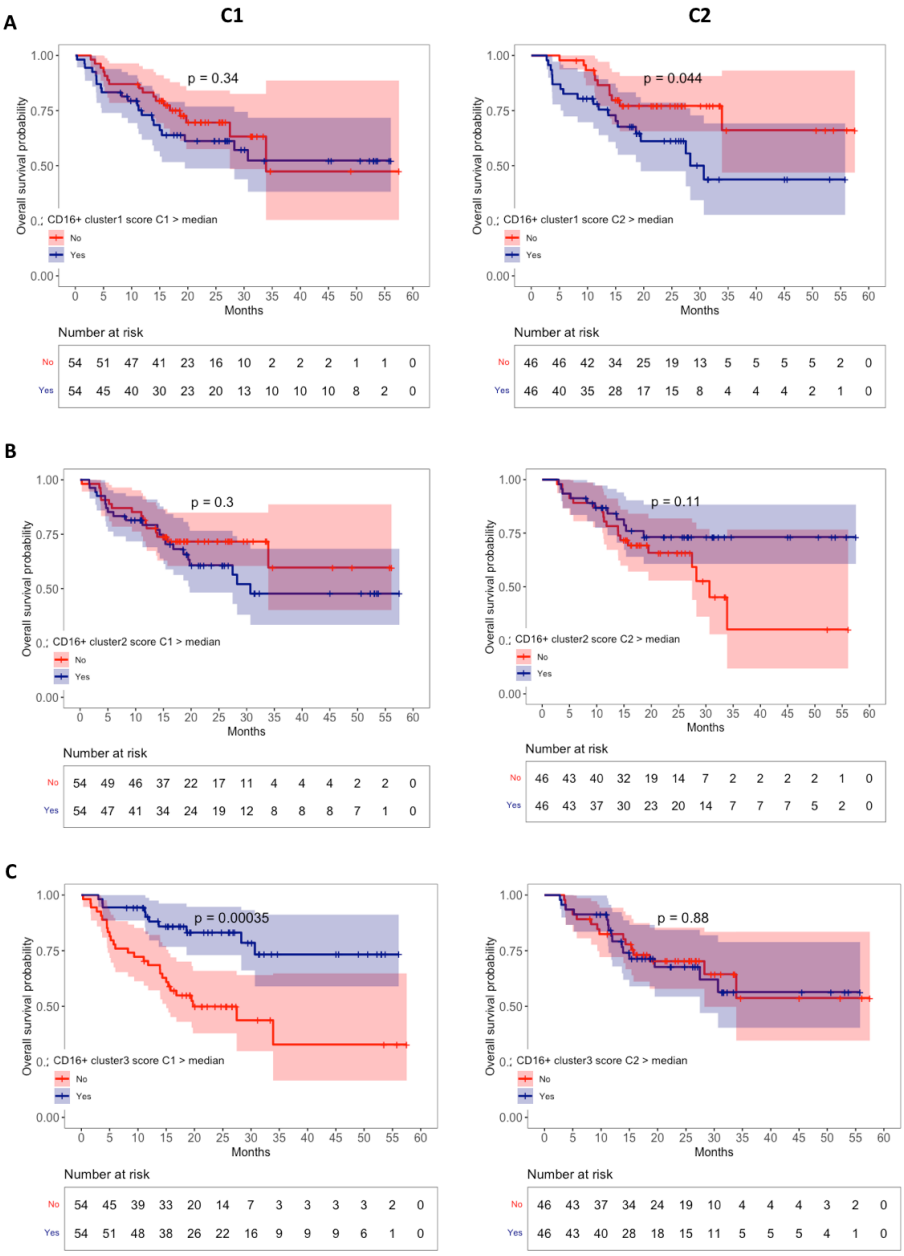


Figure 60: CD16⁺ clusters and overall survival (OS)

Kaplan Meier curves of OS in: A) CD16⁺ cluster 1; B) CD16⁺ cluster 2 and C) CD16⁺ cluster 3 expression score within the bulk RNA-seq data dichotomised into those above ('YES') or below the median value ('NO') at both pre-treatment and post-treatment at C1 and C2 (108 C1 patients, 92 C2 patients. Shaded areas indicate 95% confidence interval.

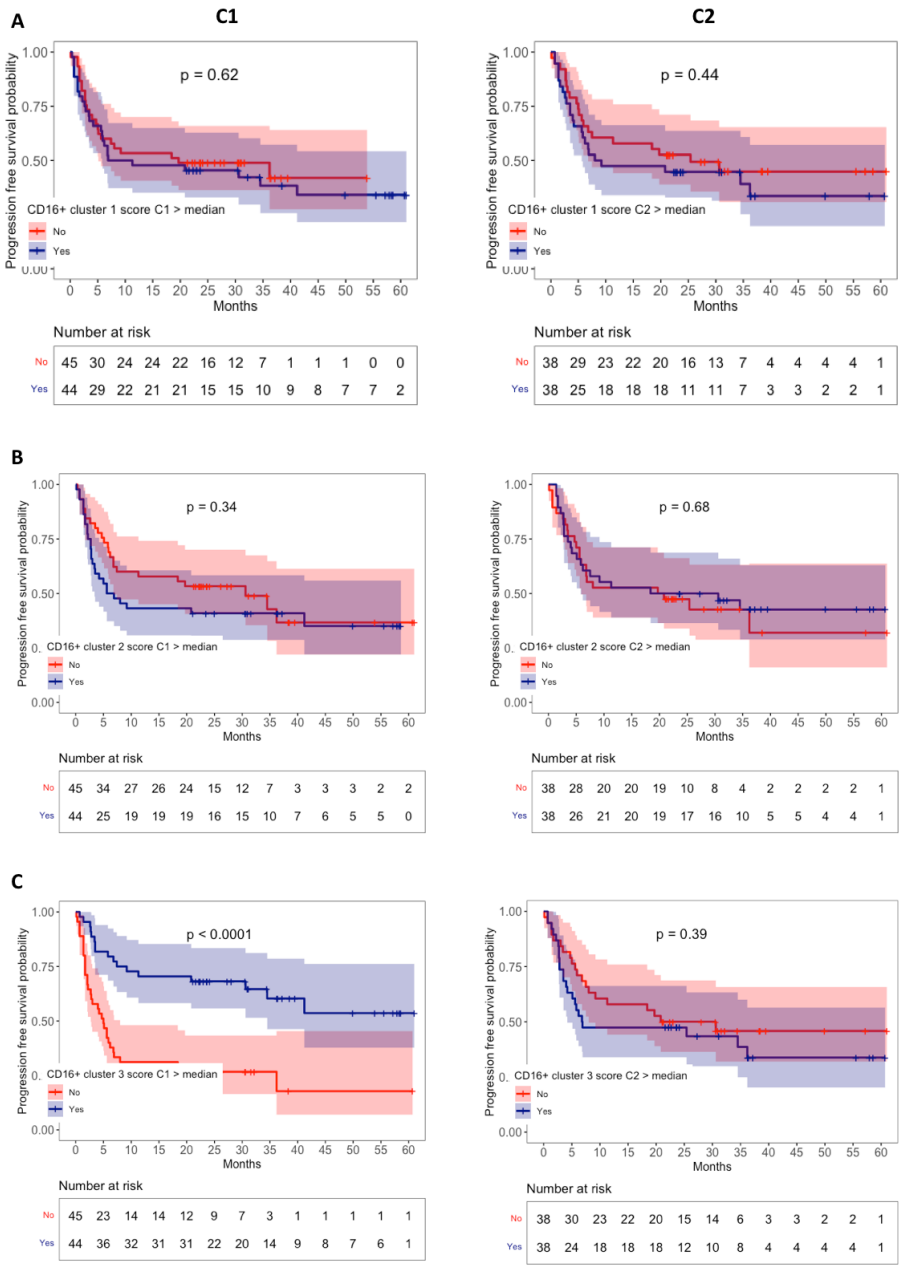


Figure 61: CD16$^+$ cluster expression scores and progression free survival (PFS)
Kaplan Meier curves of PFS according to: A) CD16$^+$ cluster 1; B) CD16$^+$ cluster 2 and C) CD16$^+$ cluster 3 expression score within the bulk RNA-seq data dichotomised into those above ('YES') or below the median value ('NO') at both pre-treatment and post-treatment at C1 and C2. Shaded areas indicate 95% confidence interval (89 C1 samples, 76 C2 samples).

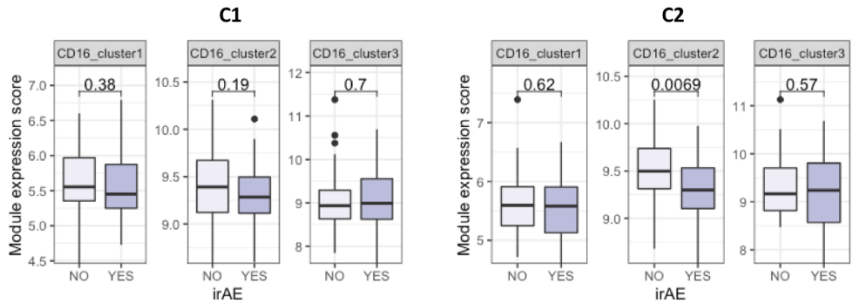


Figure 62: CD16$^+$ monocyte subsets and association with autoimmune side effects
Expression score based on top 40 most significant canonical gene markers of CD16$^+$ monocyte clusters 1, 2 and 3 in bulk RNA-seq data at baseline (C1) and following ICB (C2) and association with development of immune-related adverse events (irAEs) prior to cycle 5 ('irAEs') (n=97 C1, n=82 C2 samples).

4.3.9.1 Pseudotime analysis demonstrates dual termination states within the CD16$^+$ compartment

As transcriptional profiles in the CD16$^+$ cluster were suggestive of spectrum of differentiation, trajectory analysis was used to explore differentiation states within the CD16$^+$ monocyte compartment. Overall, there was no significant difference in pseudotime and therefore maturation state across the three CD16$^+$ subsets. However, a main bifurcation point was seen at branch point 2, where populations diverged into branch (A) characterised predominantly by CD16$^+$ clusters 1 and 2, and branch B comprising all three subsets, including a high proportion of CD16$^+$ cluster 3 cells (Figure 63A&B). This suggests that these subsets are representative of populations with similar maturity, but distinct differentiation fates.

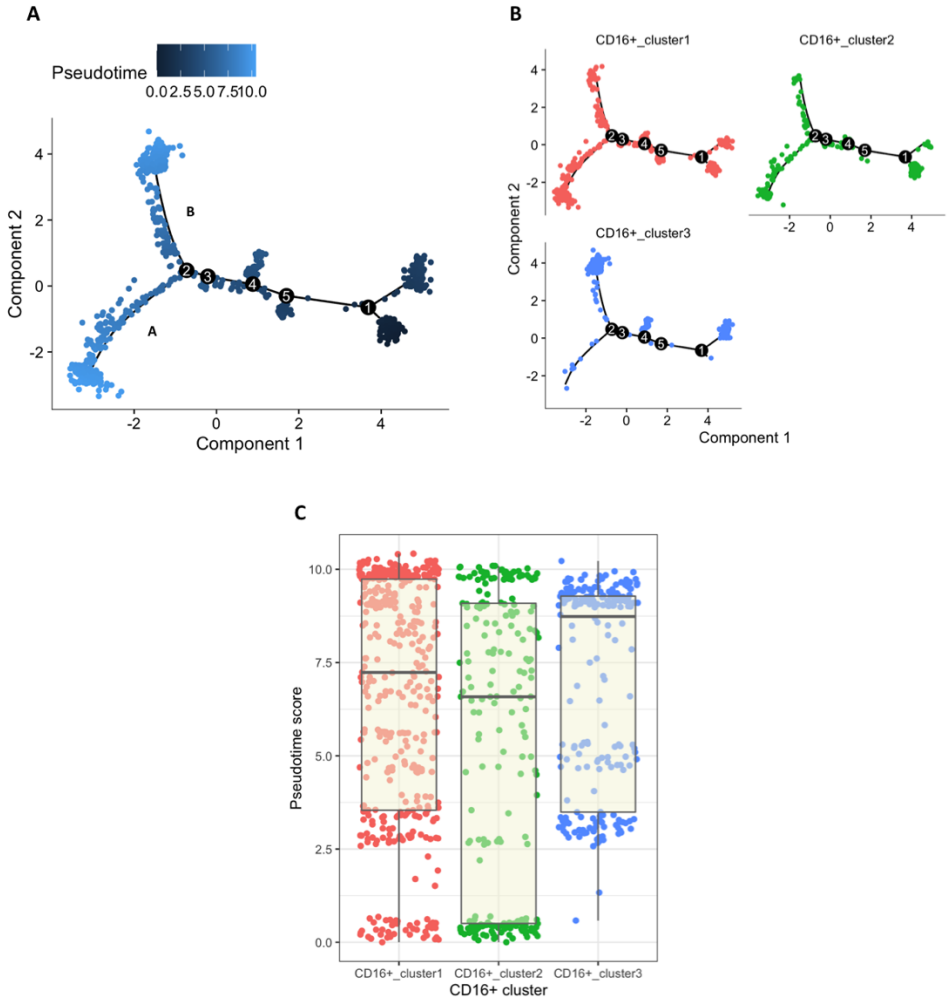


Figure 63: Pseudotime analysis of CD16$^+$ monocytes
A) Pseudotime trajectory plot across CD16$^+$ monocytes with branch points and branches A and B labelled. B) Pseudotime trajectory plots showing CD16$^+$ monocytes coloured by cluster ('CD16_cluster1', 'CD16+_cluster2' and 'CD16+_cluster3'). C) Pseudotime score by CD16$^+$ monocyte clusters (422 cells in CD16$^+$ cluster 1, 217 cells in CD16$^+$ cluster 2 and 218 cells in CD16$^+$ cluster 3).

I then identified genes with branch-dependent expression at the main bifurcation point (branch point 2) with the *Monocle BEAM* function. Genes showing branch-dependency at branch point 2 and increased expression along branch A included antigen-presenting genes and well as classical-phenotype associated genes including calcium metabolism genes *S100A8* and *S100A8, VCAN,* the immaturity marker *CD36*, and numerous ribosomal genes (Figure 64). Branch 2-dependent genes with increasing expression along branch B included markers of a non-classical phenotype including *CDKN1C, FCGR3B,* and *IFITM3 (*Figure 64*).* Whilst previous studies have demonstrated a linear

trajectory of differentiation from classical, to intermediate to non-classical monocyte subsets[433], this suggests that there may be dual termination states, with one characterised by an intermediate monocyte phenotype and high antigen presenting function, and another characterised by expression of non-classical associated genes and a pro-inflammatory phenotype.

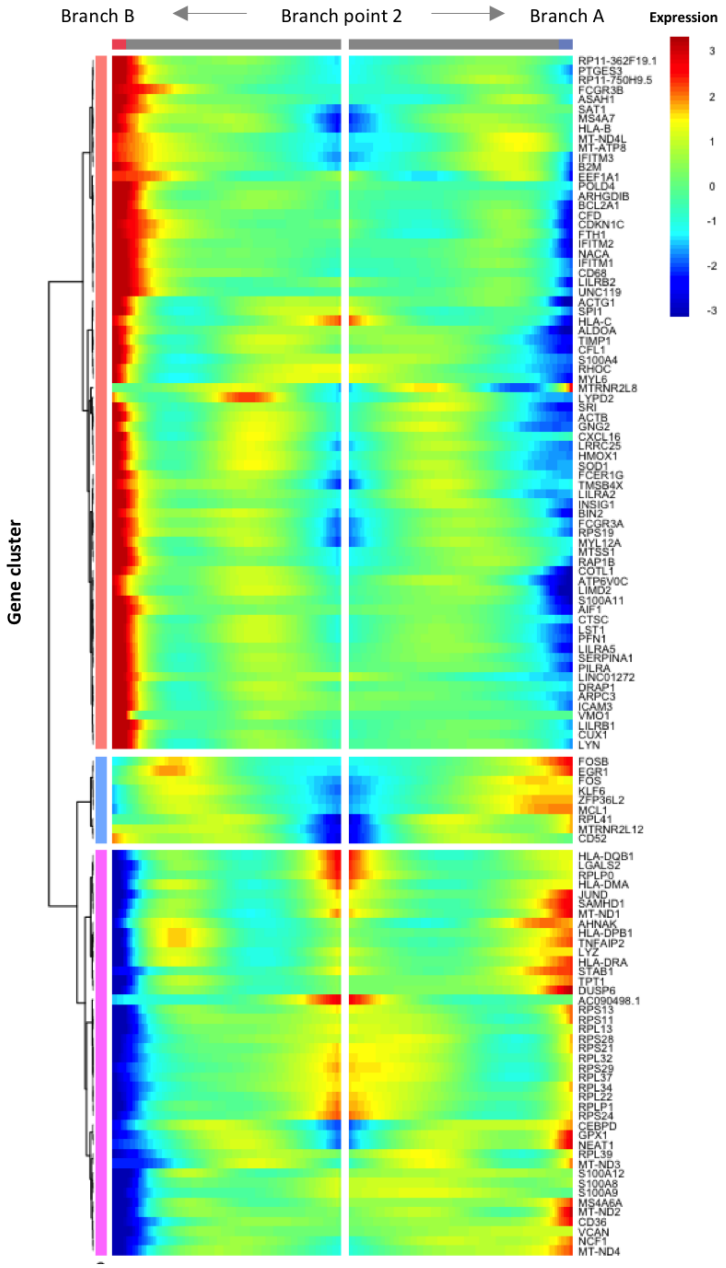


Figure 64: Expression heatmap of branch point 2

Expression heatmap to show most significant genes associated with with branch 2-dependent expression (genes selected for ordering at qval < 1e-9) (see pseudotime trajectory plot in Figure 63A). Scale indicates relative expression.

4.4 Summary and Conclusions

The key findings in this chapter are as follows:

1. Classical monocytes are characterized by transcriptomic profiles associated with cell adhesion, migration, chemotaxis and phagocytosis. $CD16^+$ monocytes exhibit a pro-inflammatory expression signature, with enrichment for antigen-presentation, complement, anti-viral, IFN-γ signaling and T cell costimulation pathways.
2. The monocyte compartment represents a spectrum of differentiation states, with near-linear maturation trajectory from a classical monocyte to a non-classical phenotype. This is associated with distinct gene expression changes including acquisition of non-classical monocyte associated expression profiles, such as antigen-presenting, complement and IFN-γ-signaling pathways.
3. As observed within the bulk RNA-seq dataset in chapter three, monocytes in MM are enriched for transcripts associated with chemotaxis and wound healing. Whilst some cluster-wise transcriptional MM-associated profiles are identified, most of these pathways are induced across classical and $CD16^+$ monocyte populations.
4. ICB is associated with maturation within the classical monocyte compartment in some individuals and acquisition of expression profiles associated with a non-classical phenotype such as antigen presentation, complement production, pro-inflammatory and IFN-γ signaling. This may be reflective of overall polarization towards a non-classical phenotype following ICB.
5. Monocyte expression profiles associated with immunosuppression are downregulated with ICB.
6. scRNA-seq profiling across a large cell population has demonstrated a previously undescribed degree of heterogeneity within the classical monocyte compartment. This is likely to represent both differential maturation states, and activation status.
7. A distinct conventional intermediate monocyte population is not readily identified with high-level unsupervised clustering. This suggests that some of the

observations in the published literature may be attributed to 'noise' or inter-individual variation. Non-classical and intermediate monocyte populations demonstrate a considerable degree of transcriptional concordance.

8. Assessment of gene expression signatures within a large bulk RNA-seq dataset from MM patients demonstrated that at baseline a non-classical monocyte signature is positively associated with OS and PFS in patients treated with ICB.
9. Following ICB a classical monocyte signature comprising classical-associated transcripts with a prominent ribosomal component which might reflect exhaustion or degeneration, may be invserly associated with the development of irAEs

In this chapter I have described novel observations regarding monocyte transcriptomic profiles, differentiation and phenotype. Unsupervised clustering approaches and subsequent cluster annotation have provided in-depth transcriptional phenotyping of conventional monocyte clusters. This approach limits some of the biases incurred with techniques such as FACS sorting that have been widely used in the literature to characterise monocyte subsets. In keeping with other studies, classical monocytes exhibited enrichment of pathways such as chemotaxis, angiogenesis, migration and adhesion, demonstrating their prominent role in the innate and anti-bacterial immune response and ability to migrate to sites of infection, whilst $CD16^{+}$ monocytes exhibited enrichment of type I and type II IFN signaling pathways, complement production and antigen presentation. This is in keeping with a more prominent role in facilitating T cell co-stimulation and adaptive anti-viral immune responses.

Dissection of a large scRNA-seq monocyte dataset has revealed previously undescribed heterogeneity within the classical monocyte compartment which is likely to represent differential maturation states, function and activation. Previous studies which have explored monocyte transcriptional phenotyping with scRNA-seq failed to identify heterogeneity within the classical compartment[104,136], however these studies are limited by very small cell population. The large monocyte cell population I have explored has allowed for high-resolution analysis of this compartment.

I have used trajectory analysis to explore monocyte differentiation, and findings are in keeping with those of previous studies which have used radioactive labelling in vivo to explore maturation profiles across the compartment[434]. Monocytes show a broadly linear differentiation trajectory from classical to CD16$^+$ monocytes. This is supported by correlation between pseudotime (a marker of maturation) and expression of non-classical and classical-associated expression profiling.

Interestingly, in keeping with studies showing overlap between transcriptional profiles of intermediate and non-classical monocytes[90,435] and scRNA-seq studies which have failed to robustly identify a distinct intermediate population[104,133–135], in this chapter I also find a considerable degree of transcriptional similarity between these two cell groups, with no clear discrete intermediate population identified at low-resolution clustering. Intermediate monocytes are defined by intermediate CD16$^+$ expression, but also the highest MHC class II antigen expression[88,93,95,99,100]. In this dataset the CD16$^+$ compartment exhibited the highest MHC class II expression, and high-resolution analysis of this cluster identified a group of cells with high MHC class II expression and comparatively (to other CD16$^+$ monocytes) low *FCGR3A* expression, which therefore may represent a conventional intermediate subset. It may be that the intermediate and non-classical populations should be considered differential activation states of the same pro-inflammatory CD16$^+$ monocyte subset due to shared transcriptional profiles.

scRNA-seq across a large monocyte population has allowed for novel dissection of MM-associated and ICB responses in a subset-wise manner. Notably, the MM-associated profiles reported in chapter three such as upregulation of chemotaxis, adhesion and a pro-thrombotic state were upregulated in both classical and CD16$^+$ monocytes. Strikingly, differential expression analysis and expression module scoring showed that ICB is associated with acquisition of a non-classical profile amongst classical monocytes, including enrichment of genes associated with antigen-presentation, complement production and anti-viral responses. Therefore, the prominent activation of anti-viral response pathways, including T cell costimulation, complement production and antigen presentation described in chapter three may be underpinned by polarization towards a pro-inflammatory non-classical phenotype within the monocyte compartment.

Trajectory analysis demonstrated that this may be underpinned by a maturation response, although only present across some individuals.

In this chapter expression signatures associated with monocyte subsets were associated with clinical outcome. However, the mechanisms underpinning these observations were more challenging to dissect. For example, in chapter three expression of activation marker *MKI67* (found to be more highly expressed in classical monocytes in chapter four) at baseline was negatively associated with irAEs (which are associated with a more favorable response to ICB). In chapter four it is following ICB that a classical-like expression profile characterised by prominent ribosomal gene expression was found to be negatively associated with irAEs. Most strikingly, at baseline a non-classical expression signature was highly significantly predictive of increased OS. This signature might reflect baseline immune recognition. It should also be noted that these associations are derived from cluster-associated expression profiles superimposed in bulk RNA-seq data. Whilst this approach facilitates inferernces regarding possible functional effects of monocyte subsets, comments regarding direct association between subset frequency or direct function are not possible.

The lack of clarity as to the inferences that can be made regarding the functional significance of these associations with clinical outcome may reflect the plasticity and diversity within the monocyte compartment, or may be attributable to sample size. Conventional methods of categorizing monocytes may not be wholly appropriate when exploring clinical association. It may be more useful to consider monocytes in terms of differentiation and activation states (e.g. cycling, pro-inflammatory, immunosuppressive) rather than as discrete subsets. Future functional work to validate these subsets and transcriptional expression profiles, would be of value in exploring these observations further.

In conclusion, the work I present in this chapter utilizes scRNA-seq to provide novel insights into monocyte subsets, function, differentiation and diversity. I describe subset-wise MM-associated and ICB-modulated expression profiles, and novel observations regarding the prominent maturation response following ICB. Building on the work I have presented in chapter three, I demonstrate that myeloid expression profiles have

potential predictive utility in MM patients receiving ICB, and that this may reflect variable activation and functional states.

5 Chapter 5: Characterising circulating epigenetic cell-free DNA profiles in patients with metastatic melanoma treated with immune checkpoint blockade

5.1 Background

Cell-free DNA (cfDNA) is DNA found outside of cells in body fluids including plasma and urine[436]. Although the majority of circulating cfDNA is derived from leukocytes[208], in patients with cancer a proportion of this cfDNA may comprise circulating tumour DNA (ctDNA)[437–439]. ctDNA concentrations are positively correlated with tumour burden[216,440] with elevated ctDNA concentration both at baseline and following treatment predictive of poor prognosis in multiple cancer types[214,215,218], including patients with metastatic melanoma (MM) and non-small cell lung cancer (NSCLC) treated with ICB[441,442]. In addition, ctDNA reflects tumour mutational profiles with good concordance[219,220] and may provide improved representation of tumour mutational profiles compared to core biopsy, overcoming issues associated with non-representative sampling due to tumour heterogeneity[221]. The prospect of assessing ctDNA as a 'liquid biopsy' is therefore an attractive one. Furthermore, analysis of mutational ctDNA profiles has utility for diagnosis[443,444] prediction[445], monitoring on-treatment response[214,220,224,309,445–448] and identifying relapse[217,222,448,449].

Aberrant DNA methylation is characteristic of malignancy[450], with global hypomethylation contributing to tumourigenesis via chromosomal instability, facilitating mitotic recombination and chromosomal rearrangements[451], and through loss of normal imprinting patterns[452,453]. A key tumourigenic driver in many cancers is promotor hypermethylation at CpG sites leading to tumour suppressor gene silencing, as well as altered regulation of genes involved in a wide range of biological process such as cell adhesion, angiogenesis, and DNA repair[241,450]. For example, promotor hypermethylation of *MHL1* in colorectal and endometrial cancers leads to MMR deficiency and MSI, driving tumourigenesis[244,245]. Whilst mutational profiles are often

disparate across cancers of the same tissue type, methylation patterns are typically comparable, providing opportunity for identification of cancer type[454].

5-hydroxymethylcytosine (5hmC) is an epigenetic DNA modification also thought to have a role in malignancy. 5hmC is a permissive marker of gene expression and during the process of DNA demethylation is generated from oxidation of 5mC under the action of the ten-eleven translocation (TET) enzyme family[265], and in turn is successively oxidised prior to base excision repair - thus 5hmC formation is a key step in the process of DNA demethylation. 5hmC also has a role in regulating differentiation in embryonic stem cells (ESCs) and is proposed to be a marker of "transcriptional flexibility"[455], being enriched within gene bodies of transcriptionally active genes[278]. 5hmC profiles are tissue-specific[275,456], and also distinct in cancer patients compared to healthy donors. In cancer, global hypo-5-hydroxymethylation is commonly observed, and has been found to be associated with poor prognosis across multiple cancer types[457–461] including melanoma[462].

Whilst most cancer-associated epigenetic profiling has been tumour-based, cfDNA carries epigenetic modifications, and epigenetic cfDNA profiles have been shown to have utility as a diagnostic and predictive biomarker in cancer and other diseases[296–298,300,304,463]. Song et al.[296] were one of the first groups to describe distinct 5hmC cfDNA profiles between cancer patients and healthy controls. These profiles were also associated with tumour stage[296]. Studies have demonstrated the utility of 5hmC cfDNA profiles as potential diagnostic biomarkers in cancers as diverse as pancreatic[464], NSCLC[301] and colorectal, thyroid, liver and gastric cancers[298]. Recently, 5hmC cfDNA profiles have been shown to have utility as predictive biomarkers. For example, Chen et al.[303] identified a 5hmC cfDNA profile robustly predictive of response to RCHOP treatment in patients with diffuse large B cell lymphoma (DLBCL). Similarly, profiling of 5mC cfDNA has been shown to be a potential diagnostic biomarker in cancers including pancreatic ductal adenocarcinoma (PDAC)[464], intracranial[254], renal[255], NSCLC and colorectal cancers[298,299]. Tissue-origin of cancer can be determined with cfDNA methylation profiling[208],which may be useful in cases of cancer of unknown primary[465], or where tissue biopsy is not feasible. However, to date, descriptions of 5hmC and 5mC

cfDNA profiles within melanoma cohorts are limited and there has been little work regarding on-treatment modulation of epigenetic cfDNA profiles.

As a large proportion of cfDNA originates from leukocytes, epigenetic cfDNA profiling is likely to represent immune-based signatures. Epigenetic modification is a key way in which innate and adaptive immune responses are regulated. This can occur via direct DNA methylation, chromatin remodelling and post-translational histone modification[466,467]. Methylation plays a role in regulation of immune cell differentiation[236] and immune function[236,466]. TET enzymes and 5hmC labelling have also been shown to have a role in regulating the development of T cells, B cells, and in myeloid differentiation[287,288,468]. Assessment of 5mc and 5hmC enrichment may therefore reflect immune activation states.

In the context of ICB, epigenetic profiles have been shown to modulate immune response and facilitate treatment resistance. For example, trimethylation of enhancer of zeste homologue 2 (EZH2) H3K27m3 is associated with inhibition of *CXCL9* and *CXCL10* expression, and reduced intra-tumoural T cell infiltration, thus facilitating immune evasion[469]. Peng et al.[469] showed that modulation of methylation to abrogate these effect led to increased T cell infiltration and enhanced the effects of anti-PD-1 therapy in mice. Global hypomethylation in melanoma is predictive of response to anti-PD-1 therapy[470]. Jung et al.[470] found that hypomethylation in melanoma was associated with immune evasion and a more robust predictor of response to immunotherapy than tumour mutational burden (TMB) in patients. It is evident therefore that epigenetic tumour profiles may have a role as predictive biomarkers in the context of ICB.

As described in Chapter Three and Four, ICB is associated with modulation of transcriptomic profiles in circulating monocytes , as well as $CD8^+$ T cells[71]. As cfDNA is likely to represent both leukocyte and tumour-derived DNA, epigenetic cfDNA profiling may be a non-invasive way to characterise tumour-based and immune responses to ICB. To date, there is no previously published work that explores on-treatment profiling of epigenetic cfDNA profiles in the context of ICB, nor that dissects its predictive utility. Such observations may provide insights into peripheral immune or tumour profiles that may have utility as potential prognostic or predictive biomarkers.

In this chapter, I describe how I used 5hmC and 5mC cfDNA sequencing to explore 5hmC and 5mC cfDNA profiles in patients with MM, to characterise on-treatment modulation and to identify epigenetic cfDNA profiles predictive of clinical response.

5.2 Aims and Objectives

The aims of this work were to:

- Characterise 5hmC and 5mC profiles in cfDNA
- Characterise epigenetic cfDNA profiles in patients with MM
- Dissect on-treatment modulation of epigenetic cfDNA profiles following ICB
- Explore epigenetic cfDNA profiles associated with clinical outcome in patients with MM treated with ICB
- Explore whether epigenetic cfDNA profiles have transcriptional correlates in circulating immune cells

5.3 Results

To explore 5hmc and 5mC cfDNA profiles, 5hmC and 5mC-enriched libraries were prepared from cfDNA derived from patients with MM receiving ICB before ('C1') and after ('C2') treatment with ICB, as well as healthy donor/control samples ('C'/'HD') (Appendix CC for 5hmC cohort details, Appendix EE for 5mC cohort details).

cfDNA 5hmC cfDNA libraries were prepared using a modified hMe-Seal method[471] to allow for low input quantities of cfDNA, and were 75bp paired-end sequenced on a NextSeq500. Samples were sequenced to a median depth of 17,106,550 reads per sample. Two non 5hmC-enriched samples ('IC') comprising pooled cfDNA from four patients were also sequenced (Appendix DD). Efficacy of 5hmC pull-down and enrichment was confirmed after sequencing by determining the proportion of 5hmC-enriched amplicon spike-in oligos in raw FASTQ files, compared to methylated and unmodified oligo spike-in read sequences (see Methods section 2.7.2.1. and Appendix CC for detailed of enrichment). Samples were taken forward for peak calling with the *MACS2* package if the enrichment of 5hmC spike-in oligos were at least two-fold greater than that of both 5mC and control (unmodified) amplicons (Appendix CC).

5mC libraries were constructed using 5mC immunoprecipitation ('cfMeDIP') as previously described[261], with the addition of bacterial spike-in oligos to assess 5mC enrichment efficacy assess with qPCR. All samples taken forward for sequencing had at least an 87% 5mC pull-down efficacy (cohort details in Appendix EE). 5mC libraries were 75bp paired-end sequenced on a NextSeq500 platform. 24 paired pre- and post ICB treatment 5mC samples were sequenced across the three batches, as well as eight healthy control samples (HD/C). Two non-enriched ('IC') samples were also sequenced, each comprising four pooled non 5mC-enriched cfDNA samples. Samples were sequenced to a median depth of 35,303,187 reads per sample.

The *MACS2* package[347] was used to identify enriched 5hmC/5mC peaks (broad peak calling, band width 300, q-value cut-off 5.00e-02), with peaks called using each individual enriched sample as the 'ChipSeq input file' and non-enriched pooled samples as the 'background control'. Samples were taken forward for downstream analysis if at least 1000 significantly enriched peaks were identified. Following this quality control (QC), a total of 13 paired pre- and post ICB-treatment 5hmC samples were available for downstream analysis, and 5 healthy donor samples. A further 10 unpaired C1 samples and 9 C2 samples met QC parameters and were included in downstream analysis. The 5hmC samples that passed QC (in total n=23 C1, n=22 C2 and n=5 HD samples) exhibited between 9,922 and 216,559 peaks per sample (median 150,320 peaks), whilst 5mC samples that passed QC (a total of n=24 C1, n=24 C2 and n=8 HD samples) exhibited between 7,516 and 185,181 peaks per sample (median 54,748 peaks).

5.3.1 Genome-wide distribution of epigenetic cfDNA profiles

To explore the enrichment profiles of 5hmC and 5mC in cfDNA across this MM dataset, the enrichment across chromosomal regions and distribution across the gene body was explored.

To visualise the distribution of 5hmC and 5mC enrichment across chromosomal regions, integrated peaksets were created by merging per-sample peaksets called by *MACS2* and then removing overlaps/duplicates using the *Granges*[472] *reduce* function. These merged peak sets, representing all 5hmC or 5mC enriched regions across the genome, were plotted across chromosomal regions using the *ChIPSeeker* function *covplot* to create an

epigenetic map of 5hmC and 5mC enriched sites. 5hmC and 5mC modifications were found across all chromosomes (Figure 65).

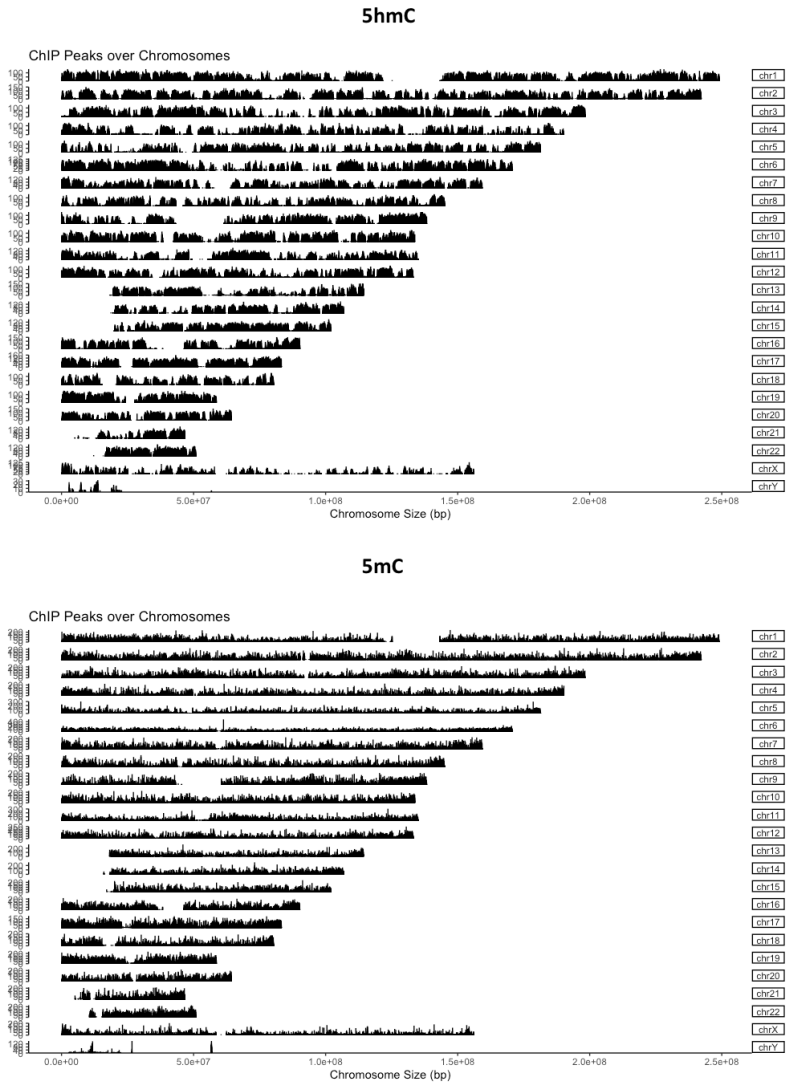


Figure 65: Plot to show distribution of 5hmC and 5mC peaks
Plot to show 5hmC (409,984 peaks) and 5mC (446,753 peaks) peak set coverage across chromosomes. Peaksets were derived from peaksets from all C1 and C2 patient and healthy donor samples merged.

5hmC is typically enriched within gene bodies and depleted within transcriptional start sites (TSS)[253,273,473]. 5mC usually exhibits depletion at the TSS with increased enrichment across the gene body and reduction in enrichment at the transcriptional end site (TES)[253,474]. To explore the distribution of 5hmC and 5mC enrichment across the gene body within cfDNA in this dataset, per-sample reads were merged and reads from all

samples were plotted across the gene body (plus a 20kb flanking region) using the *ngs.plot* package[348]. Visual inspection of tag 5hmC profiles demonstrated prominent enrichment just upstream of the transcriptional start site (TSS) in an area consistent with the promotor region, with depletion at the TSS and enrichment within the gene body (Figure 66). This is consistent with previously described 5hmC profiles[471,473]. There was no variation in the pattern of distribution across control, C1 and C2 samples on visual inspection.

5mC distribution across the gene body demonstrated depletion at the TSS and gradual accumulation across the gene body towards the TES, with depletion at the TES (Figure 66). There was no change in the pattern of distribution across treatment and between MM (C1 and C2) and healthy donors (HD). These patterns are consistent with previously described patterns of 5hmC/5mC modification across the genome and validate the data.

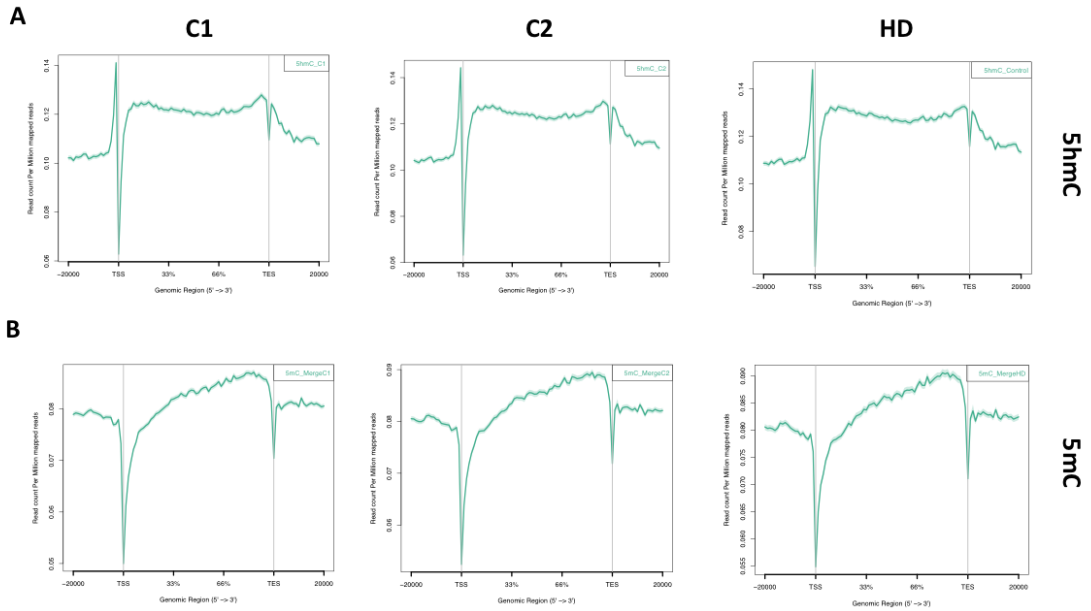


Figure 66: Metagene profiles across gene bodies to show 5hmC and 5mC distribution
Tag gene profiles showing distribution of read coverage (read count per million mapped reads) across gene bodies (from TSS to TES plus a 20kb flanking region) for all C1, C2 and healthy donor (HD) samples. Reads across gene bodies plotted using ngsplot across A) all 5hmC samples.; B) all 5mC samples.

5.3.2 Comparing the genome-wide distribution of 5hmC and 5mC modifications within cfDNA

5hmC and 5mC demonstrate distinct patterns of enrichment across functional genomic regions[278]. To explore the global distribution of epigenetic modifications within cfDNA, the 5hmC/5mC peaksets were merged to identify an integrated peakset of epigenetically enriched genomic regions derived from all patient and healthy donor samples. A 'peak' is a genomic region where 5hmC or 5mC read pileup has been found to be significantly greater than that identified in a non-enriched background sample as identified by *MACS2*. No threshold of minimum contribution per each individual was specified. The *GenomicRanges reduce* function was used to remove overlapping peaks and the *BRGenomics*[349] function *tidyChromosomes* used to remove non-standard chromosomes. A total of 409,984 discrete 5hmC peaks were identified across all samples, and 446,753 peaks found across all 5mC samples.

To determine regions of dual 5hmC and 5mC labelling, overlapping regions between the integrated 5hmC and 5mC peaksets was identified. There were 170,579 peaks that

intersected between both 5hmC and 5mC datasets (Figure 67B). Overlap analysis using the *overlapPermTest* function from *regioneR* to perform a permutation test demonstrated significant overlap between the 5mC and 5hmC peaksets (n=20 permutations, z score = 332.736, $p<0.05$), demonstrating shared sites of epigenetic modification. There were however, genomic regions with 5hmC or 5mC-wise enrichment, with 214,836 5hmC peaks not overlapping 5mC enriched regions, and 195,838 regions of 5mC enrichment without overlapping 5hmC enrichment. Next, annotation of these peaksets according to associated genic region was performed using *ChipSeeker*[350]. Genic annotation was assigned according to the nearest genic TSS (-3kb to +3kb), as is standard for most annotation packages[475]. 5hmC or 5mC peaks within intergenic or downstream regions were removed from this analysis. A total of 23,953 genes contained 5hmC peaks, and 24,887 genes contained 5mC peaks. The majority of genes demonstrated both 5hmC and 5mC labelling, with 23,066 genes showing dual 5hmC or 5mC enrichment, 887 genes showing sole 5hmC labelling, and 1,821 showing 5mC-wise labelling (Figure 67A). This demonstrates that whilst most genes exhibited both 5hmC and 5mC modifications with many locations exhibiting dual 5hmC/5mC labelling, many 5hmC and 5mC modifications are discrete to a genomic location. This may be anticipated given the relationship between 5mC and 5hmC.

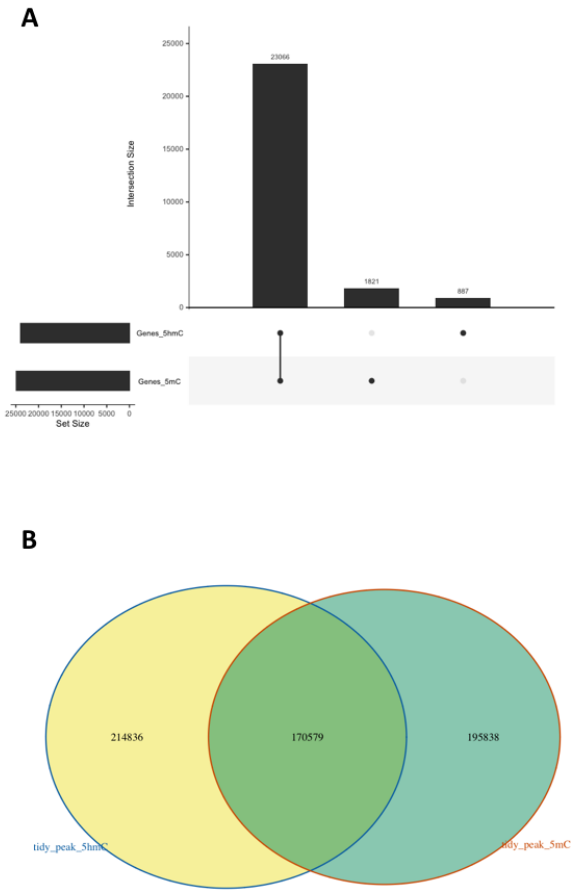


Figure 67: Overlap of 5hmC and 5mC enriched regions
A) Upset plot to show overlap between 5hmC-labelled genes (n=23,953 genes) and 5mC-labelled genes (n=24,887 genes). B) Venn diagram to show overlap between 5hmC (409,989 peaks) and 5mC (446,753 peaks) peaksets (170,583 overlapping peaks) performed with the ChipPeakAnnno function findOverlapsOfPeaks. Peaksets were derived from merging all called peaks by MACS2 per each sample (across C1, C2 and HD samples), and removing those peaks that overlap.

To explore the distinct biological pathways associated with 5hmC and 5mC modification, pathway enrichment analysis was performed across 5hmC and 5mC-labelled genes. 5hmC and 5mC peaks were annotated with the *ChipSeeker*[350] package according to the UCSC GRCh38 database[476]. The *annotatePeak* function was used to annotate peaks with functional regions and associated genes, with peaks annotated in order of priority (promoter, 5'UTR, 3'UTR, exon, intron, downstream and intergenic). Peaks within intergenic or downstream (within 3kb of the 3' area) regions were removed from the analysis.

Pathway enrichment analysis of 5hmC or 5mC-labelled gene sets was performed using *XGR*[326] interrogation of the GOBP database. A total of 66 GOBP pathways were enriched across 5hmC-enriched genes, and 27 pathways across 5mC enriched genes (P_{adj} <0.05) (Figure 68). Both gene sets showed highly significant enrichment for GTPase activity (GO:0043547), protein phosphorylation (GO:0006468) signal transduction (GO:0007165, GO:0035556) and MAPK signalling (GO:0000165) (Figure 68), demonstrating shared enrichment across key biological processes. Pathways showing 5hmC-wise enrichment included DNA repair (GO:0006281), apoptosis (GO:0006915), cell proliferation (GO:0008283), neutrophil degranulation (GO:0043312) and response to hypoxia (GO:0001666, GO:0061418) (Figure 68). GOBP pathways showing 5mC-wise enrichment include potassium ion transmembrane transport (GO:0071805), cell adhesion (GO:0007155) cytoskeletal organisation (GO:0007010) and tyrosine kinase signalling (GO:0007169) (Figure 68).

Whilst only 0.1-0.7% of the genome is 5-hydroxymethylated[271,477], this demonstrates that this modification is distributed across nearly all genic regions and that 5hmC and 5mC are likely to play a role in a diverse but critical range cellular processes including cell signalling, metabolism, proliferation and cell death.

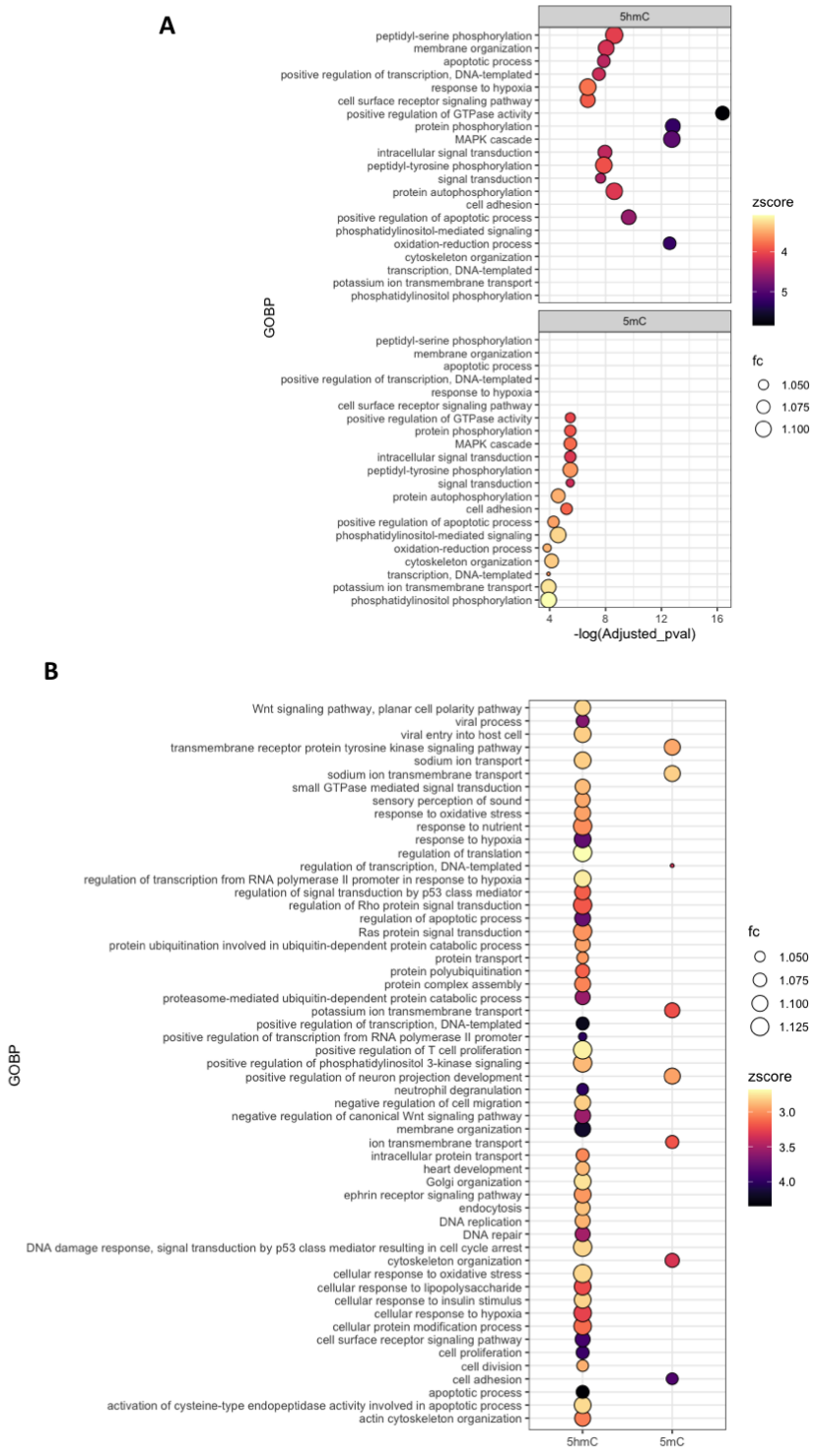


Figure 68: Pathway enrichment analysis of 5hmC and 5mC enriched genes

A) Dotplot to show top 15 most significant GOBP pathway enriched across 5hmC and 5mC enriched genes with XGR interrogation of the GOBP database (P_{adj} <0.05). B) Dotplot to show pathways enriched in only 5hmC or 5mC annotated gene sets (P_{adj}<0.05).

5.3.3 Enrichment of epigenetic modifications across functional regions in cfDNA

Next, the distribution of epigenetic enrichment across functional genomic regions within this cfDNA dataset was explored. Permutation analysis was run using the *regioneR* package to characterise enrichment of 5mC or 5hmC modification across genomic function regions (e.g. exons/introns). As previously described in section 5.3.2, peak sets for each 5hmC and 5mC per individual were merged to create integrated 5hmC and 5mC peak sets consisting of all epigenetically modified regions across conditions, with overlapping peaks removed using the *reduced* function in *GRanges*. Non-standard chromosomes were removed from analysis using the *BRGenomics*[349] function *tidyChromosomes*. The integrated peaksets comprised 409,984 5hmC peaks and 446,753 5mC peaks. The package *regioneR*[351] was then used to create a randomly generated set of genomic peaks comparable to the size of the 5hmC/5mC peakset. Following this, comparison of overlap between the patient-derived peakset and randomly-generated peakset across defined functional genomic regions according to UCSC hg38 genome build annotations retrieved using the *annotatr* package was carried out[353]. Permutation analysis was run using the *permTest RegioneR* function with 20 iterations (p-value <0.05).

5hmC was enriched across all intragenic genomic regions, most prominently across introns, but also within exonic regions, promoters and sites 1 to 5kb upstream (20 iterations, p<0.05) (Figure 69A). 5hmC was depleted within intergenic regions (Figure 69A). This is consistent with both the previously described profiles of these data (section 5.3.1), and with previous studies demonstrating prominent intragenic enrichment and intergenic depletion of 5hmC[456]. The distribution of 5mC similarly demonstrated enrichment across intragenic regions with the greatest enrichment seen across exonic regions, and depletion within intergenic regions (Figure 69A).

Methylation at CpG sites is a key way in which gene expression is regulated. CpG islands (CGI) are genomic regions containing high levels of GC dinucleotide sequences (at least 50%[225]). CGIs commonly occur within promoter regions, with approximately 70% of promoters containing a CGI[478]. Overall 70-80% of CpG sites exist in a methylated state[227]. Within promoter regions CGIs are commonly unmethylated, whilst those within

intragenic regions are typically methylated[227]. Methylation of CGIs within promoter regions regulate chromatin structure, accessibility and are a key way in which gene expression is regulated[479]. In addition, regions directly flanking CGI sites; CpG shores (2kb adjacent to CGIs) and CpG shelves (2kb adjacent to CpG shores) exhibit more dynamic methylation patterns than within CGIs[480,481] and it is thought that variation in CpG shore methylation profiles are reflective of tissue of origin[480,481]. 5hmC enrichment at CGIs associated with transcriptional modulatory function[482] as well as long interspersed nuclear element (LINE) elements has been reported[483], however 5hmC enrichment is usually lower than 5mC enrichment at these regions[483].

To explore the distribution of 5hmC and 5mC across CpG sites in cfDNA, genomic CpG sites were retrieved from the UCSC hg38 genome with *annotatr*[353] and permutation analysis was run as previously described using *regioneR*. Both 5mC and 5hmC had a similar distribution across CpG sites, and were enriched at CpG sites, but more prominently within CpG shelves and shores with depletion at inter-CpG sites (Figure 69B).

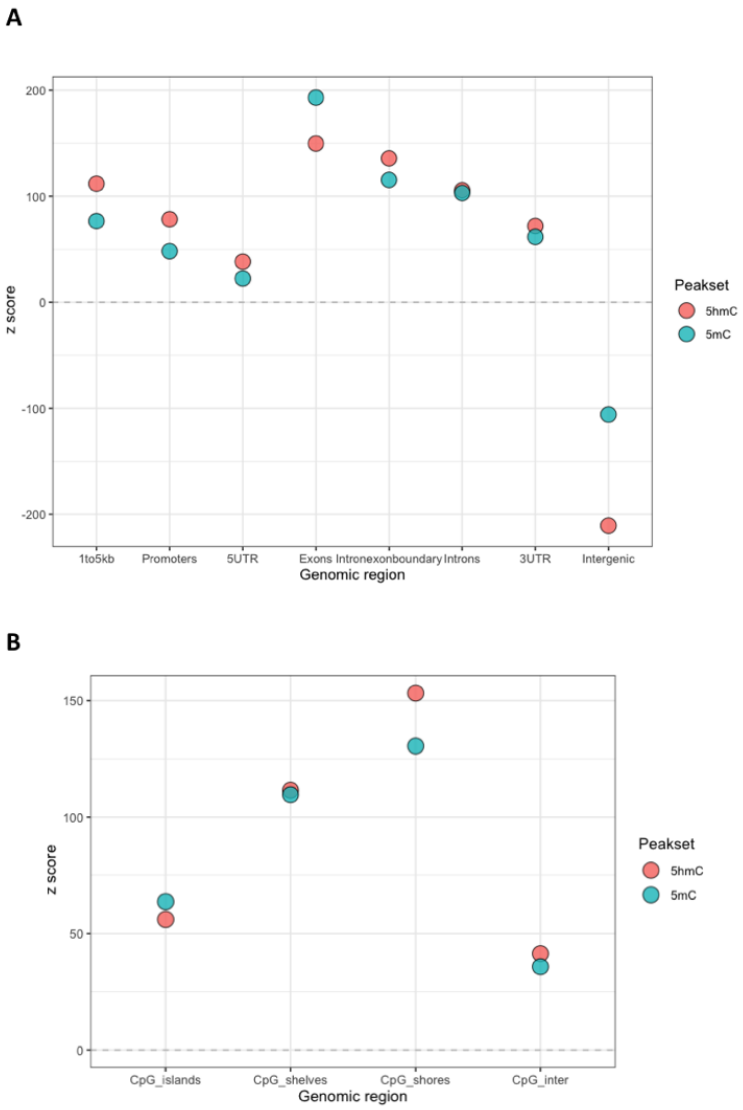


Figure 69: Enrichment of 5hmC and 5mC across functional genomic regions
Plots to show enrichment of 5hmC and 5mC across functional genomic regions. Peak sets were derived from peaks called by MACS2 per sample then merged together (n= 440,9984 5hmC peaks and n= 446,753 5mC peaks). Genomic region locations were obtained from the UCSC hg38 database. Permutation analysis was run using the regioneR package with 20 permutations (p<0.05). A) Enrichment of 5hmC and 5mC across functional genomic regions (including 1to5kb (1 to 5kb upstream of the TSS), promoters, 5'UTR, exons, intronexonboundary, introns, 3'UTR and intergenic regions). B) Enrichment of 5hmC and 5mC across CpG islands, CpG shores (region 2kb adjacent to CpG island) and CpG shelves (region 2kb adjacent to CpG shore), and inter-CpG sites (genomic regions outside of CpG islands, shores or shelves).

To facilitate direct comparison of 5hmC and 5mC enrichment across functional genomic regions, enrichment across functional regions using Fisher's exact test was performed. Overlap of 5mC or 5hmC peaksets across a genomic region of interest (e.g. promoter or

intron) was identified using the *GRanges findOverlaps* function. Genomic coordinates denoting functional genomic regions from hg38 were obtained with *annotatr*. Comparison of the overlap of the 5hmC or 5mC peaksets across the defined functional genomic regions was then assessed and the odds ratio (OR) indicating relative enrichment at these sites calculated using Fisher's exact test.

Compared to 5mC, 5hmC was significantly more enriched across exons, promoters, enhancers, 5'UTR and 3'UTR regions, and more depleted across intergenic regions, as well as intronic and 1to5kb regions upstream of the TSS (Figure 70). 5hmC was also significantly depleted at CpG shores and shelves, as well as at inter-CpG regions, in comparison to 5mC (Figure 70).

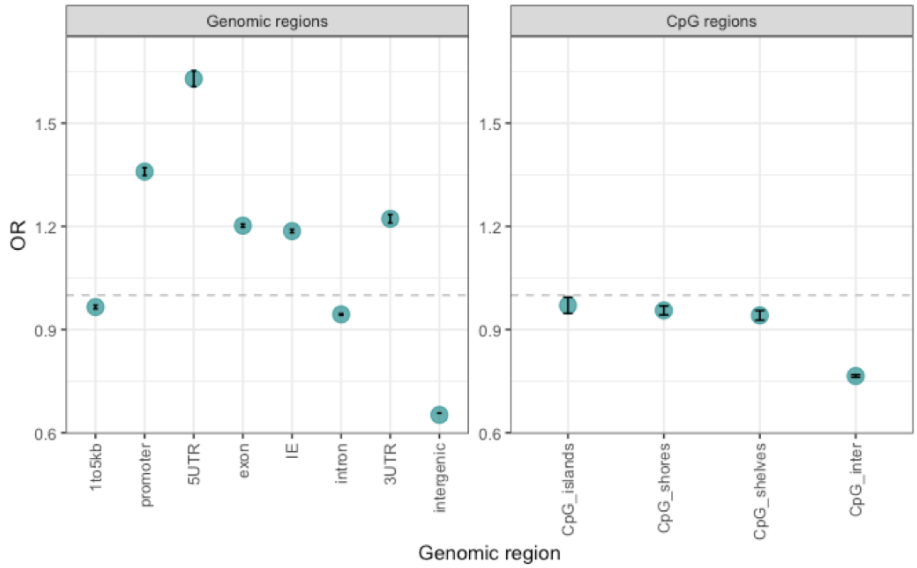


Figure 70: Enrichment of 5hmC versus 5mC
Figure to show odds ratio (OR) indicating relative enrichment of the 5hmC peakset (409,984 peaks, derived from all patient and healthy donor samples) compared to the 5mC peakset (44,6753 peaks, derived from merged peaksets for all patients and healthy donors samples) across functional regions (1 to 5 kb upstream, promoters, 5' UTR regions, exons, intron/exon boundaries 'IE', introns, 3' UTR regions and intergenic regions) and CpG sites (including CpG islands, CpG shores, CpG shelves and inter-CpG sites) (Fisher's exact test). Whiskers indicate 95% confidence interval.

5.3.4 Enrichment of epigenetic cfDNA modifications across functional regions in metastatic melanoma and with ICB

I then explored whether 5hmC and 5mC enrichment across functional genomic regions were distinct in in MM patients versus healthy donors, and modulated in response to ICB.

Firstly, I explored whether MM and ICB was associated with distinct patterns of 5hmC and 5mC enrichment across functional regions. Integrated peaks sets were created by merging all pre-treatment (C1), post-treatment (C2) and healthy donor (HD) peaksets as previously described in section 5.3.3. Fisher's exact test was used to calculate the OR of overlap of the epigenetic peakset of interest (i.e. 5hmC or 5mC) across functional genomic regions and CpG sites in MM patients versus healthy donors, and followinng ICB.

Distinct enrichment profiles of 5mC and 5hmC across functional regions were seen in MM patients versus healthy donors. In MM patients, 5mC was depleted at promoter regions and CGIs, CpG shelves and shores, and enriched at intergenic regions (Figure 71A). Despite overall depletion at intergenic regions, 5hmC was more enriched within intergenic regions in MM patients, whilst relative depletion was seen within promoter regions and introns (Figure 71B).

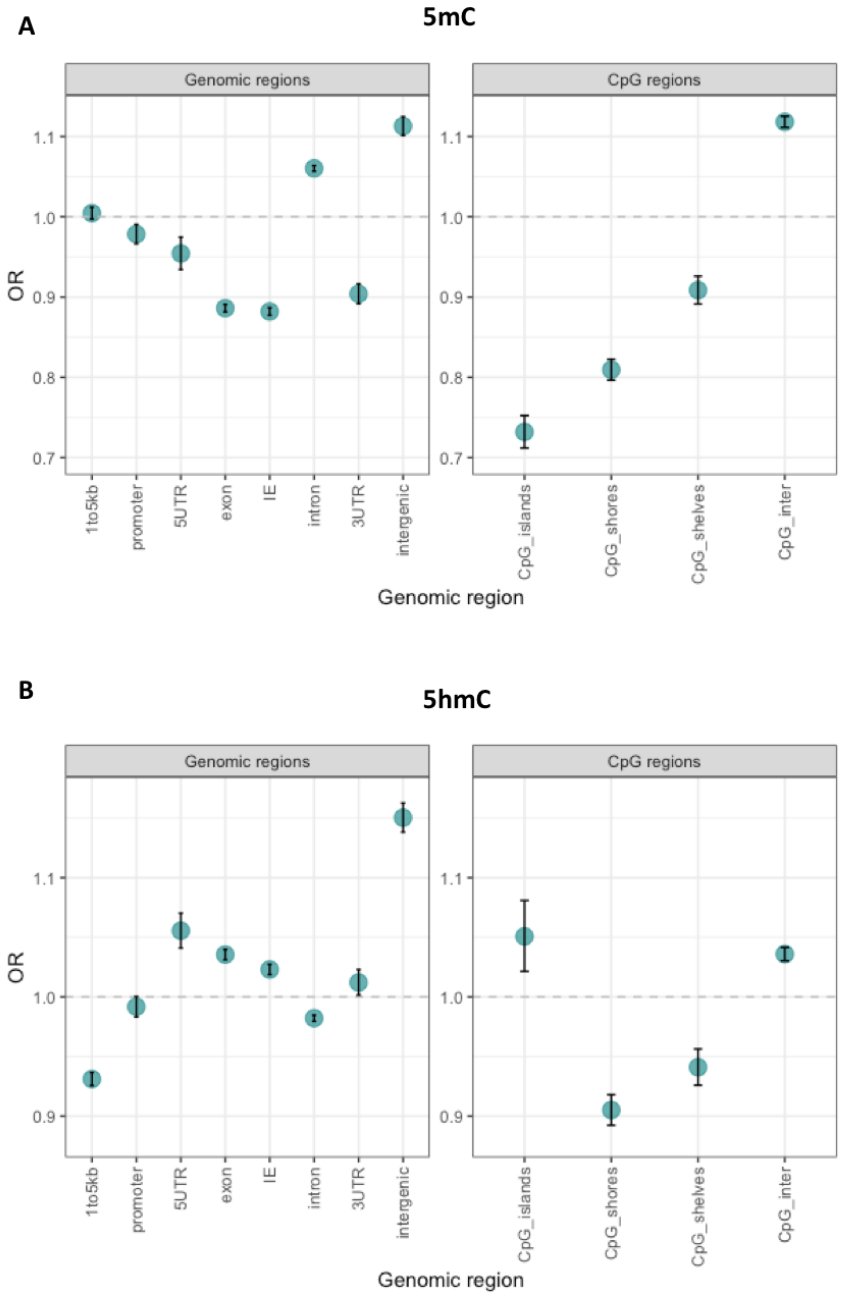


Figure 71: Enrichment of 5mC and 5hmC in cfDNA in MM
Odds ratio ('OR') to show relative enrichment of: A) 5mC (359,588 C1 peaks, 192,942 HD peaks) and; B) 5hmC (373,933 C1 peaks, 274,639 HD peaks); across functional regions obtained from the UCSC hg38 build in patients with MM at baseline versus healthy donors. Whiskers indicate 95% confidence interval (Fisher's exact test). OR with 95% confidence interval not overlapping 1 (indicated by the dotted line) are significant.

Next, the change in enrichment of 5mC and 5hmC across functional regions following ICB was characterised. Notably, there was significant reduction in promoter and CGI

methylation with ICB treatment (Figure 72A). Depletion of 5mC at these gene regulatory sites which are typically associated with transcriptional silencing, suggest capture of DNA from activated cells. ICB is associated with increased proliferation and T cell apoptosis[484], so these epigenetic profiles may reflect preferential death of activated circulating immune populations following ICB.

Following ICB treatment, 5hmC was more enriched across most genic regions including exonic regions, 5' UTRs, 3' UTRs and at intron/exon boundaries, and significantly depleted across intergenic regions (Figure 72B). Given 5hmC is a permissive marker of gene expression typically found across gene bodies, enrichment of 5hmC across intragenic sites may reflect transcriptional activation and induction of gene expression following ICB. This is suggestive of capture of immune cell derived cfDNA. In addition, the depletion of intergenic 5hmC suggests that more specifically activated cells have been captured after treatment.

Overall, the distribution of both 5mC and 5hmC enrichment in cfDNA across functional genomic regions is in keeping with increased genic expression following ICB and may reflect dynamic changes in immune responses.

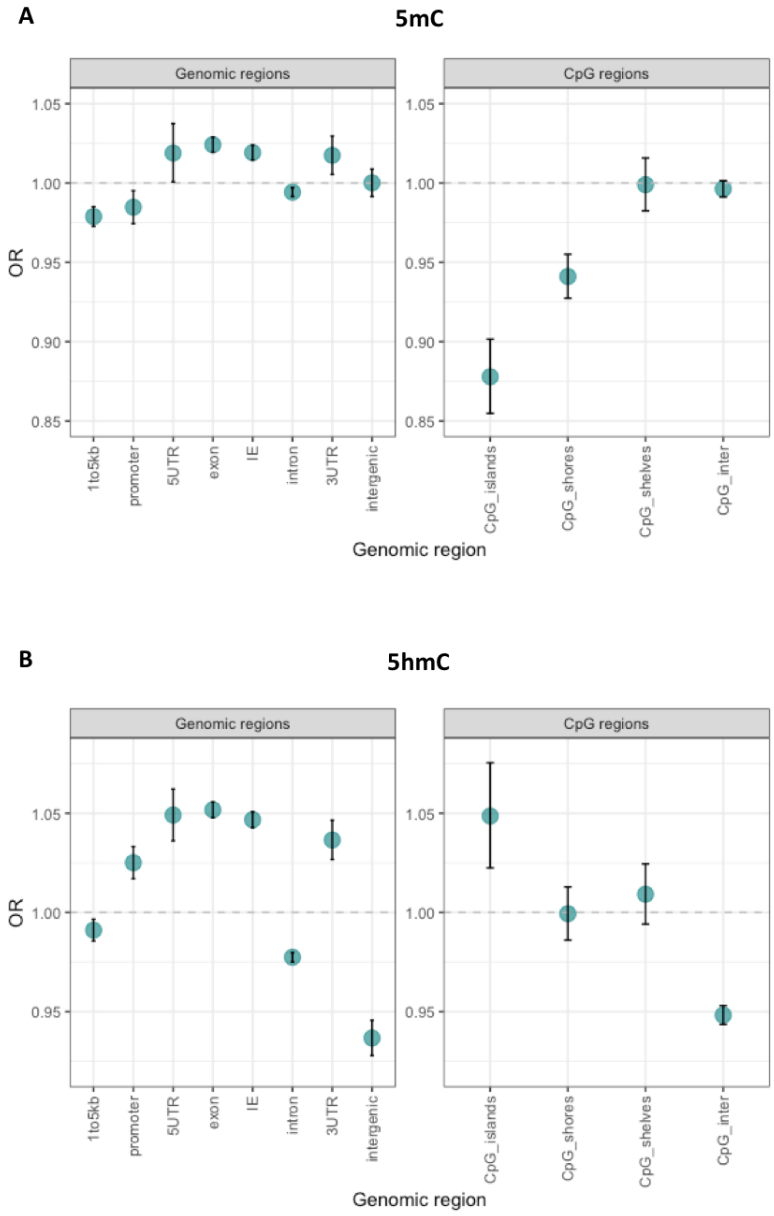


Figure 72: Enrichment of 5mC and 5hmC across functional regions with ICB
Odds ratio (OR) to show relative enrichment of A) 5mC (C1 359,588 peaks, C2 372,869 peaks) and B) 5hmC (373,933 C1 peaks, 360,602 C2 peaks) across functional regions and CpG sites following ICB. OR was calculated using Fisher's exact test. Whiskers indicate 95% confidence interval.

5.3.5 Epigenetic modifications in cfDNA at histone modification sites

A key mechanism by which gene expression is regulated is via epigenetic post-translational histone modification[485]. The ways in which histones can be modified are diverse, and include acetylation, phosphorylation and methylation[485]. Such

modifications affect gene expression, both by affecting the interaction between DNA and chromatin, therefore affecting DNA accessibility and transcriptional activity; but also by directly binding with proteins such as transcription factors and associated enzymes[485,486]. Modifications such as H3K4me3 and H3K36me3 are associated with transcriptional activation, whilst others including H3K9me3 and H3K27me3, with transcriptional repression[487] (Figure 73). Hydroxymethylation at these sites are also a key way in which chromatin is regulated. 5hmC is associated with euchromatin and markers of transcriptional activation such as histone modification H3K4me2/3[488,489]. It has been postulated that 5hmC has a permissive effect on gene expression via increasing chromatin access, or preventing inhibitory transcription factor binding[490].

Histone modification	Association
H3K27ac	Active enhancers and TSSs[491]
H3K27me3	Inactive chromatin[492]
H3K36me3	Exons[493,494] and gene bodies within transcribed genes[487]
H3K4me1	Enhancer regions[495], typically downstream[487]
H3K4me3	Active chromatin[487]
H3K9me3	Heterochromatin, inactive regions[487]

Figure 73: Table to show key histone modifications and functional associations

To explore enrichment of epigenetic modifications in cfDNA across chromatin modifications a permutation analysis approach was used with *regioneR.* Histone modification sites (broad peak calling files) from the *Roadmap Epigenomics* project[352] from peripheral T cells ($CD3^+$ cells, cell type ID E034) and monocytes (cell type ID E029) were obtained using the *AnnotationHub* package in R. The *liftOver* function was used to convert genomic ranges of histone modification into the appropriate genome build (from hg19 to hg38). *regioneR* was used to run a permutation analysis to determine enrichment of the 5hmC/5mC peaksets across histone modification sites.

The histone modification H3K27ac is associated with transcriptional activation and is a marker of active enhancer sites[491]. By comparing the 5hmC modified DNA with T cell and monocyte chromatin modifications, 5hmC enrichment was found across H3K27ac-

modified regions (Figure 74). H3K36me3 is also a marker of active transcription[493], particularly within exons[494], and has been shown to have a role in preventing transcription from cryptic start sites[493]. 5hmC is typically associated with H3K36me3 sites[468]. In this dataset, 5hmC was highly enriched across H3K36me3 sites in both monocytes and T cells, in keeping with a marker of transcriptional activation (Figure 74). H3K4me1 is associated with enhancer regions[495] and H3K4me3 is with transcriptional activity[487]. Enrichment of 5hmC was found across both of these markers of gene activation in both T cells and monocytes (Figure 74).

H3K27me3 is associated with heterochromatin and transcriptional repression[492]. It is found at low levels in cancers including ovarian, breast and pancreatic cancer, with utility as a prognostic indicator, and depletion associated with reduced OS[496]. Despite 5hmC typically a marker of transcriptional activity, 5hmC and H3K27me3 are strongly correlated[497], and both are found at high levels in terminally differentiated cells, with depletion in immature or embryonic cells[497–499]. H3K27me3 is positively correlated with global methylation at bivalent promotors, and negatively correlated with global methylation at silent promotors[500]. Both 5hmC and 5mC were enriched across this modification (Figure 74). H3K9me3 is also associated with heterochromatin, which is tightly packed and transcriptionally inactive[493]. 5hmC demonstrated significant depletion across H3K9me3 sites, whilst 5mC showed enrichment in monocytes and no significant enrichment in T cells (Figure 74).

In summary, 5hmC was prominently enriched across markers of active chromatin and transcription and active gene expression including H3K27ac, H3K9me3, H3K36me3 and H3K4me3, and depleted across markers of transcriptionally inactive chromatin including H3K9me3. This is in keeping with the previously described association between epigenetic modifications across histone modification sites[488–490]. Broadly, there was similarity between enrichment patterns of 5mC and 5mC across histone modifications of both T cell and monocyte cell types. Enrichment of 5hmC across both T cell and monocyte histone modification sites associated with transcriptional activation suggests that both T cells and monocytes form a significant ‘donor’ cell to cfDNA.

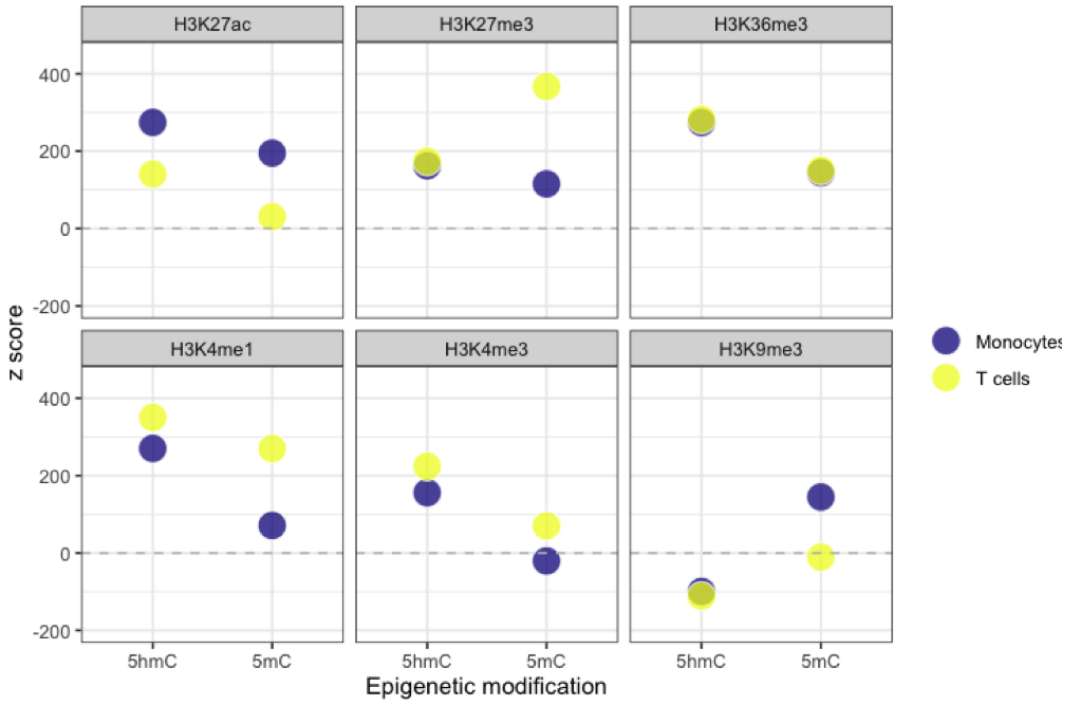


Figure 74: Enrichment of 5hmC and 5mC across histone modifications
Enrichment of 5hmC and 5mC across histone modification sites obtained from the RoadmapEpigenomics project from monocytes (cell ID E029) and peripheral T cells (cell ID E034). 5hmC (409,984 peaks) and 5mC peaksets (446,753 peaks) were derived from all samples merged across the cohort. Enrichment analysis was run using RegioneR permutation analysis (20 iterations, p-value <0.05).

To allow for direct comparison between 5hmC and 5mC enrichment across histone modifications, OR were calculated using Fisher's exact test comparing overlap of peaksets across histone modification sites in peripheral T cells (sample E034) and monocytes (sample E029) obtained from the *EpigenomeRoadmap* project. 5hmC was found to be more enriched than 5mC across histone modifications associated with transcriptional activation and permissive gene expression, such as H3K27ac and H3K4me3, and less enriched than 5mC across histone modifications associated with inactive chromatin and transcriptional repression such as H3K27me3 and H3K9me3 (Figure 75). Broadly, histone modifications showed the same pattern of enrichment across histone modification sites derived from monocytes and T cells. However, there was some variation and it may be that greater 5hmC enrichment across markers of active transcription might infer cell type of origin. For example, 5hmC was more

enriched across H3K4me1 and H3K4me3 in monocytes versus T cell, markers of enhancer regions and active chromatin respectively. This suggests that the cfDNA captured here may be predominantly derived from circulating monocyte populations. However, it is noted that 5hmC is more enriched in T cells across marker of active enhancers H3K27ac, so overall predominant cell-of-origin cannot be clearly inferred.

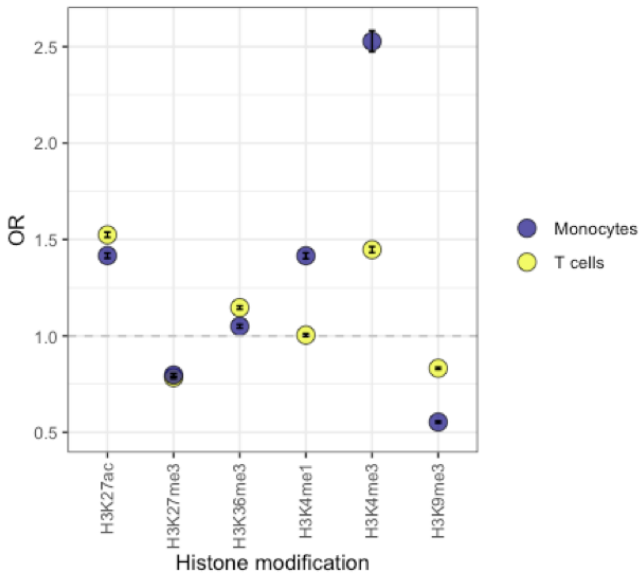


Figure 75: Odds ratios of 5hmC enrichment compared to 5mC cfDNA enrichment at histone modification sites
Plot to show odds ratio (OR) of overlap of 5hmC (409,984 peaks) peaks across histone modification sites in monocytes and T cells versus overlap of 5mC peaks (446,753 peaks). Whiskers indicate 95% confidence intervals.

To explore whether enrichment of 5hmC and 5mC across histone modifications sites was distinct in MM patients versus healthy donors, OR were calculated to compare overlap of peaksets in untreated MM samples and healthy donor peaksets. 5hmC showed greater enrichment in MM patients across histone modification sites associated with transcriptional repression, such as H3K27me3 and H3K9me3, with depletion at histone modifications associated with permissive gene expression such as H3K36me3 and H3K4me1 (Figure 76A). In MM patients, 5mC was more enriched across H3K27ac, a marker of active enhancers Figure 76A). Together, this may reflect of a state of overall repression of immune cell expression in MM.

Next, the enrichment of 5hmC and 5mC in cfDNA across histone modification sites following ICB was characterised using the same approach. Following ICB, 5hmC was more enriched across the permissive marker of gene expression H3K4me3, and depleted at H3K9me3, a modification associated with heterochromatin (Figure 76B). The dynamic modification of both 5hmC with ICB were in keeping with ICB-associated activation of immune expression. There was no significant difference between changes in 5hmC enrichment with ICB across T cell and monocytes sites, which might infer no preferential activation or apoptosis of one cell type versus another.

This suggests that MM is associated with distinct profiles of epigenetic enrichment across histone modification sites in cfDNA, which are then modulated by ICB. This may represent distinct immune activation states in MM, and dynamic transcriptional changes in circulating immune populations with ICB.

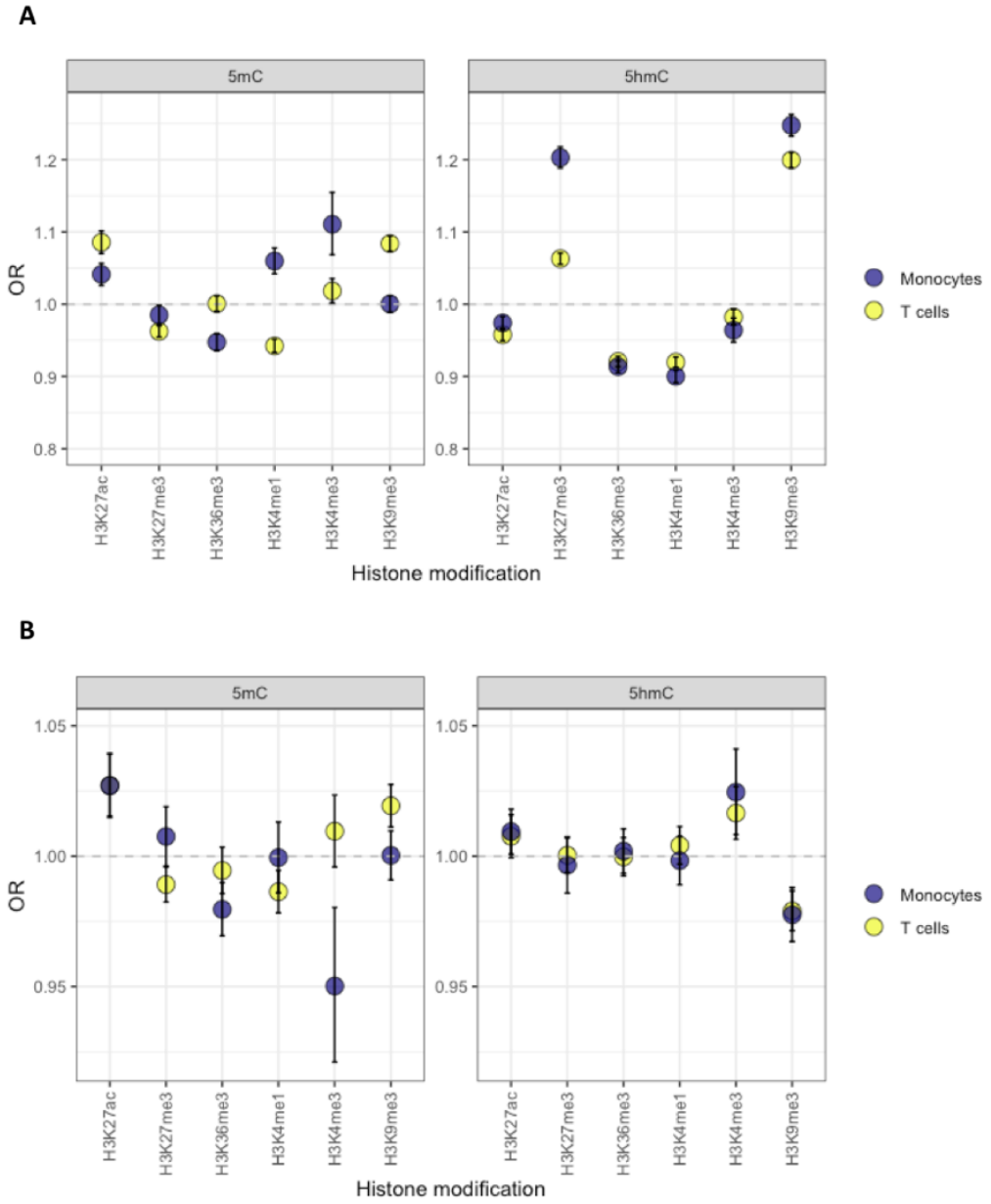


Figure 76: Enrichment of 5mC and 5hmC in cfDNA across histone modifications in MM patients and with ICB

A) Plots to show odds ratio (OR) of enrichment of 5mC and 5hmC peaksets in MM patients at baseline (5hmC C1 peaks n=373,933, 5mC C1 peaks n= 359,688) compared to healthy controls (5hmC HD peaks n=274639, 5mC HD peaks, n = 192,942). OR calculated using Fisher's exact test comparing overlaps of histone modifications across monocytes (cell type ID E029) and peripheral T cells (cell type IDE034). Whiskers indicate 95% confidence interval. B) Plots to show OR of enrichment of 5mC and 5hmC peaksets following ICB treatment versus at baseline (5mC: 359588 C1 peaks, 372869 C2 peaks; 5hmC: 373933 C1 peaks, 360602 C2 peaks). Whiskers indicate 95% confidence interval.

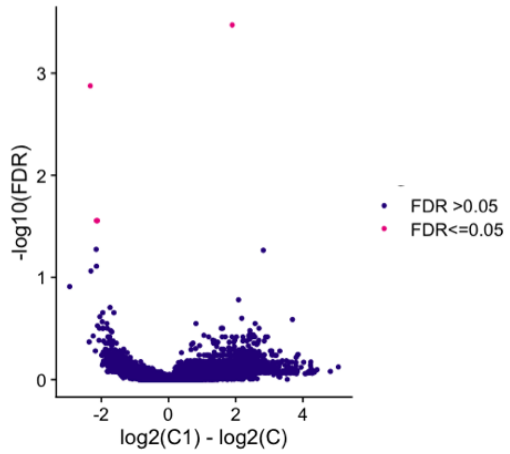
5.3.6 Characterising differential enrichment of 5hmC and 5mC cfDNA modifications in metastatic melanoma

Despite the previously described association of epigenetic profiles with cancer type[296,456,501], stage[296] and prognosis[254,255,297,298,300,314,463,464], there is little work in the literature characterising either 5hmC or 5mC cfDNA profiles in MM patient cohorts. To explore whether epigenetic cfDNA profiles were distinct in patients with MM and identify distinct genomic loci differentially modified, differential enrichment analysis was performed using the *DiffBind*[354] package in R. Regions ('peaks') of 5hmC or 5mC enrichment were identified using the *MACS2* package as previously described in section 5.3. Using the *DiffBind* package, consensus peaksets were identified comprising peaks located within at least two samples. Next, an expression matrix corresponding to reads overlapping these peaks per individual sample was created using the *dba.count* function. Differential enrichment analysis with *DESeq2* or *edgeR* was then performed (within the *DiffBind* wrapper application *dba.analyze*) to identify differentially 5-hydroxymethylated (DhMRs) or methylated regions (DMRs). Annotation of DhMRs and DMRs was performed using *ChipSeeker* as previously described.

5.3.6.1 5-hydroxymethylcytosine cfDNA profiles in MM

Differential enrichment analysis was performed to identify DhMRs using *DiffBind*. Four DhMRs in patients with MM were identified versus healthy donors (*edgeR* analysis controlling for batch across 23 C1 and 5 HD samples, FDR<0.05) (Figure 77). Annotation of DhMRs was carried out with *ChipSeeker*[350]. These included one DhMRs with greater 5hmC enrichment in MM patients, and three DhMRs with reduced 5hmC enrichment in patients (Figure 77). These DhMRs were found within *PCSK5*, which is involved in post-translational modification of integrin-α[502]; *VIRMA,* which has a role in alternative polyadenylation[503] and micro-RNA *MIR8076* (Figure 77B).

A



B

Chromosome	Start	End	Conc_C1	Conc_C	FDR	Genomic region	Gene
chr10	103741928	103744063	4.85	2.95	0.000336	Promoter (1-2kb)	*SH3PXD2A-AS1*
chr3	113424498	113425465	0.44	2.77	0.00133	Intron	*MIR8076*
chr9	75522537	75522866	0.26	2.41	0.0278	Distal Intergenic	*PCSK5*
chr8	94545910	94547687	0.87	2.99	0.0278	Promoter (2-3kb)	*VIRMA*

Figure 77: Differentially hydroxymethylated regions (DhMRs) in patients with MM
A) Volcano plot to show DhMRs between MM patients ('C1') and healthy donors ('C') (23 C1 patients, 5 healthy donors (dots indicate DhMRs, pink indicates FDR<0.05)). B) Table to show all significant DhMRs in MM patients. Conc_C1 and Conc_C2 is the mean read count across the region across conditions (Conc_C1 infers C1, Conc_C infers C), and genomic region type (FDR <0.05).

5.3.6.2 5-methylcytosine cfDNA profiles in MM

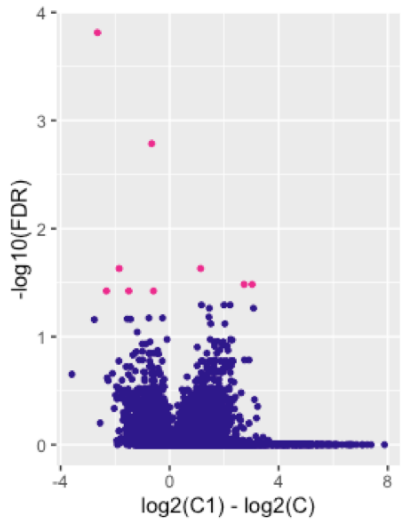
5mC cfDNA profiles have been shown to have utility as diagnostic biomarkers across a range of cancer types[254,255,314,464]. To explore whether this may be evident within this melanoma cohort, *DiffBind* was used to perform differential enrichment analysis of 5mC cfDNA profiles in patients with MM versus healthy controls. Nine 5mC peaks were differentially enriched between healthy controls versus MM patients (*edgeR* analysis controlling for batch effect with 24 C1 and 8 HD samples, FDR<0.05), including three that showed greater 5mC enrichment in MM patients, and six that showed reduced enrichment (Figure 78).

ChipSeeker was used to annotate the DMRs with genic location. Genes that showed increased 5mC labelling in patients included *TNRC6C*, protein kinase C eta gene *PRKCH* and retinoid isomerohydrolase gene *RPE65*. Genes that showed reduced 5mC labelling

in cancer patients included *PI3K3CB, FAM71E2* and diadenosine triphosphatase gene *FHIT*. *PI3K3CB* and *FHIT* are oncogenes that are recurrently mutated in cancers[504,505]. Inactivation of *FHIT* has also been shown to be associated with clonal expansion[506]. Methylation is variably associated with gene expression according to genomic location. Reduced 5mC enrichment within these genes within intronic regions (Figure 78B) may therefore be associated with reduced gene expression. This demonstrates that in MM distinct methylation patterns within tumour or circulating immune populations may facilitate dysregulation of mitogenic signalling pathways.

Previous studies have described 5mC and 5hmC cfDNA profiles with many thousands of DMRs/DhMRs in cancer patients versus healthy donors, typically within cancer-related genes. The limited number of DMRs and DhMRs identified within these data may be due to lack of power due to the relatively small sample size, particularly of healthy donor samples.

A



B

Gene	Fold change	FDR	Genomic region
LOC646813	-2.65	0.000154	Distal intergenic
FAM71E2	-0.66	0.00164	Promoter (<=1kb)
RPE65	1.14	0.0234	Distal Intergenic
FHIT	-1.85	0.0234	Intron
TNRC6C	3.03	0.0328	Exon
PRKCH	2.73	0.0328	Distal Intergenic
SNHG27	-1.5	0.0378	Intron
NUP93	-2.32	0.0378	Intron
PIK3CB	-0.59	0.0378	Intron

Figure 78: Differential 5mC cfDNA profiles in patients with MM
A) Volcano plot to show differentially enriched 5mC regions in cfDNA between baseline (C1 n = 24) and Control (C, n=8) samples (pink FDR<0.05, purple FDR >0.05) (edgeR analysis). B) Annotation of DMRs in MM at baseline versus healthy donors.

5.3.7 Characterising differential enrichment of 5hmC and 5mC modifications in ICB

To explore whether epigenetic cfDNA profiles in patients with MM were predictive of clinical outcome, differential enrichment analysis was performed with *DiffBind* across both 5hmC and 5mC datasets.

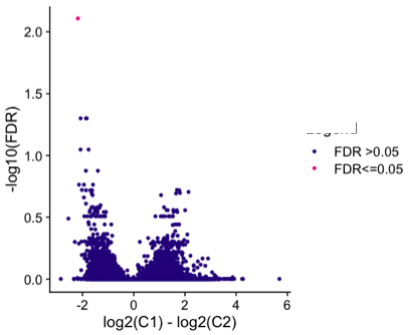
5.3.7.1 Modulation of 5-hydroxymethylcytosine cfDNA profiles with ICB

To identify differentially 5-hydroxymethylated regions (DhMRs) in cfDNA with ICB, pairwise differential enrichment analysis was carried out using the *DiffBind* package. Following identification of a consensus peakset composed of peaks found in at least two samples (n=266,103 5hmC peaks) only one DhMRs was identified following ICB treatment (13 paired samples, *edgeR* analysis, FDR<0.05). This was within *LINC01985*, which showed increased 5hmC enrichment following ICB.

Analysis of treatment-wise responses revealed one significant DhMR with sICB (10 paired samples, *edgeR* analysis, FDR <0.05) within the *HSPA12A* gene, and a single DhMR with cICB (3 patients, *DESEq2* analysis, FDR<0.05) within *AGAP1,* which showed reduced enrichment following ICB. This gene is a member of the Ras family and has a role in endosomal transport[507].

These observations show limited epigenetic modification of 5hmC cfDNA profiles with ICB, however results are likely to be limited by small sample size available for pairwise analysis and might also be confounded by a hetergeoneous response to treatment across the cohort.

A



B

Chromosome	Start	End	Conc_C1	Conc_C2	FDR	Genomic region	Gene
chr3	30297333	30298540	0.82	3	0.00781	Intron	*LINC01985*

C

Chromosome	Start	End	Conc_C1	Conc_C2	FDR	Genomic region	Gene
chr10	117000000	117000000	1.79	3.63	0.00353	Promoter (1-2kb)	*HSPA12A*

D

Chromosome	Start	End	Conc_C1	Conc_C2	FDR	Genomic region	Gene
chr2	236051819	236056507	5.61	-0.79	0.0351	Intron	*AGAP1*

Figure 79: DhMRs in patients with ICB
A) Volcano plot to show DhMRs in patients with ICB (13 patients, edgeR analysis, FDR<0.05). B) DhMRs enriched with ICB with annotations. C) DhMR in pairwise analysis of patients receiving sICB (10 paired samples, FDR<0.05). D) DhMR identified following pairwise analysis of patients receiving cICB (3 paired samples, FDR <0.05).

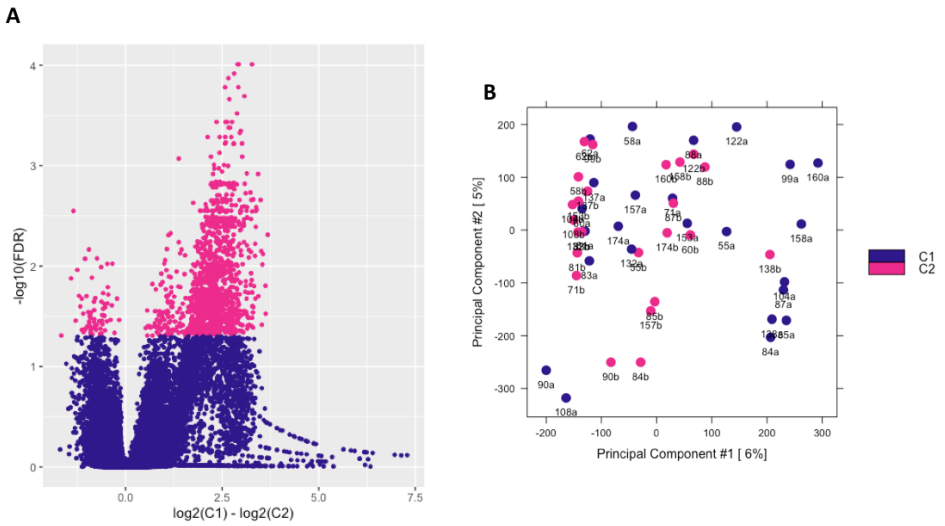
5.3.7.2 Modulation of 5-methylcytosine cfDNA profiles with ICB

Differential enrichment analysis was then carried out exploring changes in 5mC cfDNA profiles in response to ICB using the *DiffBind* package. Pairwise *edgeR* analysis (24 paired samples) identified a total of 1298 DMRs, including 69 DMRs enriched with ICB, and 1229 that were depleted (FDR<0.05) (Figure 80).

Annotation of these genomic regions was performed with *ChipSeeker*[350]. These DMRs were found across a total of 483 genes, including 69 genes with 5mC enrichment following ICB, and 281 that showed depletion. Genes that showed increased 5mC enrichment with ICB included cell signalling-related genes such as the voltage-gated potassium channel interacting protein 4 *KCNIP4,* and the RAS oncogene family member

RAB7A. Immune-related genes such as *TRAF3* also showed increased 5mC enrichment. This gene encodes the TNF receptor associated factor, which binds to and modifies signalling of the TNF receptor family, causing activation of NFκB and MAPK signalling[508]. The H3.5 histone gene *H3-5*, and histone deacetylase *HDAC4* were also 5mC-enriched with ICB.

Genes that showed depletion in 5mC labelling with ICB included genes associated with GTPase signalling, such as *GARNL3*, a GTPase activating Rap/RanGAP domain, GTPase activating protein and VPS9 domain 1 gene *GAPVD1,* and the RAB GTPase activating protein 1 gene *RABGAP1*. Other genes that contained regions of 5mC depletion with ICB include multiple zinc finger protein genes including *ZBTB43,* and *CIZ1*. Zinc finger genes have diverse range of cellular roles, including signalling, cell migration, DNA repair and the regulation of transcription[509]. Genes that showed reduced 5mC enrichment with ICB also included cell cycle regulatory genes including *CDK5RAP2*, *CIZ1* and *CDK9* (all of which showed depletion within promoter regions). This may reflect activation of cell turnover and proliferation in response to ICB.



C

Enriched DMRs

Chromosome	Start	End	FDR	Genomic Region	Gene
chr9	68769793	68770392	0.00282	Intron	*FAM122A*
chr21	17261186	17261445	0.00682	Distal Intergenic	*LINC01549*
chr6	21708825	21709117	0.00838	Intron	*CASC15*
chr22	23824612	23825645	0.00844	Promoter (1-2kb)	*SMARCB1*
chr18	78930778	78938379	0.00947	Distal Intergenic	*SALL3*
chr1	6109465	6112616	0.0095	Promoter (2-3kb)	*CHD5*
chr8	141609366	141613004	0.0102	Distal Intergenic	*MROH5*
chr7	43913004	43913832	0.0109	Intron	*URGCP*
chr1	54494949	54496169	0.013	Distal Intergenic	*ACOT11*
chr17	43609145	43609426	0.0132	Intron	*ETV4*
chr14	99184793	99188073	0.0162	Intron	*BCL11B*
chr11	42874960	42876804	0.0186	Distal Intergenic	*HNRNPKP3*
chr14	105865244	105868575	0.0192	Exon	*MIR4537*
chr3	38044873	38046342	0.0198	Exon	*DLEC1*
chr12	31774485	31775226	0.0199	Distal Intergenic	*H3-5*
chr10	4983990	4984731	0.0209	Downstream (2-3kb)	*AKR1C1*
chr7	158320364	158326107	0.0211	Intron	*PTPRN2*
chr6	131024547	131024783	0.0211	Intron	*EPB41L2*
chr19	51142410	51143619	0.0226	Promoter (<=1kb)	*SIGLEC7*
chr15	59275910	59277338	0.0226	Intron	*MYO1E*

Depleted DMRs

Chromosome	Start	End	FDR	Genomic region	Gene
chr9	126701387	126702672	9.77E-05	Exon	*LMX1B*
chr9	126414286	126416956	9.77E-05	Intron	*MVB12B*
chr9	126873032	126874418	9.77E-05	Intron	*ZBTB34*
chr9	123520640	123521982	1.21E-04	Intron	*MIR7150*
chr9	126107369	126108895	1.34E-04	Intron	*LOC101929116*
chr9	127466878	127467975	1.65E-04	Exon	*LRSAM1*
chr9	127121597	127122136	1.65E-04	Promoter (<=1kb)	*ANGPTL2*
chr9	126879737	126881126	2.03E-04	3' UTR	*ZBTB34*
chr9	126882472	126883923	2.17E-04	3' UTR	*ZBTB34*
chr9	126731952	126734104	3.02E-04	Distal Intergenic	*ZBTB43*
chr9	126451698	126452673	3.66E-04	Intron	*MVB12B*
chr9	126785577	126786477	3.66E-04	Distal Intergenic	*ZBTB43*
chr9	125076146	125077160	3.66E-04	Intron	*SCAI*
chr9	125314398	125315171	3.66E-04	Intron	*GAPVD1*
chr9	125132016	125132917	4.18E-04	Intron	*SCAI*
chr9	126918204	126918770	4.18E-04	Promoter (2-3kb)	*RALGPS1*
chr9	124989357	124991962	4.50E-04	Intron	*SCAI*
chr9	124798447	124799605	4.70E-04	Promoter (1-2kb)	*OLFML2A*
chr9	125136362	125137775	5.17E-04	Intron	*SCAI*
chr9	120542817	120544128	5.17E-04	Intron	*CDK5RAP2*

Figure 80: Differentially 5-methylated regions (DMRs) with ICB
A) Volcano plot to show DNRs following ICB (pairwise analysis with edgeR, n= 24 paired samples, FDR<0.05) Each point represents a methylated region. Purple peaks infers FDR>0.05, pink infers FDR<0.05). B) Principal component analysis plot to show first two principal components and read count matrix by cycle and by patient ID). D) Tables to show top 20 most significantly enriched or depleted differentially methylated regions including chromosome, region start and end point, FDR, genomic region and gene. Regions annotated with ChipSeeker.

Pathway enrichment analysis using *XGR* interrogation of the GOBP database across genes with increased 5mC enrichment following ICB identified pathways included positive regulation of transcription, negative regulation of cell proliferation. Pathway enrichment analysis of genes with reduced enrichment of 5mC following ICB included positive regulation of GTPase activity and signal transduction, response to oxidative stress, endocytosis and membrane organization.

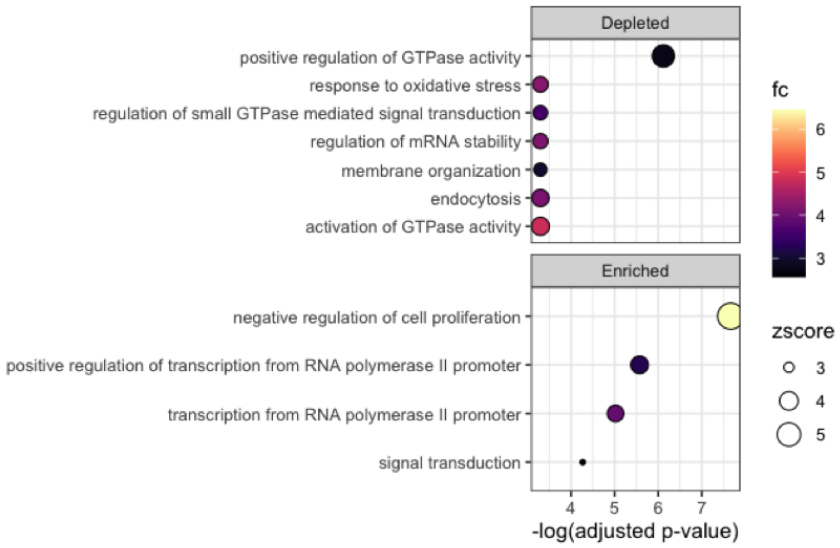


Figure 81: Differentially methylated regions with ICB
Pathway enrichment analysis using XGR interrogation of the GOBP database across genes containing DMRs with ICB with XGR interrogation of the GOBP database (P_{adj}<0.05).

It was notable that the majority of regions that showed depletion with ICB were within chromosome nine (Figure 80C), across a largely continuous sequence. To ensure that this was not due to contamination or the effect of an outlier, the analysis was repeated with a high minimum sample overlap (i.e. specifying that peaks within the consensus peakset were shared across a minimum of 30 samples, 40 and or 48 samples). With minimum overlap set to 30 or 40, again the majority of the differentially enriched regions were found within chromosome nine. Loss of chromosome 9 typically occurs in malignancies such as bladder cancer, where it is associated with silencing of *CDKN2A*[510,511]. In melanoma, deletion of chromosome 9 has been reported to be an early event in tumourigenesis[512]. Alternatively, there exist satellite regions within long arms of chromosome 1, 9 and 16 which are known to be prominently methylated[513,514]. DMRs within chromosome 9 may represent changes within these regions following ICB. In addition, BRAF fusion with partners such as CDK5RAP2 are reported in melanoma[515]. The identification of DMRs predominantly within this chromosome 9 region may reflect a similar fusion event. Abberrant methylation is typical in cancer and this is suggestive of cfDNA melanoma as opposed to leukocyte origin.

Interestingly, there were a number of DMRs within genes associated with spermatogenesis and sperm motility such as *ODF2, STRBP, SPACA9*, *ODF2L, CFAP157, CFAP77, DLEC1* and *CT66. SPACA9* encodes the sperm acrosome associated 9 protein, whilst *STRBP* encodes spermatid perinuclear RNA binding protein, and *ODF2* the outer dense fibre of sperm tails 2 protein. *CFAP157, CFAP77* and *DLEC1* are cilia and flagella-associated proteins, involved in sperm motility. This may represent DMR within cancer testis antigens (CTAs).

CTAs are genes whose expression is usually limited to gametes, typically the testis, but which are expressed in non-germ cell cancers[516]. Cancers derived from almost all tissue types have been shown to express CTAs[516]. CTAs have been shown to promote cell survival, metastasis, and genomic instability, thereby promoting tumorigenesis[517]. However, as these genes are almost tumour-specific and exhibit marked immunogenicity, these are attractive therapeutic targets[517]. Melanoma was one of the first cancers in which CTAs were identified and characterised[518,519]. Whilst databases of recognised CTAs exist, literature regarding CTAs within cfDNA is almost completely absent, similarly as is work regarding epigenetic profiling and modification of cfDNA CTAs. Epigenetic modulation of CTAs has previously been reported, and variable methylation is a way in which CTA expression is regulated[520,521]. In this dataset both *CT66* and *ODF2* have been previously documented as a CTAs according to the CTdatabase[522]. *ODF2, STRBP, SPACA9*, *ODF2L* and *CFAP157* contain DMRs within promoter regions with reduced enrichment following ICB. DMRs associated with both 5mC enrichment and depletion following ICB within possible CTAs were indetified within both promoter and intragenic regions (Figure 82). Although directional effect is not clear, it might be hypothesised that modification of 5mC profiles in CTAs regulating expression may be another mechanism through which ICB modulates immune responses.

Modification	Analysis	Treatment	Gene	Name	Region	DMR/DhMR
5mC	C1vC2	All	*ODF2*	outer dense fiber of sperm tails 2	Promoter, intron, exon, distal intergenic	Depleted
5mC	C1vC2	All	*STRBP*	spermatid perinuclear RNA binding protein	Promoter, intron, exon, distal intergenic	Depleted
5mC	C1vC2	All	*SPACA9*	sperm acrosome associated 9	Promoter, intron	Depleted
5mC	C1vC2	All	*ODF2L*	outer dense fiber of sperm tails 2 like	Promoter	Depleted
5mC	C1vC2	All	*CFAP157*	cilia and flagella associated protein 157	Promoter/intron	Depleted
5mC	C1vC2	All	*CFAP77*	cilia and flagella associated protein 77	Intron	Depleted
5mC	C1vC2	All	*CT66*	cancer/testis associated transcript 66	Intergenic	Enriched
5mC	C1vC2	All	*DLEC1*	DLEC1 cilia and flagella associated protein	Exon	Enriched
5mC	C1vC2	sICB	*TEX101*	testis expressed 101	5'UTR	Enriched
5mC	C1vC2	sICB	*TOPAZ1*	testis and ovary specific PAZ domain containing 1	Promoter	Enriched
5mC	C1vC2	sICB	*CFAP74*	cilia and flagella associated protein 74	Exon	Depleted
5mC	C1vC2	sICB	*CFAP46*	cilia and flagella associated protein 46	Promoter	Depleted
5mC	C1vC2	sICB	*DLEC1*	DLEC1 cilia and flagella associated protein	Promoter	Enriched
5mC	C1vC2	sICB	*PSLNR*	prostate enriched lncRNA	Intron	Enriched
5mC	C1vC2	sICB	*THEG*	theg spermatid protein	3'UTR	Enriched
5mC	C1vC2	cICB	*TEX48*	testis expressed 48	Intron, distal intergenic	Depleted
5mC	C1vC2	cICB	*TEX53*	testis expressed 53	Promoter	Depleted
5mC	C1vC2	cICB	*THEG5*	testis highly expressed protein 5	Distal intergenic	Enriched
5mC	C1vC2	cICB	*ODF2*	outer dense fiber of sperm tails 2	Promoter, intron, exon	Depleted
5mC	C1vC2	cICB	*SPACA9*	sperm acrosome associated 9	Exon, intron, 5'UTR, 3'UTR	Depleted
5mC	C1vC2	cICB	*STPG1*	sperm tail PG-rich repeat containing 1	Intron	Enriched
5mC	C1vC2	cICB	*STRBP*	spermatid perinuclear RNA binding protein	Promoter, intron, exon, 3'UTR, distal intergenic	Depleted
5mC	C1vC2	cICB	*CFAP157*	cilia and flagella associated protein 157	Promoter, intron	Depleted
5mC	C1vC2	cICB	*CFAP77*	cilia and flagella associated protein 77	Intron, exon	Depleted
5mC	C1 & AI_c5	All	*STRBP*	spermatid perinuclear RNA binding protein	Distal intergenic	Enriched in AI_c5
5hmC	C2 & six month progression	All	*SPATS2L*	spermatogenesis associated serine rich 2 like	Promoter	Enrichd in progressors
5hmC	C2 & AI_Ses	All	*CFAP157*	cilia and flagella associated protein 157	Promoter	Depleted in AI_c5
5hmC	C2 & AI_Ses	All	*SPATA21*	spermatogenesis associated 21	Exon	Enriched in AI_c5

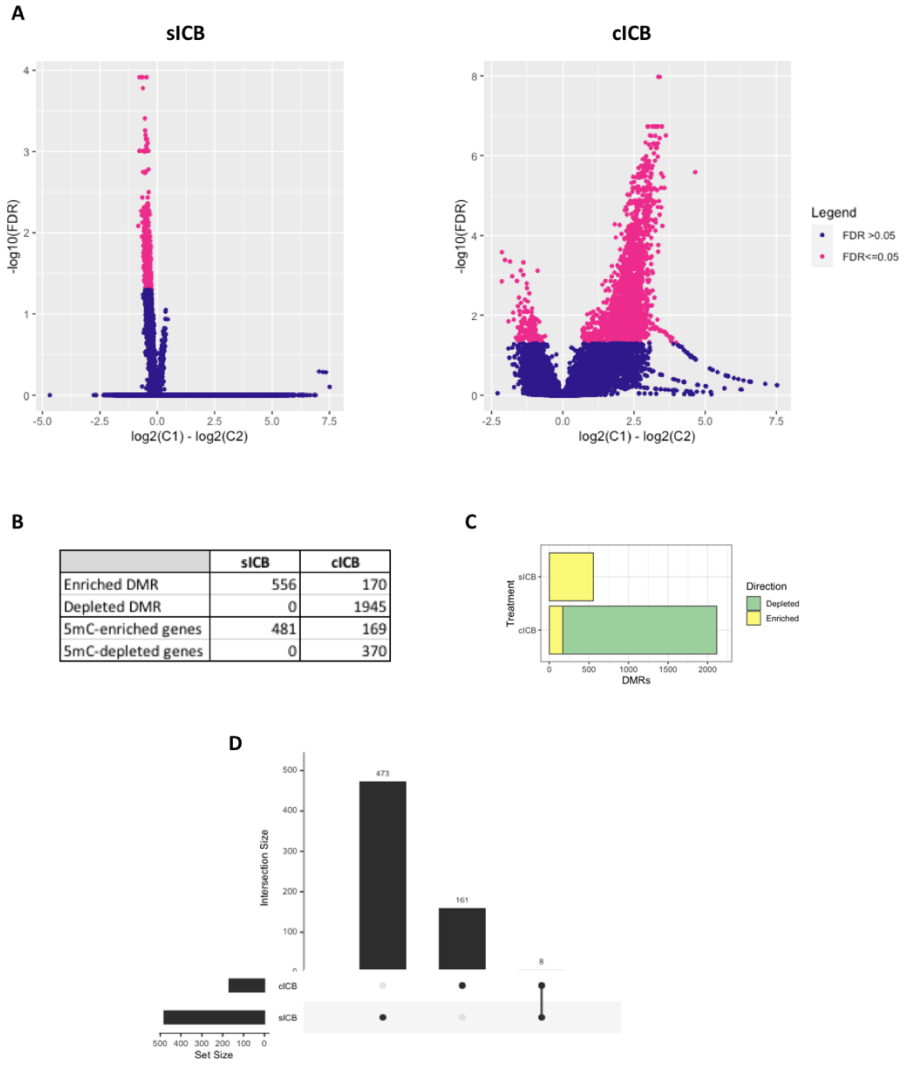
Figure 82: Cancer testis antigens (CTAs) and DMRs/DhMRs
DMRs and DhMRs found within germ cell-related genes (FDR<0.05). Genes highlighted in blue are confirmed CTAs and green denotes that other members of the gene family are confirmed CTAs. Analysis indicates in which analysis the DMR/DhMR was identified (AI_c5 denotes irAEs effects prior to cycle 5 and 6MP indicates six-month progression). 'Treatment' indicates whether analysis was performed using all samples, or only those receiving sICB or cICB. Region indicates genomic location within which DMR/DhMR lies (if DMR/DhMR is identified across multiple functional regions these are listed). DMR/DhMR denotes whether 5mC or 5hmC enrichment was depleted or enriched.

5.3.7.3 Comparing 5mC cfDNA modulation with combination versus single agent immune checkpoint blockade

We have previously demonstrated that cICB is associated with the same directional transcriptional effect as sICB in $CD8^+$ T cells[71], but that cICB is associated with a greater magnitude of effect. The same observation was found to be true of circulating monocytes in chapter three (3.3.4.1). To explore whether these effects are observed in cfDNA, analysis of DMRs by treatment type was explored using *DiffBind*. Pairwise differential enrichment analysis following ICB in patients receiving single agent ICB (sICB) identified 556 DMRs within 481 genes. Interestingly, all DMRs were 5mC enriched with sICB rather than depleted (*DESeq2* analysis, 10 paired samples, FDR<0.05) (Figure 83). Differential enrichment analysis of 5mC cfDNA profiles patients receiving combination therapy (cICB) showed a greater number of DMRs, with 2,115 significant DMRs found within 539 genes, 169 of which showed enrichment of 5mC with ICB, and 370 showed depletion of 5mC labelling with ICB (*edgeR* analysis, 14 paired samples, FDR<0.05)

(Figure 83). It was noted that the magnitude fold change of 5mC enrichment was greater in those receiving cICB versus sICB (Figure 83A).

Overlap of 5mC-enriched and depleted genes showed that differentially methylated gene sets were distinct between those receiving sICB versus cICB, with little overlap between the two cohorts (Figure 83C). DMRs in sICB and cICB cohorts were found across the genome, but notably the preponderance of DMRs on chromosome 9 was only seen in the cohort receiving cICB (Appendix FF and Appendix GG).



	sICB	cICB
Enriched DMR	556	170
Depleted DMR	0	1945
5mC-enriched genes	481	169
5mC-depleted genes	0	370

Figure 83: Comparing sICB and cICB responses
A) Volcano plots to show differentially methylated regions (DMRs) in patients receiving sICB (DESeq2, 10 paired samples, FDR<0.05) and cICB (edgeR, 14 paired samples, FDR<0.05). B) Table to show number of DMR and associated genes for sICB and cICB pairwise analysis. C) Barplot to show DMRs and 5mC enrichment versus depletion (FDR <0.05) in patients receiveing sICB versus cICB. D) Upsetplot to show intersection between the genes showing 5mC-enrichment with ICB in patients receiving sICB (481 genes) and cICB (169 genes).

Differential methylation with cICB and sICB was seen across genes involved in a wide range of biological processes (summarised in Figure 84, see Appendix FF for gene list).

Amongst those genes containing DMRs in those receiving sICB (all of which show increased 5mC enrichment with sICB), numerous immune-related genes were seen. These include *JAK2,* a component of the JAK/STAT signalling pathway facilitating

signalling and activation by type II interferon[523]; the pro-inflammatory cytokine *IL16* which facilitates chemotaxis, binds to CD4 and induces IL-2 expression activating T cells[524]; the pro-inflammatory cytokine IL-17 receptor *IL17REL,* which induces NFκB and MAPK signalling pathways[525] and which is associated with autoimmunity[526]. Other immune-related genes containing DMRs included *TNFAIP2, TNFRSF25* and *C1QTNF3*. *NFATC1*, which encodes a transcription factor associated with T cell receptor activation[527], also contained a DMR. There was no overlap between these genes and those found to be differentially regulated by ICB in Chapter Three.

Other genes that showed 5mC enrichment in those receiving sICB included genes involved in transcription factors such as *TCF25, R2F1* and *SOX10. SOX10* has been implicated in melanoma development, and antigenic expression assessed via IHC in melanoma is commonly used as a diagnostic aid in clinical practice[528]. This may reflect a tumour-related signature. Genes involved in mitogenic signalling pathways were also 5mC enriched, including RAS-like proto-oncogene B, RAS protein activators *RASA3* and *RASAL3*, as well as *MAPK11* and *AKT1*. Genes associated with cell cycle progression such as *CDK19* and *GAK* also showed 5mC enrichment in those receiving sICB. Other genes enriched for 5mC with cICB include *TERT,* which encodes telomerase reverse transcriptase which has a role in maintaining cell senescence and lifespan[529].

In those who received cICB, genes that showed enrichment of 5mC with treatment included zinc finger genes (e.g. *ZNF329, ZNF473* and *ZNF438*). Signalling associated genes including the potassium channels *KCNMB4* and *KCNIP1* were also enriched, as well as mitogenic signalling gene mitogen-activated protein kinase gene *MAP4K4* and *RHOB*, a RAS homolog. Similarly, as with sICB, *TERT* was found to exhibit 5mC enrichment following treatment. With differential enrichment analysis the samples are analysed as replicates in a pairwise manner, so the 'contribution' from each patient is difficult to dissect.

In those receiving cICB, genes that showed depletion of 5mC with treatment included those involved in GTPase signalling such as *RABGAP1, GAPVD1* and *GARNL3*. Immune-related genes including *TLR4,* which has a pro-inflammatory in response to LPS via activation of NFκB[530,531]; and *TNFSF8,* a cytokine which stimulates T cell proliferation;

and tumour necrosis factor receptor associated factor 1 gene *TRAF1,* which forms a dimer with TRAF2 and activates MAPK and NFkB signalling[532], were 5mC depleted with cICB. Similarly, complement gene *C5*, which encodes C5 which is cleaved into C5a with role in chemotaxis, and C5b which forms the membrane attack complex with a key role in lysis, was also found to be 5mC depleted following cICB[533]. This may reflect modulation of myeloid cell expression following cICB.

As with pairwise analysis across all samples, treatment-wise analysis identified DMRs in recognised and possible CTAs (Figure 82). In those receiving sICB, as with the pairwise differential methylation analysis across all samples, ciliary and flagella-associated genes *CFAP74* and *CFAP46* showed a reduction in methylated regions with ICB. In addition, genes including testis expressed 101 gene *TEX101*, and testis and ovary specific PAZ domain containing 1 gene *TOPAZ1* were also differentially methylated, but showed enrichment with treatment. In those receiving cICB, pairwise analysis again identified *CFAP157* and *CFAP77* genes, as well as testis expressed genes *TEX48* and *TEX53,* showing 5mC depletion with ICB.

Pathway enrichment analysis using XGR interrogation of the GOBP database only demonstrated enrichment of pathways across genes with enrichment of 5mC with sICB. These included positive regulation of GTPase activity (GO:0043547), protein phosphorylation (GO:0001934) and positive regulation of transcription (GO:0045893) (Figure 85). No pathways were enriched across 5mC depleted or enriched pathways with cICB.

These results suggest that ICB is associated with modulation of 5mC in cfDNA, with differential methylation likely to represent both tumour and immune signatures. The pathways enriched for genes containing DMRs following both sICB and cICB are similar. These results suggest that the greater magnitude effect in epigenomic profiles seen within immune cells with cICB versus sICB is reflected in 5mC cfDNA profiles, and that cICB is associated with predominant depletion of 5mC within cfDNA, whilst sICB is associated with enrichment (although this observation may be limited due to cohort size).

Analysis	5mC depleted/enriched	Associated cellular function	Example DMR genes
sICB C1vC2	Enriched	Immune system	*JAK2, IL16, IL17REL, TNFA1P2, TNFRSF25, C1QTNF3*
		Cell signalling	*ZC3H12C, ZNF624, ZNF467, ZNF324B, ZBTB46,*
		Mitogenic signalling	*RASA3, MAPK11, AKT1, PI3K*
		Cell maintenance/lifespan	*TERT*
		Ion transmembrane transport	*SLC12A3, SLC12A7, SLC1A7, SLC9A3, KCNF1, KCNG2, KCNJ18, KCNQ2, KCNV2, NEDD4L, TRPM5*
		Transcription/Transcriptional regulation	*E2F1, EGR1, HDAC4, ZNF516, LINC01310, MIR3187, H1FX-AS1*
sICB C1vC2	Depleted	N/A	*N/A*
cICB C1vC2	Enriched	Cell signalling	*ZNF329, ZNF473, ZNF438, GPR39, ZNF331*
		Mitogenic signalling	*MAP4K4, RHOB, PRKCG*
		Ion transmembrane transport	*CACNA2D4, KCNIP1*
		Apoptosis	*BCL2*
		Cell maintenance	*TERT*
		Transcription/Regulation of transcription	*LINC00701, LINC01524, LINC02537, MIR153-2, MIR548Q, MIR296, LINC01435, TCF7L2, ETV4*
cICB C1vC2	Depleted	Immune-related genes	*TLR4, TNFSF8, TRAF1, C5*
		Cell signalling	*ABL1, AK1, AK8, DNM1, GPR107, GAPVD1, GARNL3, HRH2, MAPKAP1, NEK6, PIP5KL1, PLPP7, PKN3, RABGAP1, RIN2, RGS3, ROCK1, ZBTB34, ZC3HC1, ZNF618*
		Cell adhesion	*CDH4, CDH9, CERCAM*
		Cell cycle regulation	*CDK5RAP2, CDK9*
		Structural genes	*COL5A1, COL21A1, COL27A1, GLE1*
		Immune-related	*C5, HMCN2, IRF4, TLR4, TRAF1, TNFSF8*
		Inflammatory response	*PTGES, PTGS1*
		Transcription/Transcriptional regulation	*GTF3C5, LMX1B, MIR147A, MIR4478*
		Metabolism	*GYS2, MRRF*
		Ion transmembrane transport	*KCTD20, KCNJ9, KCNT1*
		Apoptosis	*NIBAN2, NAIF1*

Figure 84: Key selected genes exhibiting DMRs with ICB

Table to show key genes containing DMR in pairwise differential enrichment analysis between C1 versus C2 in patients receiving sICB versus cICB and associated biological function.

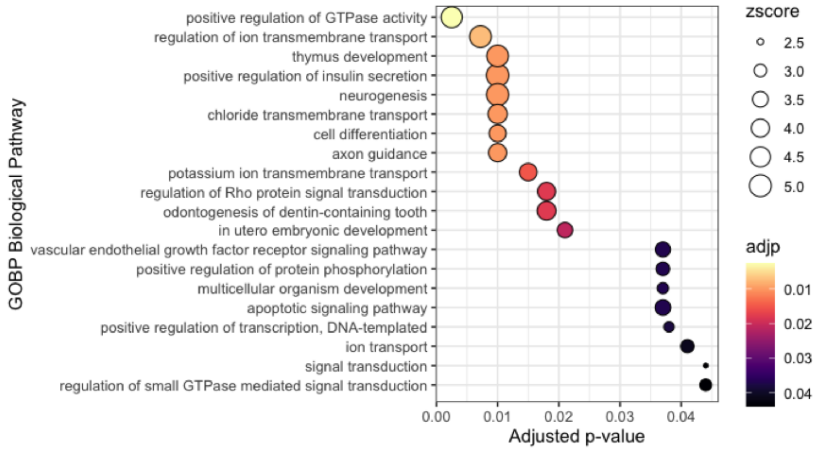


Figure 85: Pathway enrichment analysis of genes enriched for 5mC with sICB
Pathway enrichment analysis using XGR interrogation of the GOBP database of genes enriched for 5mC with sICB (556 DMR annotated to 481 genes with ChipSeeker, P_{adj}<0.05).

5.3.8 Characterising association with epigenetic cfDNA profiles with clinical outcome

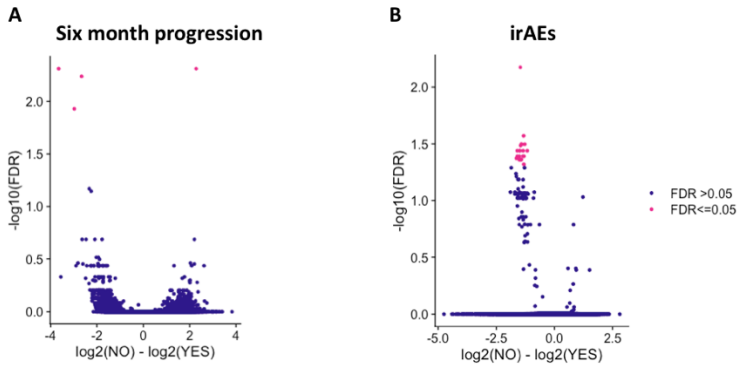
The association between epigenetic cfDNA profiles and clinical outcome in MM patients treated with ICB was next explored.

5.3.8.1 5-hydroxymethylcytosine cfDNA profiles and clinical outcome

Differential enrichment analysis across consensus peakset of 242,241 peaks (minimum samples overlap = 2) created with *DiffBind* identified four DhMRs significantly associated with progression at six months at baseline (23 patients, *edgeR* analysis controlling for batch effect, FDR<0.05) (Figure 86A). These DhMRs were found within genes including matrix metalloproteinase *MMP17*, which is expressed by monocytes with a role in inflammation[534], and also the serine/threonine kinase gene *SIK1B,* within which 5hmC was enriched in those who progressed. Enrichment of 5hmC in transcription factor *RUNX1* was associated with lack of progression at six months (Figure 86B). *RUNX1* is a transcription factor which has a key role in regulating haematopoiesis, and mutations in which are associated with myeloid haematological cancer[535].

Association with 5hmC profiles baseline in untreated samples and the development of irAEs prior to cycle 5 identified a total of 22 DhMRs (23 patients, *DESeq2* analysis, FDR<0.05) (Figure 86D), all of which demonstrated positive association between 5hmC

enrichment and the development of irAEs. DhMRs were found within a total of 17 genes. including those associated with G protein signalling such as *GPSM1,* kinase signalling such as *DIPK1B* and G protein signalling modulatory gene *GPSM*. Other genes associated with DhMRs include microRNAs such as *MIR4673* and *MIR4669.* These DhMRs were found predominantly within promoter regions in genes on chromosome 9.



C

Six month progression

Chromosome	Start	End	Conc_NO	Conc_YES	FDR	Genomic annotation	Gene
chr21	43401397	43401725	0.03	3.68	0.0049	Distal Intergenic	*SIK1B*
chr21	34883247	34887343	4.81	2.53	0.0049	Promoter (<=1kb)	*RUNX1*
chr10	72393523	72394068	0.22	2.88	0.0058	Intron	*DNAJB12*
chr8	88243279	88249266	0.03	3.01	0.0118	Intron	*MMP16*

D

irAEs

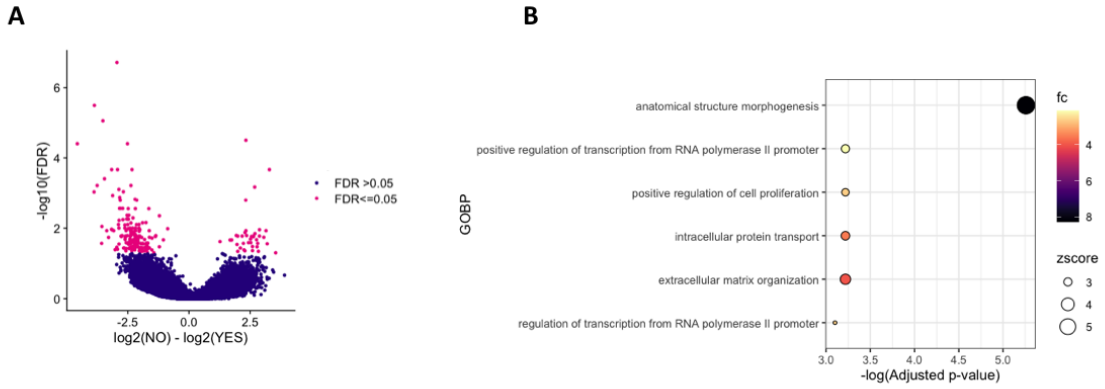
Chromosome	Start	End	Conc_NO	Conc_YES	FDR	Genomic region	Gene
chr9	133934253	133955398	4.19	5.54	0.0407	Promoter (<=1kb)	*VAV2*
chr9	136465590	136482070	3.66	5.3	0.0425	Promoter (<=1kb)	*SEC16A*
chr9	134327277	134368541	5.42	6.74	0.0269	Promoter (<=1kb)	*RXRA*
chr9	134400054	134430391	4.77	5.94	0.0363	5' UTR	*RXRA*
chr9	136219462	136236746	3.94	5.45	0.0425	Promoter (<=1kb)	*QSOX2*
chr9	137302995	137322495	3.99	5.59	0.0363	Promoter (<=1kb)	*NRARP*
chr9	136560160	136579714	4.37	5.68	0.0407	Distal Intergenic	*NALT1*
chr9	135993200	136028685	4.82	6.22	0.0319	Promoter (<=1kb)	*NACC2*
chr9	136072363	136089388	4.42	5.77	0.0363	Intron	*NACC2*
chr9	136484100	136544516	6.08	7.55	0.00668	Promoter (<=1kb)	*MIR4673*
chr9	134369559	134398722	5.08	6.35	0.0319	Promoter (<=1kb)	*MIR4669*
chr9	134044375	134061025	4.02	5.33	0.048	Promoter (<=1kb)	*LOC100130548*
chr9	136613560	136630186	4.37	5.88	0.0363	Promoter (<=1kb)	*LINC01451*
chr9	136593643	136613081	3.96	5.38	0.0436	Promoter (2-3kb)	*LINC01451*
chr9	136328399	136346424	3.83	5.35	0.0436	Promoter (<=1kb)	*GPSM1*
chr9	137620053	137642599	4.02	5.48	0.0436	Promoter (<=1kb)	*EHMT1*
chr9	136713275	136726964	3.98	5.49	0.0407	Promoter (<=1kb)	*DIPK1B*
chr9	135875073	135891540	3.76	5.36	0.0407	Promoter (<=1kb)	*CAMSAP1*
chr9	136687760	136711922	4.55	6.01	0.0327	Promoter (<=1kb)	*AGPAT2*
chr9	136666660	136686327	4.7	6.14	0.0319	Promoter (<=1kb)	*AGPAT2*
chr9	137004257	137026093	4.33	5.8	0.0363	Promoter (<=1kb)	*ABCA2*

Figure 86: DhMRs and clinical outcome at baseline

Volcano plots to show DhMRs in patients at baseline (23 C1 patients) associated with: A) six-month progression (edgeR analysis) and; B) the development of irAEs prior to cycle 5 (DESeq2 analysis). Pink points indicate FDR<0.05). C) Table to show all DhMRs at baseline associated with progression at six months (FDR<0.05) D) Table to show all DhMRs at baseline associated with irAEs prior to cycle 5 (FDR<0.05). Average read count by clinical response is inferred by ('Conc_No'/'Conc_Yes').

There was more prominent association between post-treatment 5hmC cfDNA profiles and clinical outcome. Analysis of post-treatment (C2) 5hmC cfDNA profiles and clinical outcome with *DiffBind* identified 221 DhMRs associated with six month progression, including 183 DhMRs where 5hmC enrichment was associated with progression at six

months, and 38 DhMRs where 5hmC cfDNA enrichment was greater in those who failed to progress (22 patients, *edgeR* analysis controlling for batch, FDR<0.05) (Figure 87).These DhMRs were associated with a total of 208 different genes. Greater 5hmC enrichment was associated with six-month progression in genes associated with wide range of biological processes. These include transmembrane protein genes *TMEM38A, TMEM87B* and *TMEM8B,* zinc finger genes such as *ZNF704* and *ZFPM2*, and genes involved in cell adhesion such as *CDH1* and *CDH4.* The promoter region of spermatogenesis associated serine rich 2 like gene *SPATS2L* also contained a region more 5hmC-enriched in those who progressed (Figure 82). As previously discussed, this may represent differential 5-hydroxymethylation within a CTA. Pathway enrichment analysis using *XGR* interrogation of the GOBP database across the associated DhMR genes revealed six pathways differentially enriched pathways in genes more 5hmC-enriched in those who progressed at six months. These include positive regulation of cell proliferation (GO:0008284) and extracellular matrix organisation (GO:0030198) (Figure 87B).



C

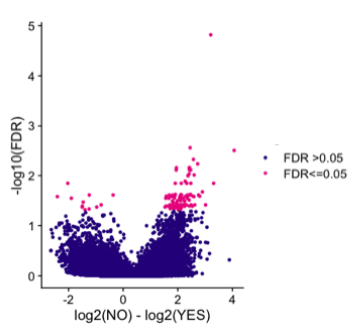
Chromosome	Start	End	Conc_NO	Conc_Yes	FDR	Genomic region	Gene
chr8	69462608	69464266	1.42	4.37	1.92E-07	Promoter (2-3kb)	*SULF1*
chr2	235871774	235876708	0.88	4.75	3.17E-06	Intron	*AGAP1*
chr2	217248505	217250362	0.42	3.94	8.72E-06	Distal Intergenic	*DIRC3-AS1*
chr6	144036516	144038303	4.1	1.78	3.11E-05	Exon	*PLAGL1*
chr12	64229294	64230045	0.87	3.39	3.90E-05	Intron	*C12orf66*
chr8	46218024	46221635	0.22	4.79	3.90E-05	Distal Intergenic	*ASNSP1*
chr19	14774966	14775952	1.17	3.5	2.14E-04	Promoter (<=1kb)	*ADGRE2*
chr22	23664018	23665367	4.04	0.75	2.14E-04	Intron	*GUSBP11*
chr15	70863604	70864445	0.52	3.44	2.14E-04	Intron	*LARP6*
chr17	72582761	72586447	1.37	4.53	2.14E-04	Promoter (<=1kb)	*LINC00673*
chr12	54697849	54699396	0.22	3.68	3.90E-04	Distal Intergenic	*DCD*
chr19	33057916	33058680	0.84	3.21	6.04E-04	Intron	*RHPN2*
chr8	56887029	56890360	0.22	3.98	6.04E-04	Distal Intergenic	*IMPAD1*
chr22	37590762	37592252	4.45	1.77	6.69E-04	Distal Intergenic	*LGALS2*
chr12	13390006	13390838	0.22	3.09	7.81E-04	Promoter (2-3kb)	*LINC01559*
chr13	105327480	105328408	0.36	3.04	9.10E-04	Distal Intergenic	*DAOA*
chr13	111895922	111898800	0.64	4.52	9.21E-04	Promoter (1-2kb)	*LINC00354*
chr12	54153705	54156829	0.39	3.51	1.17E-03	Intron	*SMUG1*
chr14	66553604	66554530	0.55	3.38	1.33E-03	Intron	*GPHN*
chr20	61421903	61422198	0.4	2.6	1.50E-03	Intron	*CDH4*

Figure 87: DhMRs post-treatment and progression at six months
A) Volcano plots to show DhMRs in cfDNA post-treatment (C2) associated with progression at six months (pink dots indicate FDR<0.05) (22 patients, edgeR analysis). B) Pathway enrichment analysis of genes containing DhMRs where greater 5hmC enrichment was associated with six-month progression (P_{adj}<0.05) C) Table to show top 20 most significant DhMRs in cfDNA at C2 associated with disease progression at six months.

A total of 113 DhMRs post-treatment were found to be associated with the development of irAEs prior to cycle 5 (21 patients, *edgeR* analysis, FDR<0.05) (Figure 88). The majority of these (101/115) showed greater enrichment in those who did not develop irAEs. DhMRs were associated with 97 genes. These include *CDK5RAP1*, a methylthiotransferase which inactivates CDK5[536] which has a role in a wide range of biological processes, such as motility and adhesion; and *SLC36A1*, which is involved in lysosome production and amino acid transport[537]. Other genes containing DhMRs

associated with irAEs include immune-related genes such as the Toll-like receptor 4 gene *TLR4*, and GTPase and kinase related genes such as *GAPVD1* and *NEK6* respectively. DhMRs found within genes such as cilia and flagella associated protein gene *CFAP157* and spermatogenesis associated 1 gene *SPATA21*, which may represent CTAs.

A



B

Chromosome	Start	End	Conc_NO	Conc_YES	FDR	Genomic region	Gene
chr14	106268326	106270165	1.05	3.08	0.0142	5' UTR	*LINC00226*
chr2	101245693	101247402	2.97	3.34	0.0243	Intron	*TBC1D8*
chr12	4043633	4044811	2.19	3.43	0.0243	Distal Intergenic	*PARP11*
chr1	16408584	16410844	1.15	3.55	0.0262	Exon	*SPATA21*
chr2	74064103	74064596	1.69	3.59	0.0283	Intron	*TET3*
chr1	12217039	12217398	1.38	2.85	0.0337	Distal Intergenic	*VPS13D*
chr17	32411753	32412713	1.72	2.53	0.0383	Exon	*PSMD11*
chr15	64264620	64266237	1.74	3.23	0.0405	Exon	*CSNK1G1*
chr13	59464903	59465461	2.28	3.25	0.0423	Distal Intergenic	*DIAPH3*
chr1	39444117	39444778	1.51	3.01	0.0423	Exon	*MACF1*
chr17	20115287	20117258	2.91	4.16	0.0463	Intron	*SPECC1*
chr1	156716581	156719067	3.49	4.88	0.0479	Downstream (2-3kb)	*ISG20L2*

C

Chromosome	Start	End	Conc_NO	Conc_YES	FDR	Genomic region	Gene
chr9	130407375	130409412	3.5	0.3	1.52E-05	3' UTR	*HMCN2*
chr11	124469699	124471100	2.85	0.4	2.75E-03	Distal Intergenic	*OR8B8*
chr8	46218024	46221635	4	-0.06	3.11E-03	Distal Intergenic	*ASNSP1*
chr9	124146072	124146592	2.51	-0.06	4.70E-03	Distal Intergenic	*NEK6*
chr9	130565443	130566092	2.66	-0.06	5.77E-03	Distal Intergenic	*FUBP3*
chr9	125325991	125326693	3.04	0.6	6.88E-03	Exon	*GAPVD1*
chr9	95806376	95807058	2.8	0.85	6.88E-03	Intron	*ERCC6L2*
chr9	123636226	123637245	2.51	0.09	6.88E-03	Intron	*DENND1A*
chr9	129529430	129531618	2.41	-0.06	7.66E-03	Distal Intergenic	*LINC00963*
chr9	124666791	124667335	2.59	0.16	7.66E-03	Intron	*MIR181A2HG*
chr9	99725529	99725839	2.49	0.55	7.66E-03	Intron	*LOC101928438*
chr20	33337158	33338391	3.49	1.07	9.21E-03	Distal Intergenic	*CDK5RAP1*
chr12	76729340	76729956	2.83	0.23	9.70E-03	Distal Intergenic	*ZDHHC17*
chr9	87273112	87273655	2.47	0.1	1.01E-02	Distal Intergenic	*C9orf170*
chr9	131851661	131852057	2.36	0.09	1.29E-02	Downstream (<1kb)	*RAPGEF1*
chr9	128003592	128004173	2.26	-0.06	1.42E-02	Distal Intergenic	*FAM102A*
chr1	71843665	71845313	2.59	0.46	1.42E-02	Intron	*NEGR1-IT1*
chr6	161291855	161293119	3.2	1.3	1.42E-02	Distal Intergenic	*AGPAT4*
chr9	124470904	124473856	3.67	1.2	1.42E-02	Intron	*NR5A1*
chr8	56887029	56890360	3.24	-0.06	1.42E-02	Distal Intergenic	*IMPAD1*

Figure 88: DhMRs post- ICB treatment (C2) and association with irAEs

A) Volcano plot to show DhMRs in cfDNA at C2 associated with development of irAEs prior to the fifth treatment cycle (C5) (edgeR analysis, 21 patients, FDR <0.05). B) Table to show DhMRs in cfDNA post-treatment associated with the development of irAEs and which show higher 5hmC enrichment in patients who develop irAEs. C) Table to show the top 20 most significant DhMRs with 5hmC inversely associated with the development of irAEs. Annotation of DhMRs was carried out with ChipSeeker. Conc_NO and Conc_YES refer to the average reads across the outcome (No progression/progression).

5.3.8.2 5-methylcytosine cfDNA profiles and clinical outcome

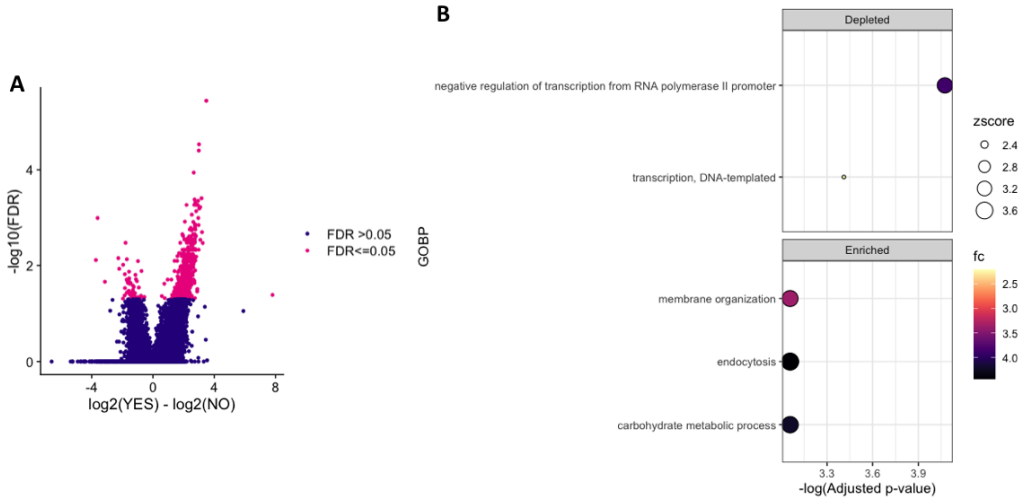
I then next explored whether 5mC cfDNA profiles were associated with clinical outcome. Differential enrichment analysis was performed using *DiffBind* as previously described.

At baseline (C1) there were 851 DMRs significantly associated with six-month progression in MM patients treated with ICB (24 patients, FDR<0.05). 51 DMR were associated with lack of progression at six months, and 799 DMRs associated with progression at six months (Figure 89).

Genes exhibiting 5mC enriched regions associated with disease progression at six months included include coiled-coiled domain gene *CCDC130*, a highly conserved protein which has a role in spliceosome formation[538] and *ARHGAP26*, which has an inhibitory role on the Rho GTPase family[539]. Low *ARHGAP26* expression in cancer cells is associated with poor prognosis[539] and induction of *ARHGAP26* expression has been shown to inhibit tumourigenic behaviour such as cell proliferation, invasion and migration, and. reduce tumour drivers such as VEGF and matrix metalloproteinase expression[539].

Genes containing DMRs where 5mC enrichment was associated with lack of progression at six months include E2 ubiquitin ligase ring finger protein *RNF121*, which has previously been shown to induce NFκB signalling[540]. Other genes containing DMRs associated with six month progression include *CSRNP3*, a transcription factor induced by IL-2[541]. *CSRNP3* has been demonstrated to be a prognostic marker in renal cell carcinoma (RCC), with reduced expression associated with reduced overall survival (OS)[542].

Pathway enrichment analysis of genes labelled with DMRs associated with progression at six months revealed enrichment of endocytosis (GO:0006897, P_{adj}= 0.047), carbohydrate metabolism (GO:0005975, P_{adj}= 0.047) and membrane organisation pathways (GO:0061024, P_{adj}= 0.047). Enrichment of transcription-associated pathways (GO:0000122, P_{adj}= 0.017, GO:0006351, P_{adj} = 0.033) were identified across genes associated with regions with reduced 5mC in progressors (Figure 89B).



C

Chromosome	Start	End	Conc_YES	Conc_NO	FDR	Genomic region	Gene
chr4	190089962	190090430	1.47	5.09	0.00102	Distal Intergenic	*DBET*
chr2	165658922	165659716	1.25	3.04	0.00333	Intron	*CSRNP3*
chr1	148385340	148386251	0.58	2.85	0.00701	Promoter (2-3kb)	*MIR6077*
chr10	51345515	51346164	0.77	2.48	0.00741	Intron	*MIR605*
chr11	54258466	54258702	0.23	3.96	0.00769	Distal Intergenic	*OR4C46*
chr15	87770293	87770676	2.68	3.67	0.00808	Distal Intergenic	*NTRK3*
chr9	1687992	1688251	1.14	3.09	0.00964	Distal Intergenic	*SMARCA2*
chr16	78465424	78466916	1	3.21	0.0117	Intron	*WWOX*
chr1	19859322	19859843	2.83	3.58	0.013	Distal Intergenic	*OTUD3*
chr11	45877252	45877500	1.42	2.99	0.0134	Intron	*CRY2*

D

Chromosome	Start	End	Conc_YES	Conc_NO	FDR	Genomic regionn	Gene
chr9	123483441	123485745	4.12	0.64	3.63E-06	Promoter (<=1kb)	*MIR7150*
chr9	123472917	123473445	3.59	0.59	2.94E-05	Intron	*DENND1A*
chr9	126744075	126744644	2.89	-0.1	3.96E-05	Distal Intergenic	*ZBTB43*
chr9	126717159	126717509	2.84	0.18	1.14E-04	Distal Intergenic	*ZBTB43*
chr9	128994622	128996401	3.39	0.23	3.91E-04	Promoter (<=1kb)	*NUP188*
chr9	127485703	127486786	3.62	0.91	4.14E-04	Exon	*LRSAM1*
chr9	127380888	127382069	4.29	1.35	4.40E-04	Intron	*GARNL3*
chr9	129266692	129267604	2.83	0.07	4.80E-04	Distal Intergenic	*LOC101929331*
chr9	124307452	124308632	3.56	0.67	5.42E-04	Intron	*NEK6*
chr9	126409541	126410661	3.64	0.72	5.42E-04	Promoter (1-2kb)	*NRON*

Figure 89: DMRs and six-month progression

A) Volcano plot to show DMRs associated with six-month progression (24 patients, pink points infers FDR<0.05). B) Dotplot to show GOBP pathways enriched across genes associated with DMRs. 'Depleted' GOBP pathways infers that DMR where 5mC shows reduced 5mC labelling with six-month progression, 'enriched' GOBP set indicates genes where 5mC is more enriched in progressors. C) Table to show top 10 most significant DMRs with higher 5mC enrichment associated with lack of disease progression at six months. D) Table to show top 10 significant DMR with higher 5mC enrichment associated with progression at six months (24 patients, FDR<0.05). Conc_NO and Conc_YES refer to the average read across the outcome (No progression/progression).

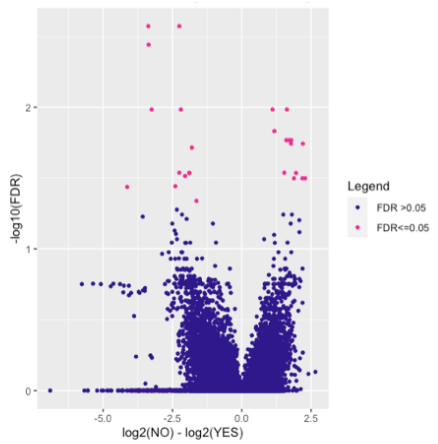
There were 26 DMRs at baseline associated with the development of irAEs, including 13 demonstrating greater 5mC enrichment associated with the development of irAEs, and

13 with lower 5mC enrichment associated with irAEs (*edgeR* analysis, 24 patients, FDR<0.05) (Figure 90).

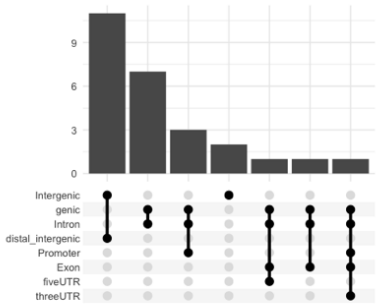
Genes containing DMRs at baseline where high 5mC enrichment was associated with irAEs include the tumour suppressor Ras-GAP protein gene *DAB2IP*, which is involved in the regulation of the innate immune system, inflammatory responses, apoptosis and cell migration[543] (Figure 90C). *DAB2IP* has a role in mediating TNF, NFκB signalling, PI3K/AKT signalling and interferon signalling, and has previously been reported to be inactivated via methylation in patients with cancer[543]. Other genes include the serine/threonine-protein kinase *NEK6*, which has a role in cell cycle regulation[544] and the PHD finger protein gene *PHF19* which has a role in haemopoietic stem cell differentiation[545].

Genes containing DMR with 5mC-enrichment associated with the absence of irAEs include ADP-ribosyltransferase 1 gene *ART1*, which is involved in post-translation modification influencing a wide range of cellular processes including cell migration, proliferation and inflammation[546]; *SEMA4B*, a semaphorin which has previously been shown to regulate basophil responses[547], and the DNA topoisomerase *TOP2A*, which has a role in regulating genomic stability[548].

A



B



C

Chromosome	Start	End	Conc_No	Conc_Yes	FDR	Genomic Annotation	Gene
chr9	123284362	1.23E+08	0.42	3.79	0.00267	Distal Intergenic	*STRBP*
chr3	159865663	1.6E+08	1.53	3.78	0.00267	Intron	*SCHIP1*
chr9	121700996	1.22E+08	0.64	4	0.00361	Promoter (1-2kb)	*DAB2IP*
chr9	116651607	1.17E+08	0.39	2.58	0.0104	5' UTR	*ASTN2*
chr9	124237347	1.24E+08	-0.17	3.08	0.0104	Distal Intergenic	*NEK6*
chr9	120881192	1.21E+08	0.7	2.5	0.0192	Intron	*PHF19*
chr9	122823885	1.23E+08	0.85	3.1	0.029	Promoter (2-3kb)	*PDCL*
chr8	56218881	56219218	1.77	3.67	0.0291	Distal Intergenic	*CHCHD7*
chr9	121543368	1.22E+08	0.75	2.63	0.0291	Distal Intergenic	*DAB2IP*
chr5	121032427	1.21E+08	1.11	3.16	0.0306	Distal Intergenic	*LOC102467226*
chr9	122784343	1.23E+08	0.4	2.79	0.0361	Distal Intergenic	*OR5C1*
chr10	41315352	41315700	-0.17	3.96	0.0366	Distal Intergenic	*LOC441666*
chr9	108560854	1.09E+08	0.4	2.03	0.0458	Distal Intergenic	*ACTL7B*

D

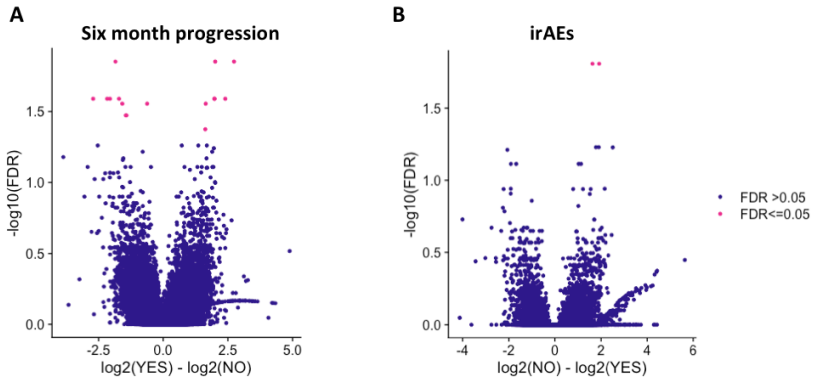
Chromosome	Start	End	Conc_No	Conc_Yes	FDR	Genomic anotation	Gene
chr15	90135034	90136605	5.07	3.96	0.0104	Intron	*SEMA4B*
chr6	108869243	1.09E+08	3.27	1.64	0.0104	Promoter (<=1kb)	*ARMC2*
chr14	92743291	92744062	4.3	3.12	0.0147	Intron	*LGMN*
chr4	75818068	75819216	4.2	2.6	0.017	Distal Intergenic	*USO1*
chr11	3653463	3654158	3.53	1.74	0.017	Intron	*ART1*
chr7	41400264	41400698	3.44	1.73	0.017	Distal Intergenic	*INHBA*
chr17	5350121	5350648	3.86	2.08	0.0181	Exon	*RABEP1*
chr16	4841923	4843432	3.51	1.31	0.0181	Promoter (<=1kb)	*GLYR1*
chr17	67066350	67067617	4.29	2.75	0.029	Downstream (2-3kb)	*CACNG1*
chr19	5972543	5973442	4.01	2.05	0.0291	Intron)	*RANBP3*
chr15	97784379	97785343	3.72	1.54	0.0318	Intron	*LINC00923*
chr1	150031977	1.5E+08	3.26	1.38	0.0318	Distal Intergenic	*OTUD7B*
chr17	40386169	40387047	2.96	0.67	0.0318	Downstream (1-2kb)	*TOP2A*

Figure 90: DMR and irAEs side effects

A) Volcano plot to show DMRs at baseline associated with irAEs. Regions significantly differentially methylated are labelled in pink (FRD<0.05). B) Upsetplot to show redundancy in genomic annotation across all significantly DMRs. C) Table to show all significant DMR at baseline associated with irAEs (edgeR analysis, 24 patients, FDR<0.05) where higher 5mC enrichment associated with irAEs. D) Table to show all significantly DMR at baseline associated with irAEs (edgeR analysis, 24 patients, FDR<0.05) with higher 5mC enrichment associated with no irAEs. Conc_No and Conc_Yes denotaes average read count across that DMR in those who do not develop irAEs 'Conc_No' versus those who do 'Conc_Yes'.

Following ICB, there were 16 DMRs associated with progression at six months (*edgeR* analysis, 24 samples, FDR <0.05) (Figure 91A&C). These included seven regions showing greater 5mC enrichment in those who progressed, and nine with greater enrichment in those that did not. These DMRs were found within genes with kinase activity such as *ROR2* and *NME1,* as well as genes such as *TUSC3* and several *LINC* genes.

Only two DMRs post-treatment were associated with irAEs (*edgeR* analysis, 24 C2 samples, FDR <0.05). These were found within *LYRM9* and *CHN2* genes (Figure 91B&D).



C

Chromosome	Start	End	Conc_YES	Conc_NO	FDR	Genomic region	Gene
chr12	55493898	55494871	2.98	0.24	0.014	Promoter (1-2kb)	*OR6C68*
chr3	138945039	138945575	2.57	0.56	0.014	Promoter (1-2kb)	*LINC01391*
chr1	207303565	207303958	1.63	3.47	0.014	Intron	*CD55*
chr6	140604948	140606013	2.68	4.39	0.0257	Intron	*MIR4465*
chr14	90817499	90817744	2.68	0.29	0.0257	Promoter (1-2kb)	*TTC7B*
chr9	91928339	91929408	2.81	0.84	0.0257	Intron	*ROR2*
chr10	39939649	39939870	0.95	3.66	0.0257	Distal Intergenic	*ACTR3BP5*
chr17	51157183	51158440	1.56	3.62	0.0257	Promoter (2-3kb)	*NME1*
chr9	109277695	109278331	3.15	1.17	0.0257	Intron	*EPB41L4B*
chr8	15890478	15890953	0.75	2.92	0.0257	Distal Intergenic	*TUSC3*
chr22	35794287	35794548	2.85	1.21	0.0279	Intron	*RBFOX2*
chr4	53237348	53239315	5.18	5.79	0.0279	Intron	*SCFD2*
chr4	154788369	154788686	1.83	3.41	0.0279	Intron	*RBM46*
chr14	26383843	26384076	2.51	3.95	0.0337	Distal Intergenic	*NOVA1*
chr15	66991970	66992345	3.03	4.45	0.0337	Promoter (1-2kb)	*LOC102723493*
chr16	54226418	54227073	3.69	2.06	0.0423	Distal Intergenic	*LINC02169*

D

Chromosome	Start	End	Conc_YES	Conc_NO	FDR	Genomic region	Gene
chr17	27889039	27889557	3.89	2.26	0.0156	Promoter (2-3kb)	*LYRM9*
chr7	29431549	29431890	3.12	1.2	0.0156	Intron	*CHN2*

Figure 91: Differentially methylated regions (DMRs) post-treatment and clinical outcome
A) Volcano plot showing DMRs post-treatment associated with progression at six months (edgeR analysis, 24 patients, FDR <0.05). B) Volcano plot to show DMRs at C2 associated with irAEs prior to cycle 5 (edgeR analysis, 24 patients, FDR <0.05). C) Table to show all DMRs post-treatment associated with six-month progression. D) Table to show all DMRs post-treatment associated with irAEs. DMRs were annotated using ChipSeeker.

5.3.9 Exploring association between epigenetic cfDNA profiles and circulating immune expression profiles

As cfDNA comprises predominantly leukocyte DNA[208] it was hypothesised that epigenetic cfDNA profiles may be correlated with expression profiles of circulating $CD8^+$ T cells and monocytes. A matrix of 5hmC and 5mC enrichment was created by assigning sequenced reads to genes using *StringTie* to create an expression-type matrix as described in section 2.9. Bulk RNA-seq data of monocytes as previously described and analysed in Chapter Three were used in this analysis. Bulk RNA-seq data from $CD8^+$ T cell

samples which had been isolated and prepared simulteously with the monocyte samples as described in section 2.3 and as previously reported[40,71] were also used in this analysis. Samples with both epigenetic and bulk RNA-seq available were selected for integrated analysis. The top 20 genes exhibiting the most 5hmC or 5mC enrichment were identified (derived from data adjusted for batch effect with *ComBat-seq* and normalised with *DESEq2*) to create a 'high 5hmC enrichment score' ('H-5hmC-S') gene set and 'high 5mC enrichment score' ('H-5mC-S') gene set (see Figure 92A&B for gene sets). The geometric mean score of H-5hmC-S and H-5mC-S enrichment per sample was calculated. *DESeq2* was then used to identify transcripts in both peripheral monocytes and $CD8^+$ T cells associated with the H-5hmC-S and H-5mC-S.

No transcripts in monocytes or $CD8^+$ T cells were associated with H-5hmC-S across either untreated (C1) samples (18 monocyte samples, 21 $CD8^+$ T cells samples), treated (C2) samples (14 $CD14^+$ samples, 22 $CD8^+$ T cell samples), or across the cohort when treated and untreated samples were analysed together, controlling for cycle (*DESeq2* controlling for age and sex, P_{adj} <0.05) (data not shown).

The analysis was repeated with 5mC data, exploring association between H-5mC-S and transcriptomic expression within monocytes and $CD8^+$ T cells. Transcriptomic expression in peripheral monocytes and $CD8^+$ T cells associated with the H-5mC-S was identified using *DESeq2* (controlling for age, sex and where appropriate, cycle). There was no association between H-5mC-S and monocytes found when a condition-wise analysis was performed (*DESeq2* analysis, 19 C1 samples, 20 C2 samples). When the analysis was run across both treated and untreated samples controlling for cycle, two transcripts were associated with H-5mC-S (19 C1 samples, 20 C2 samples, P_{adj} <0.05) (see Appendix HH for sample details). These were *FLNB* (log2FC = 2.92, P_{adj}= 0.00085*)* and *FBX8* (log2FC = 2.23, P_{adj}= 0.0085), the expression of both of which were positively associated with H-5mC-S (Figure 92E). *FBX8* is an E3 ubiquitin ligase which has been shown to have a role in colorectal cancer tumourigeneisis via its silencing preventing degradation of GSTP1[549]. When association between H-5mC-S and $CD8^+$ T cell expression was explored, one transcript was found when *DESEq2* analysis was run across both treated and untreated samples together, controlling for cycle - *SLAMF8* (log2FC= -3.75, P_{adj} = 0.021) (22 C1 samples, 23 C2 samples, *DESeq2* analysis, P_{adj} <0.05) (Figure 92, see Appendix II for

sample details). Pre-treatment, there were no transcripts associated with H-5mC-S (*DESeq2* analysis, $P_{adj}<0.05$). However, in post-treatment samples there were five transcripts associated with H-5mC-S (Figure 92). These included the chemokine receptor *CXCR3* ($log_2FC= -3.83$, $P_{adj}= 0.028$), which is expressed by effector T cells and T_H1 $CD4^+$ cells and which binds ligands CXCL9, CXCL10 and CXCL11 to activate mitogenic signalling and facilitate T_H1 cell maturation and migration to inflammatory sites[550] (Figure 92). Other transcripts associated with H-5mC-S include chromosome alignment maintaining phosphoprotein 1 *CHAMP1* ($log_2FC= -1.95$, $P_{adj}= 0.028$) and *CSTF1* ($log_2FC= -1.87$, $P_{adj}= 0.028$), which inhibits mRNA cleavage[551].

Whilst the association between epigenetic cfDNA and immune expression profiles demonstrated here are limited, the association between 5mC cfDNA enrichment scoring and immune-related genes such as *CXCR3*, which has a prominent role in T cell activation, is in keeping with cfDNA origin from circulating leukocytes and may reflect association between epigenetic cfDNA and patterns of immune activation. Whilst beyond the scope of this thesis, further in-depth characterisation of association between circulating epigenetic and immune expression profiles would be of value to chatacterise the potential of epigenetic cfDNA profiling for investigating and monitoring peripheral immune responses.

A

H-5hmC-S Genes	
RP11-392C20.1	PPWD1
LINC00992	FGF12-AS1
AC093732.1	RP11-85D18.1
RP11-429O1.1	RNU4-56P
AC078927.1	GLULP6
RNU4-42P	RNU6-98P
RP11-456O19.3	RP11-632P5.1
AC002472.11	FBXL5
DPYSL5	POU5F1P4
TTC7A	SNX7

B

H-5mC-S Genes	
ZNF501	B3GALNT1P1
SCN11A	RNU2-39P
SRSF9P1	STAT4
GUCY1A3	MIR29A
MTCP1	PCDHB9
IQCA1-AS1	RNU6-644P
AC006369.2	RP1-29C18.10
NOL9	RP11-449H3.1
CCBE1	AC093906.1
ATP6V1E1P1	AC012456.3

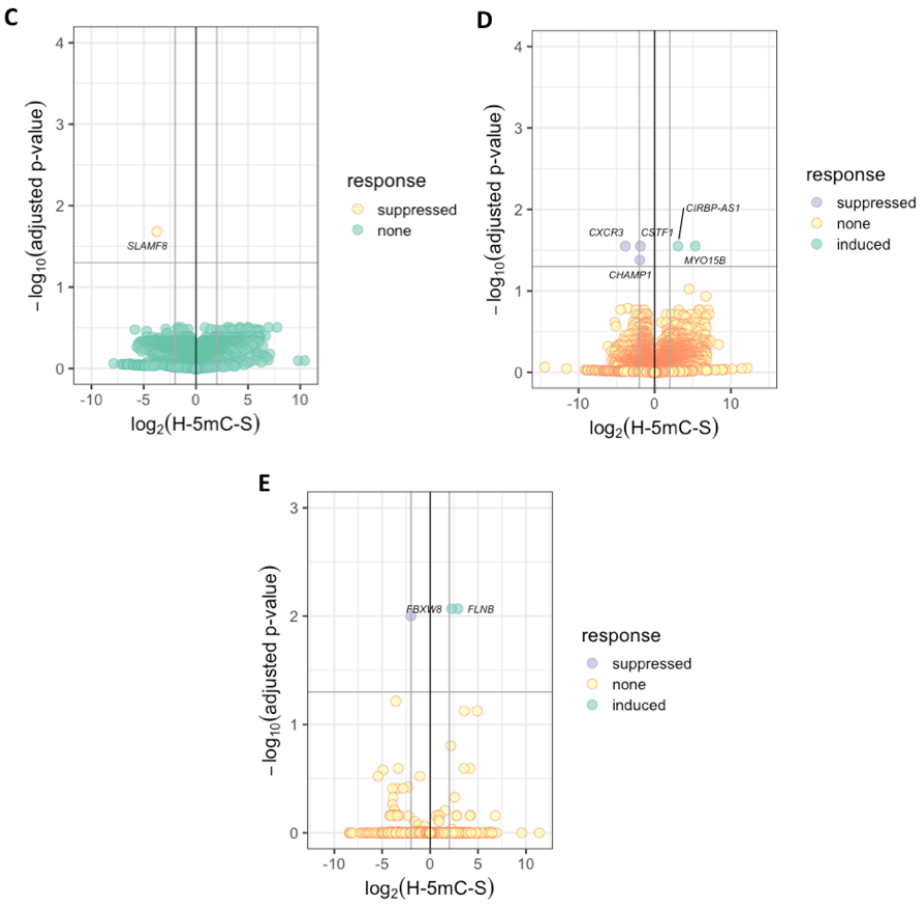


Figure 92: Integration of epigenetic cfDNA profiles with circulating immune expression profiles

Tables to show genes comprising: A) H-5hmC-S and B) H-5mC-S scores. These gene sets are the top 20 genes most enriched for 5hmC or 5mC respectively across both C1 and C2 samples. To identify these 5hmC/5mC enriched genes, read data was assigned to gene sites to create an expression matrix of 5hmC and 5mC reads. The data was adjusted for batch effect using ComBat-seq and normalised using DESEq2. The mean enrichment across all C1 and C2 samples were identified (23 C1 and 22 C2 for 5hmC, 24 C1 and 24 C2 samples for 5mC) and the top 20 most enriched genes taken forward to create the gene sets. C) Volcano plot to show $CD8^+$ T cell transcripts associated with H-5mC-S across untreated and treated samples (n=22 C1 samples, n=23 C2 samples) (DESeq2 analysis, P_{adj} <0.05). D) Volcano plot to show $CD8^+$ T cell expressed transcripts associated with H-5mC-S at C2 (DESEq2 analysis, 23 C2 patients). E) Volcano plot to show association between H-5mC-S and transcriptomic expression in peripheral monocytes across untreated and treated samples (19 C1 samples, 20 C2 samples, DESeq2 analysis).

5.4 Discussion and Conclusions

A summary of the key findings in this chapter are as follows:

- 5hmC and 5mC modifications in cfDNA are detectable and correspond to most genic regions
- Both 5hmC and 5mC-labelled genes are enriched for pathways including GTPase activity, protein phosphorylation and signal transduction. Pathways that show 5hmC-wise enrichment include apoptosis and oxidation-reduction, whilst pathways enriched from genes showing 5mC-wise enrichment include cell adhesion and ion transport. This suggests that 5hmC and 5mC modifications have a role across key biological processes.
- 5hmC and 5mC are enriched across the gene body and depleted across intergenic regions in cfDNA, corresponding to the known sites of these epigenetic changes at the genomic level.
- Both 5hmC and 5mC are enriched across CpG regions, particularly within CpG shores and shelves.
- MM is associated with reduced methylation in cfDNA across promoters and within CpG regions, whilst 5hmC is enriched across most intragenic regions. This may reflect differential immune or tumour activation states in patients with melanoma.
- ICB is associated with reduced promoter and CGI methylation and 5hmC enrichment across intragenic regions within cfDNA, in keeping with transcriptional activation of immune and tumour-derived cell populations.
- 5hmC in cfDNA is significantly enriched across histone modifications associated with transcriptional activity and gene expression including H3K36me3, H3K27ac and H3K9me3, in keeping with 5hmC as a marker of transcriptional activity.
- Differential enrichment of 5hmC is seen across histone modification sites in MM patients and following ICB. 5hmC is depleted across histone modification sites

associated with transcriptional activation such as H3K37me3 and H3K4me1 in MM, and enriched across H3K4me3, a marker of activation, following ICB.

- Epigenetic cfDNA profiles are distinct MM patients versus healthy controls, suggesting utility as diagnostic biomarker. These profiles are in keeping with this suggested role as a marker of active cancer, as well as peripheral immune status.
- 5mC cfDNA profiles are modulated with ICB, with more prominent effects seen with cICB than sICB.
- ICB may be associated with differential methylation of CTAs in cfDNA. This suggests that ICB may have a role in modulating expression of germline antigens.
- 5hmC and 5mC cfDNA profiles are associated with clinical outcome, demonstrating the utility of this approach a potential predictive biomarker.
- Preliminary work integrating epigenetic 5hmC signatures with bulk RNA-seq expression datasets from circulating immune populations suggests that epigenetic cfDNA signatures may reflect patterns of immune activation.

In this chapter, I describe how I have used cfMeDIP and hMe-Seal protocols to successfully characterise a genome-wide map of 5mC and 5hmC modifications in cfDNA respectively within a MM cohort. Characteristic 5hmC and 5mC genomic cfDNA distributions have been demonstrated, in keeping with previous studies[273,464,473,474] with most genes exhibiting dual 5hmC and 5mC modification, suggesting a role in shared key biological pathways.

These data have facilitated novel observations regarding differential enrichment and epigenetic cfDNA profiles in MM patients, and in response to ICB. Peak calling to identify epigenetically enriched genomic regions has revealed dynamic patterns of epigenetic enrichment across functional genomic regions within cfDNA in MM, which are modulated with ICB. This approach has also allowed for characterisation of epigenetic cfDNA profiles across known histone modification sites. The distribution of 5hmC across histone modification sites is in keeping with 5hmC as a permissive marker of gene expression, and suggests that captures cfDNA is derived from both T cell and monocyte populations. Differential 5hmC and 5mC enrichment across histone modification sites

between conditions may reflect dynamic changes in histone modification, which may reflect transcriptional activation or repression, but which might in particular represent transcriptional activation following ICB.

Whilst there is increasing evidence showing distinct epigenetic profiles in cfDNA in patients across a variety of different tumour types, studies exploring melanoma-associated epigenetic cfDNA profiles are limited. In this chapter, differential enrichment analysis using the *DiffBind* package and *edge/DESeq2* analysis facilitated the identification of DMRs and DhMRs in MM patients and modulation with ICB. As previously described in studies exploring other cancer types, genes associated with biological processes such as mitogenic signalling, cell-cell adhesion and transcriptional regulation exhibited DhMRs or DMRs in cfDNA in MM. DhMR and DMRs were identified within both immune and cancer-related genes, suggesting cfDNA origin from both sources.

Whilst there is a large body of work describing the diagnostic utility of epigenetic cfDNA profiling, studies exploring predictive utility and characterising on-treatment modulation are lacking. Strikingly, prominent modulation of 5mC cfDNA profiles were observed in response to ICB. 5hmC and 5mC cfDNA profiles were also associated with clinical outcome, in particular 5mC cfDNA signatures at baseline, demonstrating the potential utility of epigenetic cfDNA profiling as a predictive marker. Whilst analyses in this chapter are limited by sample size, further exploration of utilising epigenetic cfDNA profiles as predictive and diagnostic classifiers, for example via a replication cohort, and withi integration with other clinical diagnostic/predictive variables, would be of value.

Additionally, whilst the DhMRs and DMRs identified in this chapter have, where possible, been explored in terms of directional enrichment (i.e. depleted or enriched across conditions, or with clinical outcome), the effect on transcription that these DhMRs/DMRs may have are likely to be highly variable according to gene and function region. For example, 5mC within promoter regions may have inhibitory effects at promoters, and permissive effects within the gene body. Thus, whilst this analysis demonstrates that 5hmC and 5mC cfDNA profiles may have both diagnostic and

predictive utility, more detailed exploration of the functional association of these profiles to identify underpinning mechanisms would be of value.

Notably, numerous DMRs were found within recognised and potential CTAs. It is well-recognised that the expression of many CTAs is under epigenetic regulation. CTAs have pro-tumourigenic effects and are also highly immunogenic[517]. There has been a large body of work exploring the utility of cancer vaccines in targeting CTA, for example those utilising NY-ESO01 to enhance anti-tumoural T cell responses[552]. The observations in this chapter suggest that ICB may be associated with dynamic epigenetic modification across CTAs. It can be hypothesised that this may be associated with CTA gene expression, which might enhance immunogenicity. Characterisation of CTAs within cfDNA is to date limited, and epigenetic CTAs profiling within cfDNA not previously described. It would be interesting to look at the relationship between ICB and CTA expression in target tissue in future studies, with integration between this and peripheral cfDNA observations.

In view of the likely immune-based profiles and dynamic changes in enrichment across gnomic regions in keeping with dynamic transcriptional changes, integration of epigenetic cfDNA and bulk RNA-seq dataset was attempted, with preliminary work demonstrating interesting association between 5mC and monocytes and $CD8^+$ T cell expression profiles. More detailed characterisation of the association between epigenetic cfDNA profiles and immune-based expression profiles would be of value in elucidating the functional role of the observed patterns of epigenetic enrichment described in this chapter.

This work demonstrates the feasibility of epigenetic cfDNA profiling using techniques such as hMe-Seal and cfMeDIP in patients with MM, overcoming issues associated with conventional bisulfite sequencing techniques which require large input quantities. Overall, these data suggest that characterisation of epigenetic cfDNA profiles has utility as a potential diagnostic and predictive biomarker in patients with MM treated with ICB, and may potentially provide near real-time insights regarding on-treatment dynamic modification of tumour and immune profiles in response to ICB.

6 Discussion

In this thesis I describe novel findings regarding immune and cfDNA profiles in patients with MM treated with ICB. Here I discuss the key findings, contextualising them with existing literature, and highlighting areas for future work.

An elevated peripheral monocyte count is negatively prognostic in cancer[40,163–165]. In this thesis I find association between peripheral monocyte count at baseline and transcriptional profiles associated with chemotaxis, blood coagulation and angiogenesis, which suggest a mechanism for the prognostic and predictive effects of circulating myeloid populations. It also suggests that these effects are not simply quantative, but may be due to distinct expression profiles and associated function.

I find that that circulating monocytes in MM patients exhibit enrichment of pro-angiogenic pathways and those associated with innate immune signalling, including phagocytosis and response to LPS. In keeping with existing studies[176,178], chemotaxis and cell migration expression profiles were induced, whilst expression of MHC class II antigen presentation pathways were downregulated[179]. Chemotaxis-associated antigens are typically induced in monocytes in cancer, and facilitate monocyte migration to tumour sites[176,178]. Upregulation of chemotaxis markers in myeloid populations has been shown to be associated with less favourable prognosis[178,553,554]. This work suggests that in MM a pro-tumourigenic phenotype is induced in circulating monocytes.

scRNA-seq analysis demonstrated that many of the MM-associated pathways in the bulk RNA-seq dataset were induced across both classical and $CD16^+$ monocyte populations. However, pathways such as wound healing and extracellular matrix organisation were selectively enriched in $CD16^+$ monocytes. This may reflect acquisition of a pro-tumourigenic expression profile facilitating tissue infiltration and differentiation into a TAM population. Whilst expression profiling across bulk and scRNA-seq datasets suggests that MM is associated with monocyte immaturity and classical-type expression profies, trajectory analysis suggests that this effect may be limited to the larger, classical monocyte compartment. Further work to characterise the predictive and prognostic

role of myeloid immaturity, or cycling monocyte populations as opposed to simply classical populations, would be informative.

Consistent with other studies demonstrating the potential diagnostic utility of myeloid expression profiles[193], LDA analysis showed that MM-associated myeloid expression profiles have potential utility as a diagnositic classifier. Further work integrating this gene set with other clinical biomarkers, such as serum lactate dehydrogenase (LDH)[555], may increase the efficacy of this classifier model.

I demonstrate that ICB is associated with transcriptomic modulation in circulating monocytes, with induction of MHC class II antigen presentation, type I and II interferon signalling and complement pathways. This is in keeping with acquisition of a pro-inflammatory phenotype which may have co-stimulatory effects across both innate and adaptive immune pathways. ICB was associated with dynamic effects across monocyte sub-populations, however classical monocytes demonstrated acquisition of a non-classical monocyte-associated transcriptional profile, with induction of MHC class II, antigen presentation, complement production, and IFN-signalling expression profiles. Transcriptiomic modulation within the bulk RNA-seq data may therefore predominantly reflect maturation of the classical compartment, although this response appears to be variable across classical monocyte subsets. This transcriptional modulation is not associated with change in monocyte subset size, which remain constant following ICB.

In this thesis, scRNA-seq across a large monocyte population has allowed for detailed dissection of transcriptomic profiles of conventional monocyte groups. Classical monocytes prominently exhibited enrichment for transcripts associated with chemotaxis, innate immune response, and angiogenesis, whilst $CD16^{+}$ monocytes demonstrated a prominent pro-inflammatory transcriptomic phenotype, characterised by high expression of anti-viral response pathways, TNF, MHC class II antigen presentation and T cell costimulatory pathways. Whilst classical monocytes show expression profiles in keeping with prominent migratory capacity and chemotactic function, $CD16^{+}$ monocytes exhibit profiles demonstrating prominent IFN response and T cell costimulatory effects. This provides clarity regarding subset-wise monocyte transcriptomic phenotyping, as many studies report conflicting results.

In this thesis I identify heterogeneity within the classical monocyte compartment, which has not previously been identified. Goudot et al.[136] identified two monocyte subsets with Drop-Seq scRNA-seq; a $CD16^-$ classical and $CD16^+$ subset, and subsequent high-resolution clustering of the $CD16^-$ compartment did not reveal any further distinct classical subsets. Villani et al.[104] also identified a homogenous classical subset (denoted 'Mono1'). However, these studies are limited by small sample sizes. Goudot et al. performed unsupervised clustering across only 429 monocytes, and Villani et al. across 337 cells. In this thesis dissection of heterogeneity within a large cell population (n=24,457 cells) has allowed for dissection of classical monocyte heterogeneity. These classical subsets were characterised by variable expression profiles, which are likely to reflect differential activation and maturation states. For example, some classical subsets were characterised by classical-type expression profiles (classical subset 1 and 4), whilst others exhibited a more pro-inflammatory non-classical type phenotype (e.g classical subset 6). Classical subset 8 may represent a MPA population and subset 5 a degenerating, exhausted population. The variation within the classical compartment may underpin the non-concordance in the literature between studies describing monocyte profiling and function. Also notable was the predominance of a single or small number of classical subsets comprising each patient sample. Inter-individual variation regarding classical activation/maturation states has not previously been described. Further work to validate these identified classical subsets, and to explore variability of these classical subpopulations in a larger MM cohort and across multiple healthy donors, would be of value.

In keeping with existing studies which have used scRNA-seq to characterise monocyte subsets[104,134,135], high-level unsupervised clustering approaches did not readily distinguish a robust intermediate population distinct from classical and non-classical monocyte compartments. Instead, clustering within the $CD16^+$ population identified a high MHC Class II^{hi}*FCGR3A*^{low} population ('$CD16^+$ cluster 1'), which may represent an intermediate, or intermediate-like cell group. The transcriptional intersection between these populations suggests that it may be more appropriate to consider intermediate and non-classical populations as variable activation states of the same cell group.

Previous studies have used radioactive labelling in healthy individuals to explore differentiation states within the monocyte compartment[149]. I performed trajectory analysis across the scRNA-seq dataset and demonstrated monocyte differentiation from an immature classical population to a more mature $CD16^+$ non-classical type population, characterised by a pro-inflammatory transcriptional phenotype. Trajectory analysis reveals additional complexity, with a possibly degenerate/exhausted classical population representing an additional branch and termination point, and the $CD16^+$ monocyte compartment exhibiting dual differentiation fates. Monocytes are a reservoir for some TAMs which have tumour modulatory effects[156,158,159]. Whilst beyond the scope of this thesis and dataset, it would be interesting to characterise the trajectory across monocyte and tissue-based macrophage populations (such as TAMs) to more fully explore patterns of differentiation and activation.

In this thesis I have also explored the predictive role of monocyte expression signatures. In chapter three, of particular note was the association between baseline transcriptomic profiles and irAEs. *MKI67*, a marker of proliferative activity, was found to be negatively predictive of irAEs at baseline, with irAEs known to be associated with a more favourable clinical response to ICB in MM patients[40]. Our group have demonstrated that baseline monocyte count is negatively predictive of OS within this cohort[40]. This suggests both that proliferating monocytes at baseline have an immunosuppressive effect, and the development of irAEs is predicted by the pre-treatment immune environment. Strikingly, a non-classical expression signature ('$CD16^+$ cluster 3') at baseline was associated with both prolonged OS and PFS with ICB. This expression profile may reflect baseline myeloid pro-inflammatory and co-stimulatory function, which contributes to immune recognition with anti-tumourigenic effects. Whilst this demonstrates a clear link between the baseline myeloid immune environment and response to ICB, mechanisms underpinning this association require further dissection.

Additionally, following ICB, intermediate population size and an intermediate-type expression signature were associated with reduced OS post-treatment. The association between an intermediate-type expression profile and OS was seen at C2, whilst intermediate subset size (as determined by flow cytometry) was associated with OS at the later C4. Typically, expression of MHC class II antigens are induced by cytotoxicity

and anti-viral T cell responses, which are both associated with more favourable response to ICB[408]. In this context it might be hypothesised that the persistence of a co-stimulatory myeloid signature reflects persistence of a tumour population and response failure. However, further work to explore the mechanisms underpinning this association are required.

There are limitations to the approaches I have used to characterise circulating monocyte populations. Whilst allowing for dissection of monocyte subsets, profiles and function, the data I present here is expression-based. Thus, whilst inferences can be made regarding associated biological functions and pathways, these do not directly describe antigenic profiles or demonstrate biological function. Whilst there is good concordance between transcriptomic and antigen profiles in peripheral monocytes[97], it would be informative to validate some of the findings described in chapters three and four. For example, techniques such as flow cytometry might be of use to validate antigenic expression of the expression markers that define classical or $CD16^+$ subsets described in this thesis, as well as the change in activation and differentiation state with ICB, and association with clinical outcome. In addition, I used magnetic-activated cell sorting (MACS) to isolate $CD14^+$ cells, and whilst this is associated with a high purity (>90%)[556], there may be some level of contamination in the isolated $CD14^+$ population. Whilst this can be addressed with robust filtering techniques in the scRNA-seq analysis, this may be more problematic within the bulk RNA-seq dataset. This isolation technique may also affect cellular transcriptional profiles. Whilst mitigated in part by the uniform application of this technique across all patient samples, this may affect the expression-based profiling described in this thesis.

ICB is characterised by effector $CD8^+$ T cell activation[408]. The induction of pro-inflammatory, antigen presentation and anti-viral type expression profiles in circulating monocytes I describe here may be secondary to $CD8^+$ effector T cell activation. These myeloid-based profiles in turn may be co-stimulatory of the T cell compartment. Recent work from this group demonstrate that $CD8^+$ T cell cytotoxicity profiles and T cell clone size are both key predictors of response to ICB[408]. Integration of the monocyte expression profiles I describe in this thesis with $CD8^+$ T cell profiles, for example by exploring monocyte transcriptomic profiles associated with $CD8^+$ T cell cytotoxicity and

TCR clonal diversity, may characterise mechanisms underpinning the observations described in chapter three and four, and provide insights into the co-stimulatory effects of $CD8^+$ T cells and myeloid populations.

In recent years the development of library preparation assays for low input cfDNA has facilitated the characterisation of mutational and epigenetic cfDNA profiles across many different cancer cohorts (Figure 5). These studies have demonstrated the potential utility of epigenetic cfDNA profiles as diagnostic and prognostic biomarkers. Studies characterising epigenetic cfDNA profiles in melanoma cohorts have been limited, and the data presented in this thesis are one of the few describing epigenetic cfDNA profiles in such a cohort.

I find that in keeping with other published work[278,279], 5hmC in cfDNA is enriched across gene bodies, and depleted at intergenic regions. Genome-wide epigenetic cfDNA profiling reveals the abundance of 5hmC and 5mC modifications across the genome. I find that epigenetic cfDNA profiles are distinct in MM patients versus healthy controls, are modulated by ICB, and both 5mC and 5hmC cfDNA profiles are associated with clinical response to ICB. This work demonstrates that epigenetic cfDNA profiling may have utility as a diagnostic and predictive biomarker in MM, allowing for assessment of immune and tumour-based profiles in a biopsy-free manner.

Enrichment of 5hmC across T cell and monocyte histone modifications associated with gene activation is supportive of cfDNA immune cell origin. Gene sets associated with DhMR and DMRs also suggest that epigenetic cfDNA profiles represent both tumour-based and immune signatures. However, further dissection of the association between these epigenetic cfDNA signatures with both immune expression and tumour profiles would be informative to determine the utility of epigenetic cfDNA characterisation for monitoring immune and tumour responses.

Limitations of the epigenetic cfDNA profiling work I present in chapter five include the small cohort size, particularly regarding healthy donor samples, and for the pairwise analysis of ICB-associated 5hmC cfDNA modulation. This may limit identification of MM-associated epigenetic profiles and on-treatment 5hmC responses. Future work to expand the cohort would allow for replication of these findings and would facilitate

characterisation of these profiles as inputs to a clinical diagnostic or predictive classifier. Integration of epigenetic cfDNA profiles with both monocyte and $CD8^+$ T cell expression profiles, as well as other predictive biomarkers (e.g. monocyte count) could be explored to develop a robust clinical diagnostic or predictive tool.

A novel finding in chapter five is that many of the identified DMRs/DhMRs were found within germ cell-associated genes, which may represent cancer testis antigens (CTAs). CTAs are germ-cell associated genes first characterised in melanoma and which are prominently immunogenic[517]. Differential epigenetic modification within these genes following ICB may reflect a novel way in which ICB modulates immunogenicity. Epigenetic modification of CTAs in response to ICB has never previously been described and further work validating this observation and exploring the possible effects on gene expression and immunogenicity, are required.

A limitation of this work is the inference that can be made regarding the biological effect of epigenetic modulation in MM and ICB. There is variable association between epigenetic modification and transcriptomic expression. For example, whilst 5hmC is broadly associated with transcriptomic expression, this association is variable. 5mC is also variably associated with gene expression, typically repressive at promoter regions, but broadly associated with gene expression within the gene body[232,233]. Annotation of DhMRs and DMRs within genic regions and consideration of the direction of epigenetic modulation and pathway enrichment analysis provides some insight into possible functional effect. However, it is likely to provide limited insight into directional or biological effects of differential hydroxymethylation or methylation. Further work to characterise the association between epigenetic modulation in cfDNA, and immune cell expression and function, would be of value.

In conclusion, this thesis provides novel insights into peripheral myeloid and epigenetic cfDNA profiles in MM and following ICB. This work demonstrates that both immune and epigenetic cfDNA profiles are distinct in MM and are modulated by ICB, and that these profiles are associated with clinical outcome. Characterising circulating monocyte and epigenetic cfDNA profiles may therefore provide novel opportunities for peripheral-based diagnosis and prognostication.

APPENDICES

RNA-seq sample metadata

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
1	C1	1	CD14	One	42	2	sICB	1	0
2	C1	2	CD14	One	53	2	sICB	0	1
3	C1	3	CD14	One	79	2	sICB	1	0
4	C1	4	CD14	One	68	2	sICB	0	1
5	C1	5	CD14	One	74	2	sICB	1	0
6	C1	6	CD14	One	88	2	sICB	1	0
7	C1	7	CD14	One	61	1	sICB	1	0
8	C1	8	CD14	One	78	2	sICB	1	0
9	C1	9	CD14	One	72	1	sICB	0	0
10	C1	11	CD14	One	79	1	sICB	0	1
11	C1	12	CD14	One	61	1	sICB	1	0
12	C1	13	CD14	One	52	2	sICB	1	0
13	C1	14	CD14	One	75	1	sICB	1	1
14	C1	16	CD14	One	79	2	sICB	0	0
15	C1	18	CD14	One	77	1	sICB	1	1
16	C1	20	CD14	One	69	2	sICB	0	1
17	C1	27	CD14	One	47	1	sICB	0	0
18	C1	28	CD14	One	76	1	sICB	1	0
19	C1	30	CD14	One	73	1	sICB	1	1
20	C1	10	CD14	One	71	1	sICB	1	1
21	C1	15	CD14	One	60	1	sICB	0	1
22	C1	17	CD14	One	42	1	sICB	0	1
23	C1	19	CD14	One	29	1	sICB	1	0
24	C1	21	CD14	One	43	2	sICB	1	0
25	C1	22	CD14	One	76	1	sICB	1	0
26	C1	24	CD14	One	80	2	sICB	1	0
27	C1	25	CD14	One	67	1	sICB	*NA*	1
28	C1	29	CD14	One	59	1	sICB	1	0
29	C1	31	CD14	One	79	2	sICB	0	1
30	C1	33	CD14	One	64	2	cICB	0	1
31	C1	36	CD14	One	79	1	sICB	0	0
32	C1	32	CD14	One	87	1	sICB	1	0
33	C1	72	CD14	One	50	1	cICB	0	1
34	C1	73	CD14	One	64	1	sICB	1	0
35	C1	74	CD14	One	71	1	sICB	0	0
36	C1	75	CD14	One	67	2	sICB	0	0
37	C1	76	CD14	One	70	1	cICB	0	1
38	C1	77	CD14	One	57	1	cICB	0	1
39	C1	79	CD14	One	81	2	sICB	1	0
40	C1	80	CD14	One	82	1	sICB	0	1
41	C1	83	CD14	One	45	2	cICB	0	1
42	C1	84	CD14	One	43	2	sICB	0	1
43	C1	85	CD14	One	56	1	cICB	1	1
44	C1	94	CD14	One	81	2	sICB	0	0
45	C1	96	CD14	One	78	1	sICB	1	1
46	C1	97	CD14	One	70	2	cICB	0	1
47	C1	98	CD14	One	58	2	cICB	0	1
48	C1	99	CD14	One	62	2	cICB	1	1
49	C1	100	CD14	One	58	2	cICB	0	1
50	C1	101	CD14	One	71	2	sICB	0	0

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
51	C1	102	CD14	One	80	2	sICB	0	0
52	C1	103	CD14	One	59	2	cICB	0	1
53	C1	104	CD14	One	55	1	cICB	0	1
54	C1	105	CD14	One	29	2	cICB	0	1
55	C1	106	CD14	One	51	2	cICB	1	1
56	C1	71	CD14	One	84	2	sICB	0	0
57	C1	82	CD14	One	57	1	cICB	0	1
58	C1	86	CD14	One	65	2	cICB	0	1
59	C1	87	CD14	One	68	1	cICB	1	1
60	C1	88	CD14	One	69	1	sICB	0	0
61	C1	90	CD14	One	66	1	cICB	0	1
62	C1	91	CD14	One	47	2	cICB	1	1
63	C1	92	CD14	One	59	2	cICB	1	1
64	C1	107	CD14	Two	44	1	cICB	1	0
65	C1	108	CD14	Two	63	1	cICB	*NA*	1
66	C1	109	CD14	Two	64	1	cICB	0	1
67	C1	112	CD14	Two	51	1	cICB	0	1
68	C1	114	CD14	Two	88	1	sICB	1	0
69	C1	115	CD14	Two	60	1	cICB	1	0
70	C1	116	CD14	Two	60	1	cICB	1	1
71	C1	117	CD14	Two	68	2	cICB	0	1
72	C1	120	CD14	Two	74	2	cICB	0	1
73	C1	121	CD14	Two	63	2	cICB	0	1
74	C1	122	CD14	Two	71	1	cICB	1	1
75	C1	123	CD14	Two	86	1	sICB	0	0
76	C1	124	CD14	Two	61	1	cICB	0	1
77	C1	126	CD14	Two	74	1	sICB	0	1
78	C1	128	CD14	Two	74	2	cICB	0	1
79	C1	129	CD14	Two	74	1	cICB	0	1
80	C1	130	CD14	Two	70	1	cICB	0	1
81	C1	132	CD14	Two	48	1	cICB	0	1
82	C1	133	CD14	Two	79	1	sICB	*NA*	1
83	C1	135	CD14	Two	68	1	cICB	*NA*	1
84	C1	136	CD14	Two	75	1	sICB	*NA*	1
85	C1	137	CD14	Two	95	1	sICB	*NA*	0
86	C1	138	CD14	Two	56	2	cICB	1	0
87	C1	139	CD14	Two	60	1	cICB	*NA*	1
88	C1	140	CD14	Two	44	1	cICB	*NA*	1
89	C1	141	CD14	Two	90	2	sICB	*NA*	1
90	C1	142	CD14	Two	87	1	sICB	*NA*	1
91	C1	144	CD14	Two	55	2	cICB	0	1
92	C1	145	CD14	Two	80	1	sICB	*NA*	0
93	C1	146	CD14	Two	34	2	cICB	1	1
94	C1	148	CD14	Two	69	1	cICB	*NA*	1
95	C1	173	CD14	Three	32	2	cICB	*NA*	*NA*
96	C1	174	CD14	Three	81	1	sICB	*NA*	*NA*
97	C1	175	CD14	Three	60	1	cICB	*NA*	*NA*
98	C1	182	CD14	Three	57	1	cICB	*NA*	*NA*
99	C1	150	CD14	Three	57	2	cICB	*NA*	*NA*
100	C1	162	CD14	Three	83	2	sICB	*NA*	*NA*

Appendix A: Bulk RNA-seq cohort details (1)

Cohort details for bulk RNA-seq data. Cycle indicates treatment cycle, ID indicates patient ID, cell type indicates cell type sequenced (CD14 cells indicate monocytes), batch infers sequencing batch (one, two or three), Rx infers treatment type (sICB – single agent anti-PD-1, or cICB – combination agents anti-PD-1 and anti-CTLA-4 therapy). 6_MP infers progression at six months, AI_c5 infers development of irAEs prior to the fifth treatment cycle. For sex, 1 denotes male, 2 denotes female). For 6_MP and AI_c5, 1 denotes yes, 0 denotes no.

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
101	C1	113	CD14	Three	31	2	cICB	1	0
102	C1	119	CD14	Three	80	1	sICB	1	1
103	C1	125	CD14	Three	55	1	cICB	1	1
104	C1	153	CD14	Three	58	2	cICB	0	*NA*
105	C1	156	CD14	Three	71	1	cICB	0	*NA*
106	C1	157	CD14	Three	72	2	sICB	0	*NA*
107	C1	158	CD14	Three	68	1	cICB	1	*NA*
108	C1	159	CD14	Three	58	2	cICB	*NA*	*NA*
109	C1	160	CD14	Three	96	2	sICB	*NA*	*NA*
110	C1	164	CD14	Three	60	1	cICB	*NA*	*NA*
111	C1	166	CD14	Three	68	2	cICB	1	*NA*
112	C1	169	CD14	Three	44	1	cICB	1	*NA*
113	C1	171	CD14	Three	73	2	cICB	*NA*	*NA*
114	C1	172	CD14	Three	85	1	sICB	*NA*	*NA*
115	C2	1	CD14	One	42	2	sICB	1	0
116	C2	2	CD14	One	53	2	sICB	0	1
117	C2	3	CD14	One	79	2	sICB	1	0
118	C2	4	CD14	One	68	2	sICB	0	1
119	C2	5	CD14	One	74	2	sICB	1	0
120	C2	6	CD14	One	88	2	sICB	1	0
121	C2	7	CD14	One	61	1	sICB	1	0
122	C2	8	CD14	One	78	2	sICB	1	0
123	C2	9	CD14	One	72	1	sICB	0	0
124	C2	11	CD14	One	79	1	sICB	0	1
125	C2	12	CD14	One	61	1	sICB	1	0
126	C2	13	CD14	One	52	2	sICB	1	0
127	C2	14	CD14	One	75	1	sICB	1	1
128	C2	16	CD14	One	79	2	sICB	0	0
129	C2	18	CD14	One	77	1	sICB	1	1
130	C2	20	CD14	One	69	2	sICB	0	1
131	C2	23	CD14	One	59	1	sICB	0	0
132	C2	27	CD14	One	47	1	sICB	0	0
133	C2	28	CD14	One	76	1	sICB	1	0
134	C2	30	CD14	One	73	1	sICB	1	1
135	C2	33	CD14	One	64	2	cICB	0	1
136	C2	36	CD14	One	79	1	sICB	0	0
137	C2	32	CD14	One	87	1	sICB	1	0
138	C2	72	CD14	One	50	1	cICB	0	1
139	C2	73	CD14	One	64	1	sICB	1	0
140	C2	74	CD14	One	71	1	sICB	0	0
141	C2	75	CD14	One	67	2	sICB	0	0
142	C2	76	CD14	One	70	1	cICB	0	1
143	C2	77	CD14	One	57	1	cICB	0	1
144	C2	81	CD14	One	72	2	sICB	0	0
145	C2	80	CD14	One	82	1	sICB	0	1
146	C2	83	CD14	One	45	2	cICB	0	1
147	C2	84	CD14	One	43	2	sICB	0	1
148	C2	85	CD14	One	56	1	cICB	1	1
149	C2	92	CD14	One	59	2	cICB	1	1
150	C2	94	CD14	One	81	2	sICB	0	0

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
151	C2	96	CD14	One	78	1	sICB	1	1
152	C2	97	CD14	One	70	2	cICB	0	1
153	C2	99	CD14	One	62	2	cICB	1	1
154	C2	100	CD14	One	58	2	cICB	0	1
155	C2	101	CD14	One	71	2	sICB	0	0
156	C2	102	CD14	One	80	2	sICB	0	0
157	C2	103	CD14	One	59	2	cICB	0	1
158	C2	104	CD14	One	55	1	cICB	0	1
159	C2	105	CD14	One	29	2	cICB	0	1
160	C2	106	CD14	One	51	2	cICB	1	1
161	C2	71	CD14	One	84	2	sICB	0	0
162	C2	82	CD14	One	57	1	cICB	0	1
163	C2	86	CD14	One	65	2	cICB	0	1
164	C2	87	CD14	One	68	1	cICB	1	1
165	C2	88	CD14	One	69	1	sICB	0	0
166	C2	90	CD14	One	66	1	cICB	0	1
167	C2	91	CD14	One	47	2	cICB	1	1
168	C2	98	CD14	Two	58	2	cICB	0	1
169	C2	107	CD14	Two	44	1	cICB	1	0
170	C2	108	CD14	Two	63	1	cICB	*NA*	1
171	C2	109	CD14	Two	64	1	cICB	0	1
172	C2	112	CD14	Two	51	1	cICB	0	1
173	C2	114	CD14	Two	88	1	sICB	1	0
174	C2	116	CD14	Two	60	1	cICB	1	1
175	C2	120	CD14	Two	74	2	cICB	0	1
176	C2	122	CD14	Two	71	1	cICB	1	1
177	C2	124	CD14	Two	61	1	cICB	0	1
178	C2	126	CD14	Two	74	1	sICB	0	1
179	C2	128	CD14	Two	74	2	cICB	0	1
180	C2	129	CD14	Two	74	1	cICB	0	1
181	C2	132	CD14	Two	48	1	cICB	0	1
182	C2	133	CD14	Two	79	1	sICB	*NA*	1
183	C2	135	CD14	Two	68	1	cICB	*NA*	1
184	C2	136	CD14	Two	75	1	sICB	*NA*	1
185	C2	137	CD14	Two	95	1	sICB	*NA*	0
186	C2	138	CD14	Two	56	2	cICB	1	0
187	C2	139	CD14	Two	60	1	cICB	*NA*	1
188	C2	140	CD14	Two	44	1	cICB	*NA*	1
189	C2	142	CD14	Two	87	1	sICB	*NA*	1
190	C2	144	CD14	Two	55	2	cICB	0	1
191	C2	145	CD14	Two	80	1	sICB	*NA*	0
192	C2	146	CD14	Two	34	2	cICB	1	1
193	C2	148	CD14	Two	69	1	cICB	*NA*	1
194	C2	172	CD14	Three	85	1	sICB	*NA*	*NA*
195	C2	173	CD14	Three	32	2	cICB	*NA*	*NA*
196	C2	174	CD14	Three	81	1	sICB	*NA*	*NA*
197	C2	175	CD14	Three	60	1	cICB	*NA*	*NA*
198	C2	182	CD14	Three	57	1	cICB	*NA*	*NA*
199	C2	150	CD14	Three	57	2	cICB	*NA*	*NA*
200	C2	162	CD14	Three	83	2	sICB	*NA*	*NA*

Appendix B Bulk RNA-seq cohort details (2)

Cohort details for bulk RNA-seq data. Cycle indicates treatment cycle, ID indicates patient ID, cell type indicates cell type sequenced (CD14 cells indicate monocytes), batch infers sequencing batch (one, two or three), Rx infers treatment type (sICB – single agent anti-PD-1, or cICB – combination agents anti-PD-1 and anti-CTLA-4 therapy). 6_MP infers progression at six months, AI_c5 infers development of irAEs prior to the fifth treatment cycle. For sex, 1 denotes male, 2 denotes female). For 6_MP and AI_c5, 1 denotes yes, 0 denotes no.

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
201	C2	113	CD14	Three	31	2	cICB	1	0
202	C2	119	CD14	Three	80	1	sICB	1	1
203	C2	125	CD14	Three	55	1	cICB	1	1
204	C2	153	CD14	Three	58	2	cICB	0	*NA*
205	C2	156	CD14	Three	71	1	cICB	0	*NA*
206	C2	157	CD14	Three	72	2	sICB	0	*NA*
207	C2	158	CD14	Three	68	1	cICB	1	*NA*
208	C2	160	CD14	Three	96	2	sICB	*NA*	*NA*
209	C2	164	CD14	Three	60	1	cICB	*NA*	*NA*
210	C2	166	CD14	Three	68	2	cICB	1	*NA*
211	C2	169	CD14	Three	44	1	cICB	1	*NA*
212	C2	171	CD14	Three	73	2	cICB	*NA*	*NA*
213	C4	1	CD14	One	42	2	sICB	1	0
214	C4	2	CD14	One	53	2	sICB	0	1
215	C4	6	CD14	One	88	2	sICB	1	0
216	C4	3	CD14	One	79	2	sICB	1	0
217	C4	11	CD14	One	79	1	sICB	0	1
218	C4	12	CD14	One	61	1	sICB	1	0
219	C4	13	CD14	One	52	2	sICB	1	0
220	C4	14	CD14	One	75	1	sICB	1	1
221	C4	16	CD14	One	79	2	sICB	0	0
222	C4	33	CD14	One	64	2	cICB	0	1
223	C4	30	CD14	One	73	1	sICB	1	1
224	C4	27	CD14	One	47	1	sICB	0	0
225	C4	36	CD14	One	79	1	sICB	0	0
226	C4	73	CD14	One	64	1	sICB	1	0
227	C4	113	CD14	Three	31	2	cICB	1	0
228	C4	119	CD14	Three	80	1	sICB	1	1
229	C4	153	CD14	Three	58	2	cICB	0	*NA*
230	C4	157	CD14	Three	72	2	sICB	0	*NA*
231	C4	159	CD14	Three	58	2	cICB	*NA*	*NA*
232	C4	164	CD14	Three	60	1	cICB	*NA*	*NA*
233	C4	166	CD14	Three	68	2	cICB	1	*NA*
234	C4	169	CD14	Three	44	1	cICB	1	*NA*
235	HD	580	CD14	Three	52	2	NA	NA	NA
236	HD	581	CD14	Three	50	1	NA	NA	NA
237	HD	582	CD14	Three	62	2	NA	NA	NA
238	HD	583	CD14	Three	34	2	NA	NA	NA
239	HD	584	CD14	Three	54	2	NA	NA	NA
240	HD	717	CD14	One	24	2	NA	NA	NA
241	HD	724	CD14	One	47	2	NA	NA	NA
242	HD	725	CD14	One	59	2	NA	NA	NA
243	HD	726	CD14	One	51	1	NA	NA	NA
244	HD	727	CD14	One	32	2	NA	NA	NA
245	HD	728	CD14	One	28	1	NA	NA	NA
246	HD	730	CD14	One	32	2	NA	NA	NA
247	HD	731	CD14	One	26	2	NA	NA	NA
248	HD	732	CD14	One	46	2	NA	NA	NA
249	HD	735	CD14	Three	31	1	NA	NA	NA
250	HD	737	CD14	Three	58	1	NA	NA	NA

No.	Cycle	ID	Cell type	Batch	Age	Sex	Rx	6_MP	AI_c5
251	HD	738	CD14	Three	61	2	NA	NA	NA
252	HD	739	CD14	Three	31	1	NA	NA	NA
253	HD	740	CD14	Three	44	2	NA	NA	NA
254	HD	741	CD14	Three	22	2	NA	NA	NA
255	HD	742	CD14	Three	50	2	NA	NA	NA
256	HD	743	CD14	Three	36	2	NA	NA	NA
257	HD	744	CD14	Three	30	1	NA	NA	NA
258	HD	745	CD14	Three	55	2	NA	NA	NA
259	HD	746	CD14	Three	38	1	NA	NA	NA
260	HD	755	CD14	One	54	1	NA	NA	NA
261	HD	756	CD14	Three	53	2	NA	NA	NA
262	HD	760	CD14	Three	54	1	NA	NA	NA
263	HD	761	CD14	Three	54	2	NA	NA	NA
264	HD	773	CD14	Three	55	2	NA	NA	NA
265	HD	774	CD14	Three	54	1	NA	NA	NA
266	HD	776	CD14	Three	54	1	NA	NA	NA
267	HD	777	CD14	Three	34	1	NA	NA	NA
268	HD	780	CD14	Three	54	1	NA	NA	NA
269	HD	785	CD14	Three	34	1	NA	NA	NA
270	HD	794	CD14	Three	36	1	NA	NA	NA
271	HD	795	CD14	Three	36	1	NA	NA	NA
272	HD	796	CD14	Three	54	2	NA	NA	NA
273	HD	797	CD14	Three	34	1	NA	NA	NA
274	HD	798	CD14	Three	55	2	NA	NA	NA
275	HD	799	CD14	Three	31	1	NA	NA	NA
276	HD	800	CD14	One	54	2	NA	NA	NA
277	HD	811	CD14	Three	51	2	NA	NA	NA
278	HD	812	CD14	Three	56	2	NA	NA	NA
279	HD	813	CD14	Three	64	2	NA	NA	NA

Appendix C: Bulk RNA-seq cohort details (3)

Cohort details for bulk RNA-seq data. Cycle indicates treatment cycle, ID indicates patient ID, cell type indicates cell type sequenced (CD14 cells indicate monocytes), batch infers sequencing batch (one, two or three), Rx infers treatment type (sICB – single agent anti-PD-1, or cICB – combination agents anti-PD-1 and anti-CTLA-4 therapy). 6_MP infers progression at six months, AI_c5 infers development of irAEs prior to the fifth treatment cycle. For sex, 1 denotes male, 2 denotes female). For 6_MP and AI_c5, 1 denotes yes, 0 denotes no.

GOBP Pathway enrichment analysis of pathways induced in MM patients versus healthy donors

	GOBP	Pathway	FC	z score	Adjusted p-val	Direction
1	GO:0043312	neutrophil degranulation	3.2	10.1	5.40E-14	Induced
2	GO:0006954	inflammatory response	2.95	7.77	1.60E-08	Induced
3	GO:0032496	response to lipopolysaccharide	3.15	5.7	2.70E-04	Induced
4	GO:0006915	apoptotic process	1.97	4.71	1.10E-03	Induced
5	GO:0006935	chemotaxis	3.29	5.17	1.40E-03	Induced
6	GO:0045087	innate immune response	2.05	4.56	1.60E-03	Induced
7	GO:0035556	intracellular signal transduction	2.1	4.51	1.80E-03	Induced
8	GO:0042542	response to hydrogen peroxide	4.58	5.14	3.80E-03	Induced
9	GO:0051642	centrosome localization	8.73	5.98	4.10E-03	Induced
10	GO:0050728	negative regulation of inflammatory response	3.63	4.69	4.10E-03	Induced
11	GO:0002576	platelet degranulation	2.91	4.45	4.10E-03	Induced
12	GO:0006468	protein phosphorylation	1.91	4.05	4.10E-03	Induced
13	GO:0048041	focal adhesion assembly	6.11	5.18	5.00E-03	Induced
14	GO:0015701	bicarbonate transport	4.45	4.73	5.00E-03	Induced
15	GO:0070373	negative regulation of ERK1 and ERK2 cascade	4	4.6	5.00E-03	Induced
16	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	2.61	4.22	5.00E-03	Induced
17	GO:0018105	peptidyl-serine phosphorylation	2.68	4.21	5.00E-03	Induced
18	GO:0006955	immune response	1.95	3.97	5.00E-03	Induced
19	GO:0031100	animal organ regeneration	4.35	4.64	5.30E-03	Induced
20	GO:0035914	skeletal muscle cell differentiation	4.35	4.64	5.30E-03	Induced
21	GO:0030705	cytoskeleton-dependent intracellular transport	6.79	5.08	7.00E-03	Induced
22	GO:0008285	negative regulation of cell proliferation	1.9	3.7	7.00E-03	Induced
23	GO:0060326	cell chemotaxis	3.91	4.26	9.60E-03	Induced
24	GO:0006801	superoxide metabolic process	6.11	4.72	1.00E-02	Induced
25	GO:0051091	positive regulation of sequence-specific DNA binding transcription factor activity	2.85	3.89	1.00E-02	Induced
26	GO:0030593	neutrophil chemotaxis	3.44	4.04	1.10E-02	Induced
27	GO:0045944	positive regulation of transcription from RNA polymerase II promoter	1.48	3.24	1.30E-02	Induced
28	GO:0006909	phagocytosis	3.8	3.89	1.90E-02	Induced
29	GO:0045444	fat cell differentiation	3.14	3.71	1.90E-02	Induced
30	GO:0051591	response to cAMP	3.72	3.82	2.00E-02	Induced
31	GO:0048010	vascular endothelial growth factor receptor signaling pathway	3.1	3.66	2.00E-02	Induced
32	GO:0030206	chondroitin sulfate biosynthetic process	4.89	4.02	2.30E-02	Induced
33	GO:0006469	negative regulation of protein kinase activity	2.64	3.42	2.30E-02	Induced
34	GO:0016477	cell migration	2.2	3.21	2.60E-02	Induced
35	GO:0090023	positive regulation of neutrophil chemotaxis	4.37	3.68	2.90E-02	Induced
36	GO:0031623	receptor internalization	3.86	3.65	2.90E-02	Induced
37	GO:0090002	establishment of protein localization to plasma membrane	3.76	3.57	2.90E-02	Induced
38	GO:0060291	long-term synaptic potentiation	3.67	3.49	2.90E-02	Induced
39	GO:0045429	positive regulation of nitric oxide biosynthetic process	3.67	3.49	2.90E-02	Induced
40	GO:0007189	adenylate cyclase-activating G-protein coupled receptor signaling pathway	3.67	3.49	2.90E-02	Induced
41	GO:0007190	activation of adenylate cyclase activity	3.67	3.49	2.90E-02	Induced
42	GO:0007200	phospholipase C-activating G-protein coupled receptor signaling pathway	3.1	3.46	2.90E-02	Induced
43	GO:0007160	cell-matrix adhesion	2.66	3.29	2.90E-02	Induced
44	GO:0051056	regulation of small GTPase mediated signal transduction	2.25	3.09	2.90E-02	Induced
45	GO:0042127	regulation of cell proliferation	2.11	3.03	2.90E-02	Induced
46	GO:0016567	protein ubiquitination	1.68	2.88	2.90E-02	Induced
47	GO:0043065	positive regulation of apoptotic process	1.8	2.9	3.00E-02	Induced
48	GO:0070527	platelet aggregation	3.58	3.41	3.10E-02	Induced
49	GO:0045071	negative regulation of viral genome replication	3.58	3.41	3.10E-02	Induced
50	GO:0097192	extrinsic apoptotic signaling pathway in absence of ligand	4.07	3.48	3.20E-02	Induced

	GOBP	Pathway	FC	z score	Adjusted p-val	Direction
51	**GO:0046627**	negative regulation of insulin receptor signaling pathway	4.07	3.48	3.20E-02	Induced
52	**GO:0007204**	positive regulation of cytosolic calcium ion concentration	2.27	3	3.20E-02	Induced
53	**GO:0009612**	response to mechanical stimulus	3.06	3.18	3.50E-02	Induced
54	**GO:0070098**	chemokine-mediated signaling pathway	2.79	3.11	3.50E-02	Induced
55	**GO:0046777**	protein autophosphorylation	2.06	2.84	3.70E-02	Induced
56	**GO:0019731**	antibacterial humoral response	3.82	3.3	3.80E-02	Induced
57	**GO:1900026**	positive regulation of substrate adhesion-dependent cell spreading	3.82	3.3	3.80E-02	Induced
58	**GO:0032760**	positive regulation of tumor necrosis factor production	3.33	3.2	3.80E-02	Induced
59	**GO:0071363**	cellular response to growth factor stimulus	3.26	3.13	3.90E-02	Induced
60	**GO:0051092**	positive regulation of NF-kappaB transcription factor activity	2.17	2.83	3.90E-02	Induced
61	**GO:0050679**	positive regulation of epithelial cell proliferation	2.9	3.02	4.00E-02	Induced
62	**GO:0030036**	actin cytoskeleton organization	2.24	2.82	4.00E-02	Induced
63	**GO:0070301**	cellular response to hydrogen peroxide	2.85	2.97	4.30E-02	Induced
64	**GO:0071260**	cellular response to mechanical stimulus	2.57	2.84	4.70E-02	Induced
65	**GO:1902042**	negative regulation of extrinsic apoptotic signaling pathway via death domain receptors	3.49	3.05	4.80E-02	Induced
66	**GO:0043407**	negative regulation of MAP kinase activity	3.49	3.05	4.80E-02	Induced
67	**GO:0051384**	response to glucocorticoid	2.76	2.87	4.80E-02	Induced
68	**GO:0000187**	activation of MAPK activity	2.26	2.72	4.80E-02	Induced

Appendix D: GOBP pathways enriched in MM patients

GOBP pathways enriched across all transcripts induced in monocytes in MM patients at baseline (C1) versus healthy donors (P_{adj}<0.05). XGR interrogation of the GOBP database was used to perform pathway enrichment analysis

GOBP Pathway enrichment analysis of pathways suppressed in MM patients versus healthy donors

	GOBP	Pathway	FC	z score	Adjusted p-val	Direction
1	GO:0006614	SRP-dependent cotranslational protein targeting to membrane	13.4	25.3	2.30E-46	Suppressed
2	GO:0006413	translational initiation	10.2	22.7	5.80E-43	Suppressed
3	GO:0006412	translation	8.32	20.9	2.30E-40	Suppressed
4	GO:0006364	rRNA processing	7.58	20.1	2.10E-39	Suppressed
5	GO:0042776	mitochondrial ATP synthesis coupled proton transport	10.8	9.14	1.10E-06	Suppressed
6	GO:0050852	T cell receptor signaling pathway	3.54	6.64	3.40E-06	Suppressed
7	GO:0042274	ribosomal small subunit biogenesis	10.7	8.59	5.50E-06	Suppressed
8	GO:0030490	maturation of SSU-rRNA	9.95	7.68	4.90E-05	Suppressed
9	GO:0006123	mitochondrial electron transport, cytochrome c to oxygen	9.37	7.4	6.80E-05	Suppressed
10	GO:0002181	cytoplasmic translation	9.37	7.4	6.80E-05	Suppressed
11	GO:0070126	mitochondrial translational termination	3.92	5.86	8.20E-05	Suppressed
12	GO:0000027	ribosomal large subunit assembly	8.38	6.9	1.40E-04	Suppressed
13	GO:0000028	ribosomal small subunit assembly	9.1	6.73	3.20E-04	Suppressed
14	GO:0016032	viral process	2.21	4.53	6.60E-04	Suppressed
15	GO:0071353	cellular response to interleukin-4	6.37	5.76	7.60E-04	Suppressed
16	GO:0043928	exonucleolytic nuclear-transcribed mRNA catabolic process involved in deadenylation-dependent decay	5.52	5.57	7.60E-04	Suppressed
17	GO:0043966	histone H3 acetylation	4.88	5.39	7.60E-04	Suppressed
18	GO:0032212	positive regulation of telomere maintenance via telomerase	5.35	5.45	8.70E-04	Suppressed
19	GO:0042769	DNA damage response, detection of DNA damage	5.06	5.22	1.30E-03	Suppressed
20	GO:0000462	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	5.31	5.06	2.30E-03	Suppressed
21	GO:0019835	cytolysis	6.21	5.24	2.40E-03	Suppressed
22	GO:0042273	ribosomal large subunit biogenesis	6.21	5.24	2.40E-03	Suppressed
23	GO:0006283	transcription-coupled nucleotide-excision repair	3.43	4.46	2.50E-03	Suppressed
24	GO:0000398	mRNA splicing, via spliceosome	2.21	4.02	2.50E-03	Suppressed
25	GO:0031295	T cell costimulation	3.25	4.24	3.90E-03	Suppressed
26	GO:0006418	tRNA aminoacylation for protein translation	4.55	4.51	4.90E-03	Suppressed
27	GO:0006417	regulation of translation	3.66	4.27	4.90E-03	Suppressed
28	GO:0050870	positive regulation of T cell activation	5.99	4.66	7.60E-03	Suppressed
29	GO:0000289	nuclear-transcribed mRNA poly(A) tail shortening	4.88	4.4	7.60E-03	Suppressed
30	GO:0006260	DNA replication	2.44	3.66	8.10E-03	Suppressed
31	GO:0006783	heme biosynthetic process	5.69	4.5	8.60E-03	Suppressed
32	GO:0000387	spliceosomal snRNP assembly	4.71	4.28	8.60E-03	Suppressed
33	GO:1902600	hydrogen ion transmembrane transport	3.3	3.89	8.60E-03	Suppressed
34	GO:0099132	ATP hydrolysis coupled cation transmembrane transport	4.55	4.17	9.70E-03	Suppressed
35	GO:0006120	mitochondrial electron transport, NADH to ubiquinone	3.79	3.88	1.20E-02	Suppressed
36	GO:0000722	telomere maintenance via recombination	4.27	3.97	1.30E-02	Suppressed
37	GO:0050776	regulation of immune response	2.16	3.35	1.30E-02	Suppressed
38	GO:0000715	nucleotide-excision repair, DNA damage recognition	4.95	4.06	1.40E-02	Suppressed
39	GO:0030307	positive regulation of cell growth	2.81	3.5	1.40E-02	Suppressed
40	GO:0000122	negative regulation of transcription from RNA polymerase II promoter	1.52	3.04	1.40E-02	Suppressed
41	GO:1901796	regulation of signal transduction by p53 class mediator	2.37	3.29	1.60E-02	Suppressed
42	GO:0006968	cellular defense response	3.09	3.44	1.80E-02	Suppressed
43	GO:0045727	positive regulation of translation	3.09	3.44	1.80E-02	Suppressed
44	GO:0046487	glyoxylate metabolic process	4.38	3.69	2.10E-02	Suppressed
45	GO:0045648	positive regulation of erythrocyte differentiation	4.38	3.69	2.10E-02	Suppressed
46	GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	2.98	3.33	2.10E-02	Suppressed
47	GO:0050853	B cell receptor signaling pathway	2.77	3.27	2.10E-02	Suppressed
48	GO:0006091	generation of precursor metabolites and energy	3.19	3.32	2.20E-02	Suppressed
49	GO:0016925	protein sumoylation	2.73	3.22	2.20E-02	Suppressed
50	GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	2.56	3.16	2.20E-02	Suppressed

	GOBP	Pathway	FC	z score	Adjusted p-val	Direction
51	**GO:0008380**	RNA splicing	2.07	2.96	2.30E-02	Suppressed
52	**GO:0000060**	protein import into nucleus, translocation	4.06	3.48	2.40E-02	Suppressed
53	**GO:0000245**	spliceosomal complex assembly	4.06	3.48	2.40E-02	Suppressed
54	**GO:0009615**	response to virus	2.32	2.94	2.80E-02	Suppressed
55	**GO:0006099**	tricarboxylic acid cycle	3.79	3.28	3.10E-02	Suppressed
56	**GO:0090503**	RNA phosphodiester bond hydrolysis, exonucleolytic	3.79	3.28	3.10E-02	Suppressed
57	**GO:0034080**	CENP-A containing nucleosome assembly	3.25	3.13	3.10E-02	Suppressed
58	**GO:0045599**	negative regulation of fat cell differentiation	3.25	3.13	3.10E-02	Suppressed
59	**GO:0031047**	gene silencing by RNA	2.25	2.84	3.10E-02	Suppressed
60	**GO:0045785**	positive regulation of cell adhesion	3.17	3.06	3.40E-02	Suppressed
61	**GO:0032981**	mitochondrial respiratory chain complex I assembly	2.84	2.96	3.50E-02	Suppressed
62	**GO:0007498**	mesoderm development	3.56	3.1	3.60E-02	Suppressed
63	**GO:0097421**	liver regeneration	3.56	3.1	3.60E-02	Suppressed
64	**GO:0006959**	humoral immune response	2.79	2.91	3.60E-02	Suppressed
65	**GO:0050821**	protein stabilization	2.03	2.67	3.60E-02	Suppressed
66	**GO:0042102**	positive regulation of T cell proliferation	2.75	2.86	3.80E-02	Suppressed
67	**GO:0007062**	sister chromatid cohesion	2.25	2.71	3.80E-02	Suppressed
68	**GO:0006409**	tRNA export from nucleus	3.45	3.02	3.90E-02	Suppressed
69	**GO:0006270**	DNA replication initiation	3.45	3.02	3.90E-02	Suppressed

Appendix E: GOBP pathways depleted in MM patients

GOBP pathways enriched across all transcripts downregulated in monocytes in MM patients at baseline (C1) versus healthy donors (P_{adj}<0.05). XGR interrogation of the GOBP database was used to perform pathway enrichment analysis.

Monocyte profiles associated with baseline monocyte count – pathway enrichment analysis of positively associated transcripts

	GOBP ID	Pathway	Fold change	z score	Adjusted p-val	Direction
1	GO:0043312	neutrophil degranulation	4	12.6	4.60E-20	Positive
2	GO:0007155	cell adhesion	2.24	5.06	6.70E-04	Positive
3	GO:0034113	heterotypic cell-cell adhesion	7.96	6.15	2.50E-03	Positive
4	GO:0002227	innate immune response in mucosa	7.96	6.15	2.50E-03	Positive
5	GO:0019731	antibacterial humoral response	6.38	5.74	2.50E-03	Positive
6	GO:0007160	cell-matrix adhesion	3.81	5.08	2.50E-03	Positive
7	GO:0033627	cell adhesion mediated by integrin	9.12	6.12	2.90E-03	Positive
8	GO:0048010	vascular endothelial growth factor receptor signaling pathway	4.11	4.95	2.90E-03	Positive
9	GO:0007229	integrin-mediated signaling pathway	3.82	4.88	2.90E-03	Positive
10	GO:0042542	response to hydrogen peroxide	4.86	5.05	3.50E-03	Positive
11	GO:0050830	defense response to Gram-positive bacterium	4.17	4.75	4.00E-03	Positive
12	GO:0030335	positive regulation of cell migration	2.73	4.28	4.00E-03	Positive
13	GO:0002576	platelet degranulation	3.01	4.27	5.20E-03	Positive
14	GO:0070527	platelet aggregation	4.98	4.81	5.30E-03	Positive
15	GO:0019730	antimicrobial humoral response	4.32	4.61	5.30E-03	Positive
16	GO:0007204	positive regulation of cytosolic calcium ion concentration	2.94	4.17	5.30E-03	Positive
17	GO:0045087	innate immune response	1.9	3.58	9.10E-03	Positive
18	GO:0007030	Golgi organization	3.46	4.04	1.10E-02	Positive
19	GO:0006928	movement of cell or subcellular component	3.28	3.86	1.40E-02	Positive
20	GO:0006468	protein phosphorylation	1.83	3.34	1.40E-02	Positive
21	GO:0006954	inflammatory response	1.92	3.34	1.50E-02	Positive
22	GO:0035987	endodermal cell differentiation	5.4	4.31	1.60E-02	Positive
23	GO:0006888	ER to Golgi vesicle-mediated transport	2.42	3.36	2.20E-02	Positive
24	GO:0015701	bicarbonate transport	3.98	3.73	2.60E-02	Positive
25	GO:0016575	histone deacetylation	3.98	3.73	2.60E-02	Positive
26	GO:0006909	phagocytosis	3.89	3.66	2.70E-02	Positive
27	GO:0022617	extracellular matrix disassembly	3.07	3.41	2.70E-02	Positive
28	GO:0050900	leukocyte migration	2.16	3.12	2.70E-02	Positive
29	GO:0051591	response to cAMP	3.81	3.59	2.80E-02	Positive
30	GO:0046718	viral entry into host cell	2.96	3.28	3.20E-02	Positive
31	GO:0006936	muscle contraction	2.74	3.21	3.20E-02	Positive
32	GO:0009611	response to wounding	3.19	3.31	3.30E-02	Positive
33	GO:0008285	negative regulation of cell proliferation	1.74	2.77	3.90E-02	Positive
34	GO:0006935	chemotaxis	2.45	3	4.00E-02	Positive
35	GO:0042493	response to drug	1.84	2.78	4.00E-02	Positive
36	GO:0051056	regulation of small GTPase mediated signal transduction	2.28	2.87	4.50E-02	Positive
37	GO:0007596	blood coagulation	2.08	2.77	4.70E-02	Positive
38	GO:0032496	response to lipopolysaccharide	2.15	2.78	4.80E-02	Positive
39	GO:0010628	positive regulation of gene expression	1.86	2.66	4.80E-02	Positive

Appendix F: GOBP pathways associated with baseline peripheral monocyte count.
Pathway enrichment analysis using XGR interrogation of the GOBP pathway across all monocyte-expressed transcripts positively associated with peripheral monocyte count (P_{adj}<0.05).

	GOBP ID	Pathway	FC	z score	Adjusted p-val	Direction
1	GO:0006364	rRNA processing	6.84	15.3	5.70E-23	Negative
2	GO:0006260	DNA replication	4.8	8.31	7.00E-08	Negative
3	GO:0006270	DNA replication initiation	9.26	8.74	2.90E-06	Negative
4	GO:0006418	tRNA aminoacylation for protein translation	7.86	7.47	4.80E-05	Negative
5	GO:1904874	positive regulation of telomerase RNA localization to Cajal body	12.2	8	1.10E-04	Negative
6	GO:0000462	maturation of SSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA)	8.15	7.21	1.10E-04	Negative
7	GO:0000722	telomere maintenance via recombination	7.64	6.92	1.40E-04	Negative
8	GO:0010501	RNA secondary structure unwinding	6.11	6.32	2.00E-04	Negative
9	GO:0019083	viral transcription	3.96	5.68	2.00E-04	Negative
10	GO:0009168	purine ribonucleoside monophosphate biosynthetic process	11.8	7.14	5.00E-04	Negative
11	GO:1900034	regulation of cellular response to heat	4.42	5.5	5.00E-04	Negative
12	GO:0034080	CENP-A containing nucleosome assembly	5.82	5.76	6.30E-04	Negative
13	GO:0000082	G1/S transition of mitotic cell cycle	3.65	5.1	6.30E-04	Negative
14	GO:0006409	tRNA export from nucleus	6.49	5.8	8.10E-04	Negative
15	GO:0071897	DNA biosynthetic process	7.34	5.83	1.20E-03	Negative
16	GO:0006541	glutamine metabolic process	8.49	5.85	1.80E-03	Negative
17	GO:0075733	intracellular transport of virus	4.8	4.99	1.80E-03	Negative
18	GO:0006412	translation	2.8	4.39	1.80E-03	Negative
19	GO:0007077	mitotic nuclear envelope disassembly	4.86	4.72	3.90E-03	Negative
20	GO:0042273	ribosomal large subunit biogenesis	6.95	5.13	4.30E-03	Negative
21	GO:0016925	protein sumoylation	3.67	4.26	4.90E-03	Negative
22	GO:0016032	viral process	2.15	3.7	5.30E-03	Negative
23	GO:0006297	nucleotide-excision repair, DNA gap filling	6.37	4.84	5.60E-03	Negative
24	GO:0042769	DNA damage response, detection of DNA damage	5.1	4.52	5.80E-03	Negative
25	GO:0006457	protein folding	2.56	3.73	6.70E-03	Negative
26	GO:0008033	tRNA processing	5.88	4.58	7.10E-03	Negative
27	GO:0006413	translational initiation	2.78	3.77	7.10E-03	Negative
28	GO:0090502	RNA phosphodiester bond hydrolysis, endonucleolytic	3.96	4.01	9.00E-03	Negative
29	GO:0006400	tRNA modification	5.27	4.23	1.10E-02	Negative
30	GO:0006283	transcription-coupled nucleotide-excision repair	3.35	3.7	1.20E-02	Negative
31	GO:0007005	mitochondrion organization	3.26	3.61	1.40E-02	Negative
32	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	2.99	3.52	1.40E-02	Negative
33	GO:0006302	double-strand break repair	3.51	3.61	1.60E-02	Negative
34	GO:0000723	telomere maintenance	4.37	3.67	2.10E-02	Negative
35	GO:0045859	regulation of protein kinase activity	4.37	3.67	2.10E-02	Negative
36	GO:0061077	chaperone-mediated protein folding	3.67	3.48	2.20E-02	Negative
37	GO:0006461	protein complex assembly	2.57	3.16	2.20E-02	Negative
38	GO:0006281	DNA repair	2.13	2.96	2.40E-02	Negative
39	GO:0051259	protein oligomerization	3.22	3.08	3.50E-02	Negative
40	GO:0007049	cell cycle	1.9	2.68	3.50E-02	Negative
41	GO:0000398	mRNA splicing, via spliceosome	1.93	2.67	3.70E-02	Negative
42	GO:0031295	T cell costimulation	2.78	2.88	4.10E-02	Negative
43	GO:0032508	DNA duplex unwinding	3.47	3.02	4.20E-02	Negative
44	GO:0008283	cell proliferation	1.71	2.5	4.20E-02	Negative
45	GO:0031100	animal organ regeneration	3.4	2.96	4.50E-02	Negative

Appendix G: GOBP pathways negatively associated with baseline monocyte count

Pathway enrichment analysis using XGR interrogation of the GOBP pathway across all monocyte-expressed transcripts negatively associated with peripheral monocyte count (P_{adj}<0.05).

Pathway enrichment analysis of GOBP pathways induced with ICB

	GOBP	Pathway	FC	z score	Adjp	Direction
1	GO:0060333	interferon-gamma-mediated signaling pathway	7.9	16.7	2.30E-27	Induced
2	GO:0002479	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	6.8	14.7	2.00E-22	Induced
3	GO:0060337	type I interferon signaling pathway	6.9	13.8	8.30E-20	Induced
4	GO:0051607	defense response to virus	3.9	10.6	7.10E-15	Induced
5	GO:0050852	T cell receptor signaling pathway	4	10.7	7.20E-15	Induced
6	GO:0038061	NIK/NF-kappaB signaling	5.1	9.9	2.00E-11	Induced
7	GO:0033209	tumor necrosis factor-mediated signaling pathway	3.8	9.14	3.50E-11	Induced
8	GO:0000209	protein polyubiquitination	2.9	8.22	3.90E-10	Induced
9	GO:0043488	regulation of mRNA stability	3.9	8.6	8.00E-10	Induced
10	GO:0045071	negative regulation of viral genome replication	6	9.28	1.50E-09	Induced
11	GO:0002474	antigen processing and presentation of peptide antigen via MHC class I	7.2	9.58	1.70E-09	Induced
12	GO:0009615	response to virus	3.7	8.2	2.70E-09	Induced
13	GO:0031145	anaphase-promoting complex-dependent catabolic process	4.2	8.24	5.60E-09	Induced
14	GO:0002223	stimulatory C-type lectin receptor signaling pathway	3.2	7.24	8.70E-08	Induced
15	GO:0016579	protein deubiquitination	2.4	6.47	3.80E-07	Induced
16	GO:0032480	negative regulation of type I interferon production	5.5	7.48	1.20E-06	Induced
17	GO:0006123	mitochondrial electron transport, cytochrome c to oxygen	7.6	7.9	2.00E-06	Induced
18	GO:0090090	negative regulation of canonical Wnt signaling pathway	2.7	6.1	4.10E-06	Induced
19	GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	2.4	5.91	5.80E-06	Induced
20	GO:0090263	positive regulation of canonical Wnt signaling pathway	2.8	5.74	2.20E-05	Induced
21	GO:0051092	positive regulation of NF-kappaB transcription factor activity	2.7	5.69	2.20E-05	Induced
22	GO:0045087	innate immune response	1.9	5.29	2.20E-05	Induced
23	GO:0043123	positive regulation of I-kappaB kinase/NF-kappaB signaling	2.5	5.59	2.40E-05	Induced
24	GO:0006120	mitochondrial electron transport, NADH to ubiquinone	4.3	6.23	2.80E-05	Induced
25	GO:0055085	transmembrane transport	2.2	5.38	2.80E-05	Induced
26	GO:0070125	mitochondrial translational elongation	3.1	5.75	3.20E-05	Induced
27	GO:0032981	mitochondrial respiratory chain complex I assembly	3.7	5.86	4.80E-05	Induced
28	GO:0006919	activation of cysteine-type endopeptidase activity involved in apoptotic process	3.2	5.59	6.50E-05	Induced
29	GO:0034341	response to interferon-gamma	5.4	6.24	7.60E-05	Induced
30	GO:0016032	viral process	2	4.82	1.50E-04	Induced
31	GO:0032755	positive regulation of interleukin-6 production	3.9	5.54	1.80E-04	Induced
32	GO:0038095	Fc-epsilon receptor signaling pathway	2.2	4.8	2.60E-04	Induced
33	GO:0006955	immune response	1.8	4.53	3.00E-04	Induced
34	GO:0032727	positive regulation of interferon-alpha production	7.8	6.2	4.30E-04	Induced
35	GO:0000165	MAPK cascade	1.9	4.43	5.80E-04	Induced
36	GO:0006099	tricarboxylic acid cycle	4.3	5.26	6.40E-04	Induced
37	GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	2.7	4.66	7.70E-04	Induced
38	GO:0006986	response to unfolded protein	3.6	4.97	8.00E-04	Induced
39	GO:0070555	response to interleukin-1	4.1	4.99	1.10E-03	Induced
40	GO:0010803	regulation of tumor necrosis factor-mediated signaling pathway	4.1	4.99	1.10E-03	Induced
41	GO:0001916	positive regulation of T cell mediated cytotoxicity	6.5	5.49	1.40E-03	Induced
42	GO:0007249	I-kappaB kinase/NF-kappaB signaling	3	4.46	1.90E-03	Induced
43	GO:0006954	inflammatory response	1.7	3.93	2.10E-03	Induced
44	GO:0051085	chaperone mediated protein folding requiring cofactor	6	5.19	2.20E-03	Induced
45	GO:0097296	activation of cysteine-type endopeptidase activity involved in apoptotic signaling pathway	6	5.19	2.20E-03	Induced
46	GO:1902600	hydrogen ion transmembrane transport	2.9	4.39	2.20E-03	Induced
47	GO:0031295	T cell costimulation	2.7	4.3	2.20E-03	Induced
48	GO:0008637	apoptotic mitochondrial changes	5	4.96	2.30E-03	Induced
49	GO:0050860	negative regulation of T cell receptor signaling pathway	5	4.96	2.30E-03	Induced
50	GO:2001238	positive regulation of extrinsic apoptotic signaling pathway	4.3	4.7	2.80E-03	Induced
51	GO:0002230	positive regulation of defense response to virus by host	4.8	4.76	3.20E-03	Induced
52	GO:0035455	response to interferon-alpha	6.5	5.01	4.30E-03	Induced
53	GO:0006801	superoxide metabolic process	4.5	4.57	4.30E-03	Induced
54	GO:0001836	release of cytochrome c from mitochondria	4.5	4.57	4.30E-03	Induced
55	GO:0032729	positive regulation of interferon-gamma production	3.2	4.21	4.30E-03	Induced
56	GO:0043312	neutrophil degranulation	1.6	3.6	4.30E-03	Induced
57	GO:0032728	positive regulation of interferon-beta production	4	4.41	4.50E-03	Induced
58	GO:0042026	protein refolding	5.2	4.69	4.60E-03	Induced
59	GO:0032469	endoplasmic reticulum calcium ion homeostasis	5.2	4.69	4.60E-03	Induced
60	GO:0050776	regulation of immune response	2	3.71	4.70E-03	Induced

	GOBP	Pathway	FC	z score	Adjp	Direction
61	GO:0006139	nucleobase-containing compound metabolic process	3.5	4.21	5.10E-03	Induced
62	GO:0031398	positive regulation of protein ubiquitination	2.8	3.98	5.10E-03	Induced
63	GO:0097190	apoptotic signaling pathway	2.6	3.92	5.20E-03	Induced
64	GO:0097191	extrinsic apoptotic signaling pathway	3.2	4.1	5.40E-03	Induced
65	GO:0019068	virion assembly	5.9	4.69	6.00E-03	Induced
66	GO:0010529	negative regulation of transposition	5.9	4.69	6.00E-03	Induced
67	GO:2000406	positive regulation of T cell migration	5.9	4.69	6.00E-03	Induced
68	GO:0032479	regulation of type I interferon production	4.9	4.46	6.00E-03	Induced
69	GO:0019835	cytolysis	4.1	4.24	6.70E-03	Induced
70	GO:0050718	positive regulation of interleukin-1 beta secretion	4.1	4.24	6.70E-03	Induced
71	GO:0006457	protein folding	1.9	3.53	7.10E-03	Induced
72	GO:0055119	relaxation of cardiac muscle	5.4	4.41	8.70E-03	Induced
73	GO:0032693	negative regulation of interleukin-10 production	5.4	4.41	8.70E-03	Induced
74	GO:0032456	endocytic recycling	3.9	4.08	8.70E-03	Induced
75	GO:0042102	positive regulation of T cell proliferation	2.7	3.71	8.70E-03	Induced
76	GO:0055114	oxidation-reduction process	1.5	3.24	8.70E-03	Induced
77	GO:0007259	JAK-STAT cascade	3.5	3.89	9.80E-03	Induced
78	GO:0032760	positive regulation of tumor necrosis factor production	2.9	3.73	1.00E-02	Induced
79	GO:0031293	membrane protein intracellular domain proteolysis	4.3	4.07	1.10E-02	Induced
80	GO:0045088	regulation of innate immune response	5	4.15	1.20E-02	Induced
81	GO:0010248	establishment or maintenance of transmembrane electrochemical gradient	5	4.15	1.20E-02	Induced
82	GO:0006883	cellular sodium ion homeostasis	5	4.15	1.20E-02	Induced
83	GO:0031648	protein destabilization	3.1	3.69	1.20E-02	Induced
84	GO:0043066	negative regulation of apoptotic process	1.5	3.15	1.20E-02	Induced
85	GO:0042776	mitochondrial ATP synthesis coupled proton transport	4.1	3.9	1.40E-02	Induced
86	GO:0050870	positive regulation of T cell activation	4.1	3.9	1.40E-02	Induced
87	GO:0031532	actin cytoskeleton reorganization	2.8	3.49	1.50E-02	Induced
88	GO:0019079	viral genome replication	4.6	3.93	1.60E-02	Induced
89	GO:2000045	regulation of G1/S transition of mitotic cell cycle	4.6	3.93	1.60E-02	Induced
90	GO:0071480	cellular response to gamma radiation	3.5	3.67	1.60E-02	Induced
91	GO:0050830	defense response to Gram-positive bacterium	2.5	3.37	1.60E-02	Induced
92	GO:0006401	RNA catabolic process	3.9	3.74	1.70E-02	Induced
93	GO:0042346	positive regulation of NF-kappaB import into nucleus	3.9	3.74	1.70E-02	Induced
94	GO:0008217	regulation of blood pressure	2.4	3.31	1.70E-02	Induced
95	GO:0050766	positive regulation of phagocytosis	3.4	3.55	1.90E-02	Induced
96	GO:0042127	regulation of cell proliferation	1.8	3.02	2.00E-02	Induced
97	GO:1903543	positive regulation of exosomal secretion	4.3	3.72	2.10E-02	Induced
98	GO:0050901	leukocyte tethering or rolling	4.3	3.72	2.10E-02	Induced
99	GO:0060828	regulation of canonical Wnt signaling pathway	3.7	3.58	2.10E-02	Induced
100	GO:0008283	cell proliferation	1.5	2.89	2.10E-02	Induced
101	GO:0097193	intrinsic apoptotic signaling pathway	3.2	3.43	2.20E-02	Induced
102	GO:1901998	toxin transport	3	3.36	2.20E-02	Induced
103	GO:0007005	mitochondrion organization	2.3	3.13	2.20E-02	Induced
104	GO:0006635	fatty acid beta-oxidation	2.7	3.25	2.30E-02	Induced
105	GO:0006464	cellular protein modification process	2.1	3.05	2.30E-02	Induced
106	GO:0042981	regulation of apoptotic process	1.7	2.91	2.30E-02	Induced
107	GO:0033617	mitochondrial respiratory chain complex IV assembly	4.1	3.53	2.70E-02	Induced
108	GO:0048535	lymph node development	4.1	3.53	2.70E-02	Induced
109	GO:0050727	regulation of inflammatory response	2.3	3.02	2.90E-02	Induced
110	GO:0006977	DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	2.3	3.02	2.90E-02	Induced
111	GO:0071222	cellular response to lipopolysaccharide	1.9	2.89	2.90E-02	Induced
112	GO:0030866	cortical actin cytoskeleton organization	3.4	3.3	3.00E-02	Induced
113	GO:0031100	animal organ regeneration	2.6	3.09	3.00E-02	Induced
114	GO:0008015	blood circulation	2.6	3.09	3.00E-02	Induced
115	GO:0046718	viral entry into host cell	2.1	2.92	3.10E-02	Induced
116	GO:0043171	peptide catabolic process	3.8	3.35	3.30E-02	Induced
117	GO:1902041	regulation of extrinsic apoptotic signaling pathway via death domain receptors	3.8	3.35	3.30E-02	Induced
118	GO:0030890	positive regulation of B cell proliferation	2.7	3.08	3.30E-02	Induced
119	GO:0046034	ATP metabolic process	2.9	3.1	3.50E-02	Induced
120	GO:0045454	cell redox homeostasis	2.2	2.85	3.60E-02	Induced

	GOBP	Pathway	FC	z score	Adjp	Direction
121	**GO:0006695**	cholesterol biosynthetic process	2.7	3	3.70E-02	Induced
122	**GO:0006458**	'de novo' protein folding	3.6	3.19	4.00E-02	Induced
123	**GO:0031668**	cellular response to extracellular stimulus	3.6	3.19	4.00E-02	Induced
124	**GO:1901214**	regulation of neuron death	3.6	3.19	4.00E-02	Induced
125	**GO:0008625**	extrinsic apoptotic signaling pathway via death domain receptors	2.8	3	4.00E-02	Induced
126	**GO:0032689**	negative regulation of interferon-gamma production	3.1	3.05	4.10E-02	Induced
127	**GO:0032735**	positive regulation of interleukin-12 production	3.1	3.05	4.10E-02	Induced
128	**GO:0033198**	response to ATP	3.4	3.04	4.70E-02	Induced
129	**GO:0046655**	folic acid metabolic process	3.4	3.04	4.70E-02	Induced
130	**GO:0002028**	regulation of sodium ion transport	3.4	3.04	4.70E-02	Induced
131	**GO:0042098**	T cell proliferation	3.4	3.04	4.70E-02	Induced
132	**GO:2000778**	positive regulation of interleukin-6 secretion	3.4	3.04	4.70E-02	Induced
133	**GO:1902895**	positive regulation of pri-miRNA transcription from RNA polymerase II promoter	3.4	3.04	4.70E-02	Induced
134	**GO:0006754**	ATP biosynthetic process	3	2.94	4.70E-02	Induced
135	**GO:0070372**	regulation of ERK1 and ERK2 cascade	3	2.94	4.70E-02	Induced
136	**GO:0006952**	defense response	2.2	2.72	4.70E-02	Induced
137	**GO:0006091**	generation of precursor metabolites and energy	2.3	2.73	4.90E-02	Induced
138	**GO:0032481**	positive regulation of type I interferon production	2.3	2.73	4.90E-02	Induced
139	**GO:0048146**	positive regulation of fibroblast proliferation	2.3	2.73	4.90E-02	Induced

Appendix H: GOBP pathways enriched in induced transcripts following ICB

Pathway enrichment analysis using XGR interrogation of the GOBP database was performed across all transcripts found to be induced with ICB (DESEq2 pairwise analysis, P_{adj}<0.05).

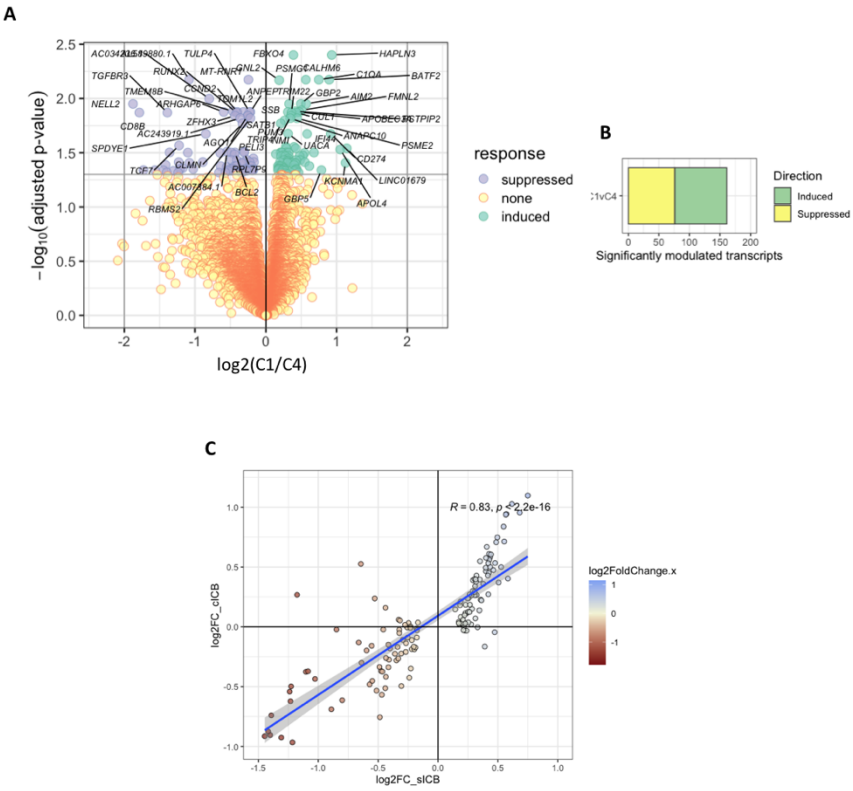
Pathway enrichment analysis of GOBP pathways suppressed with ICB

	GOBP	Pathway	FC	z score	Adjp	Direction
1	**GO:0006468**	protein phosphorylation	2.13	6.33	6.30E-06	Suppressed
2	**GO:0051056**	regulation of small GTPase mediated signal transduction	3.1	6.55	1.30E-05	Suppressed
3	**GO:0035735**	intraciliary transport involved in cilium assembly	4.68	6.1	6.00E-04	Suppressed
4	**GO:0035556**	intracellular signal transduction	1.95	4.93	6.50E-04	Suppressed
5	**GO:0060271**	cilium assembly	2.43	4.86	1.70E-03	Suppressed
6	**GO:0035023**	regulation of Rho protein signal transduction	3.16	5.04	2.10E-03	Suppressed
7	**GO:0043065**	positive regulation of apoptotic process	1.99	4.51	2.10E-03	Suppressed
8	**GO:0018105**	peptidyl-serine phosphorylation	2.46	4.63	2.30E-03	Suppressed
9	**GO:0043547**	positive regulation of GTPase activity	1.66	4.25	2.30E-03	Suppressed
10	**GO:0007257**	activation of JUN kinase activity	4.22	5.13	3.00E-03	Suppressed
11	**GO:0042552**	myelination	3.81	4.95	3.30E-03	Suppressed
12	**GO:0006351**	transcription, DNA-templated	1.33	3.89	3.50E-03	Suppressed
13	**GO:0030336**	negative regulation of cell migration	2.66	4.21	8.00E-03	Suppressed
14	**GO:0001525**	angiogenesis	2.03	3.98	8.00E-03	Suppressed
15	**GO:0010507**	negative regulation of autophagy	3.9	4.56	8.80E-03	Suppressed
16	**GO:0007611**	learning or memory	3.51	4.16	1.90E-02	Suppressed
17	**GO:0042073**	intraciliary transport	6	4.72	2.00E-02	Suppressed
18	**GO:0006470**	protein dephosphorylation	2.25	3.66	2.00E-02	Suppressed
19	**GO:0046777**	protein autophosphorylation	2.07	3.62	2.00E-02	Suppressed
20	**GO:0018107**	peptidyl-threonine phosphorylation	2.75	3.79	2.10E-02	Suppressed
21	**GO:0048013**	ephrin receptor signaling pathway	2.54	3.75	2.10E-02	Suppressed
22	**GO:0021915**	neural tube development	3.77	3.9	3.10E-02	Suppressed
23	**GO:0001933**	negative regulation of protein phosphorylation	2.77	3.65	3.10E-02	Suppressed
24	**GO:0045893**	positive regulation of transcription, DNA-templated	1.52	3.15	3.10E-02	Suppressed
25	**GO:0016337**	single organismal cell-cell adhesion	2.37	3.46	3.20E-02	Suppressed
26	**GO:0001822**	kidney development	2.39	3.35	4.20E-02	Suppressed
27	**GO:0030335**	positive regulation of cell migration	1.92	3.15	4.40E-02	Suppressed
28	**GO:0010039**	response to iron ion	4.59	3.87	4.60E-02	Suppressed
29	**GO:0048663**	neuron fate commitment	4.59	3.87	4.60E-02	Suppressed

Appendix I: GOBP pathways enriched across transcripts suppressed ICB

Table to show regulars of XGR interrogation of the GOBP database was carried out across all transcripts significantly downregulated with ICB (P_{adj}<0.05).

Transcriptional modulation at cycle four (C4) compared to baseline in peripheral monocytes with cICB



Appendix J: Transcriptional modulation in monocytes at cycle 4 (C4) with cICB

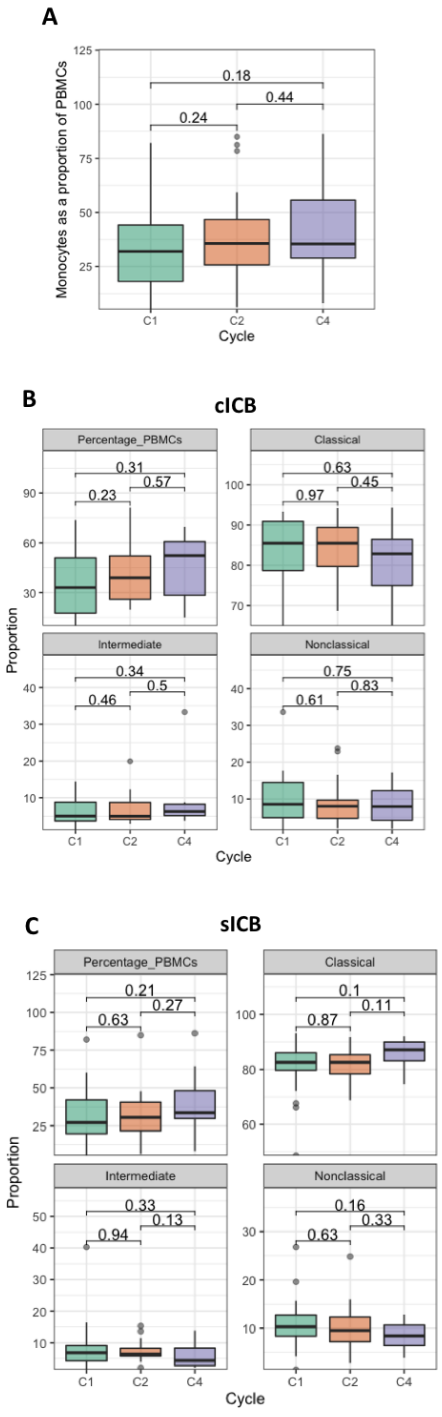
A) Volcano plot to show significantly differentially expressed genes at day 63 (C4) versus baseline (C1) in patients receiving cICB (pairwise analysis n= 15 patients, P_{adj}<0.05). B) Bar plot showing significantly differentially induced (n=85) and downregulated (n-76) transcripts with pairwise C1vC4 analysis, P_{adj}<0.05. C) Pearson's correlation between the log2 fold change (log2FC) significantly differentially modulated transcripts (n=161) in patients receiving sICB (15 patients) with the log2 fold change (log2FC) in those receiving cICB (7 patients).

Genes used for MM diagnostic classifier

Diagnostic classifier genes	
METTL7B	*MOGS*
SPRY1	*CHI3L1*
NR4A2	*KLHL2*
ADAMTS2	*TPST1*
IL1R2	*NLRP6*
CXCR1	*MMP9*
FOSB	*HCAR3*
FAM20A	*WDR61*
FCGR3B	*CEP104*
SOCS3	*MAOA*
ARID3A	*BICD2*
C16orf54	*NACC2*
MGAM	*HHEX*
KLF6	*CXCR2*
ANXA3	*MIR27A*
CDKN2B	*MIR23A*
FRAT1	*PTAFR*
MME	*MSLN*
HCAR2	*SENP1*
TIPARP	*CNTNAP3*

Appendix K: Transcripts used for MM diagnostic classifier

Transcripts derived from the top 40 most significantly differentially expressed gene in patients with MM versus healthy donors. This gene set was used to create a diagnostic classifier with the MASS package.



Appendix L: Monocyte proportion with flow cytometry

A) Monocyte population size as a proportion of all PBMCs on flow cytometry by cycle (46 C1 samples, 46 C2 samples, 18 C4 samples). B) Proportion of classical, intermediate and non-classical monocytes (of total monocyte cluster size) and size of monocyte population as a proportion of all PBMCs on flow cytometry by cycle and treatment: B) with cICB (25 C1 samples, 25 C2 samples, 6 C4 samples patients) and C) with sICB (21 C1 samples, 21 C2 samples, 12 C4 samples).

Single cell RNA-seq Cohort details

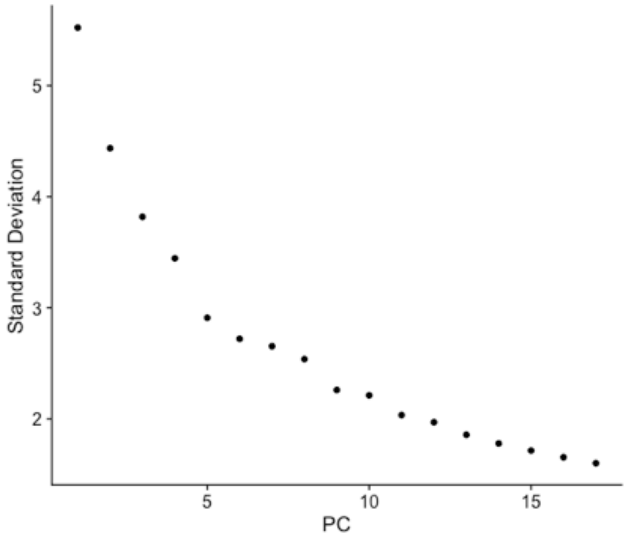
Patient ID	Age	Sex	Treatment	Treatment cycle	No. $CD14^+$ cells
ID1	82	M	sICB	C1	1752
				C2	1568
ID2	72	F	sICB	C1	475
				C2	1073
ID3	57	M	cICB	C1	1182
				C2	1498
ID4	45	F	cICB	C1	653
				C2	749
ID5	43	F	sICB	C1	1275
				C2	727
ID6	56	M	cICB	C1	695
				C2	1629
ID7	68	M	cICB	C1	1611
				C2	2232
ID8	69	M	sICB	C1	1678
				C2	1513
HD_1	51	F	NA	Control	4147*
HD_2	56	F	NA	Control	
HD_3	64	F	NA	Control	

Appendix M: scRNA-seq cohort data

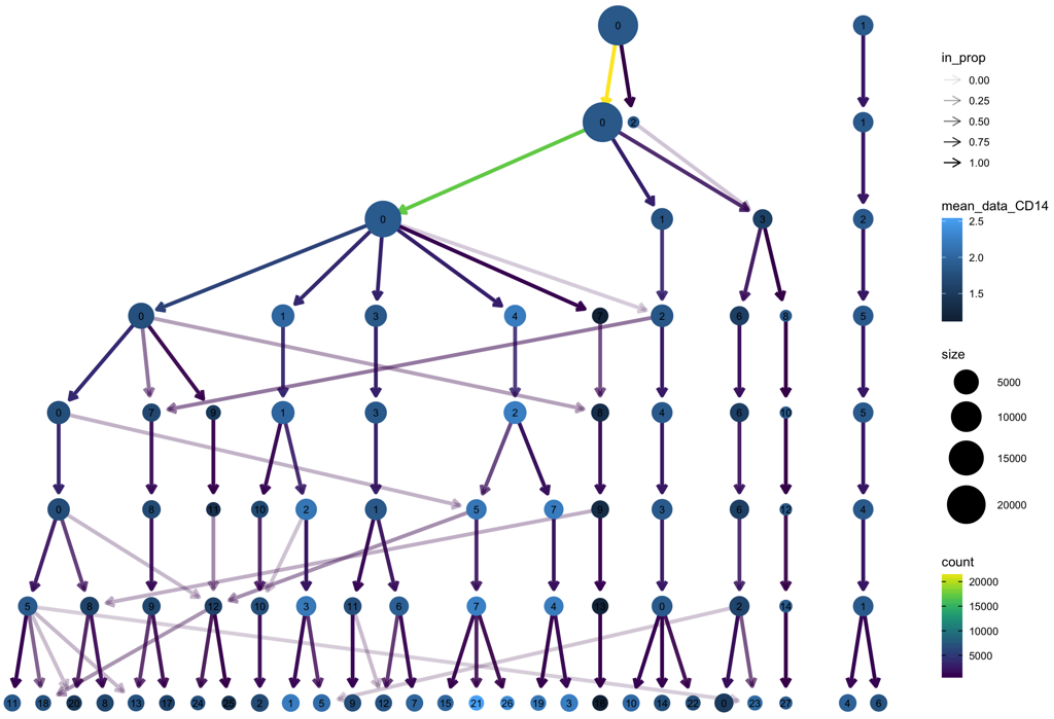
Table to show cohort details for scRNA-seq data. Note C1 denotes sample taken immediately pre-treatment (day1) and C2 taken day 21 post first treatment. Treatment – cICB denotes combination therapy (ipilimumab/nivolumab), whilst sICB denotes single agent therapy (pembrolizumab).

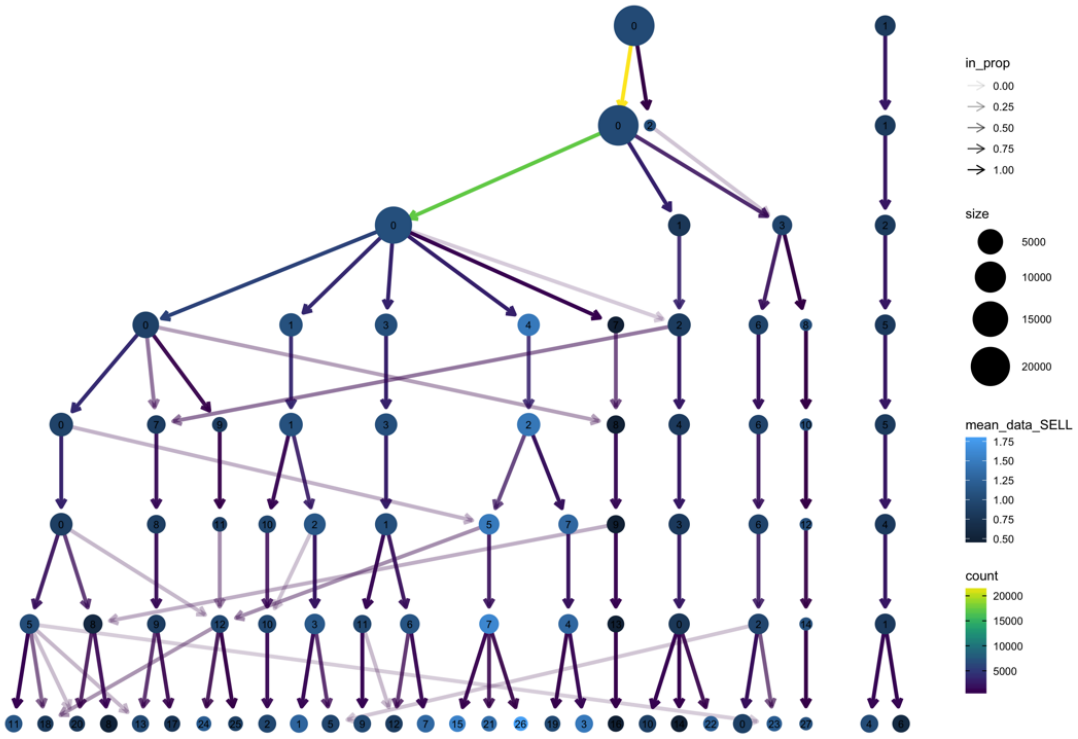
**Control samples were pooled with library preparation, sequencing and analysis occurring on a merged Control sample from the three individuals.*

Identifying the dimensionality across the dataset



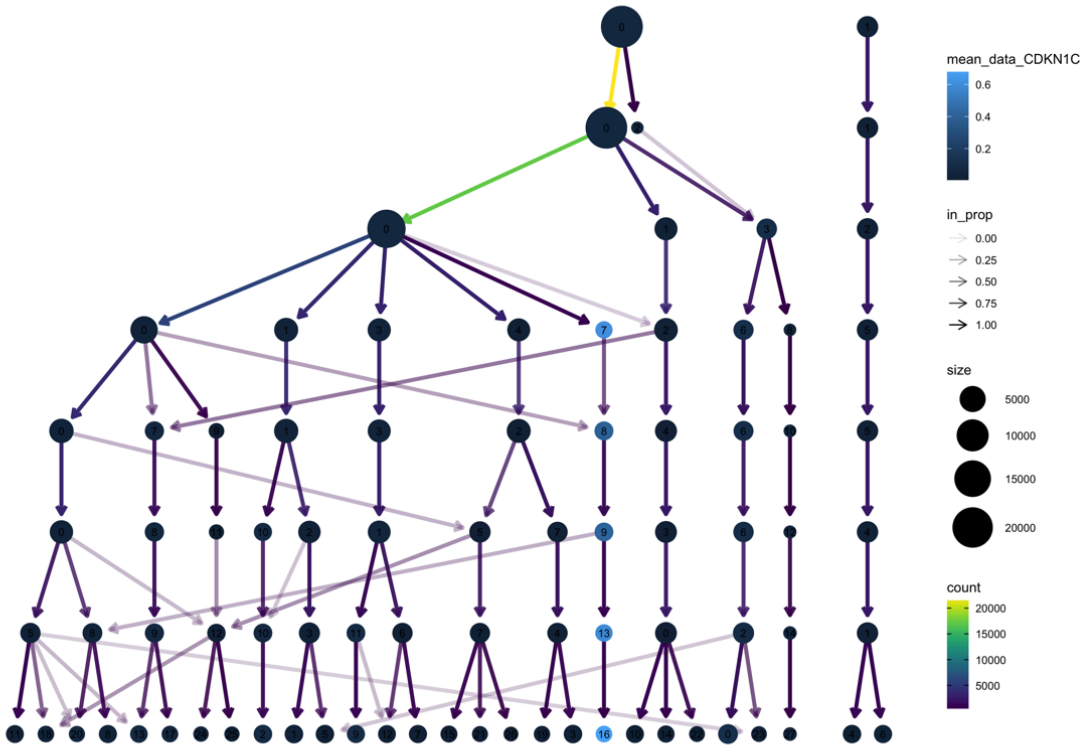
Appendix N: Elbowplot to show principal components
Elbowplot to show standard deviations of successive principal components (PCs) across all monocyte scRNA-seq data (n=24,457 cells).

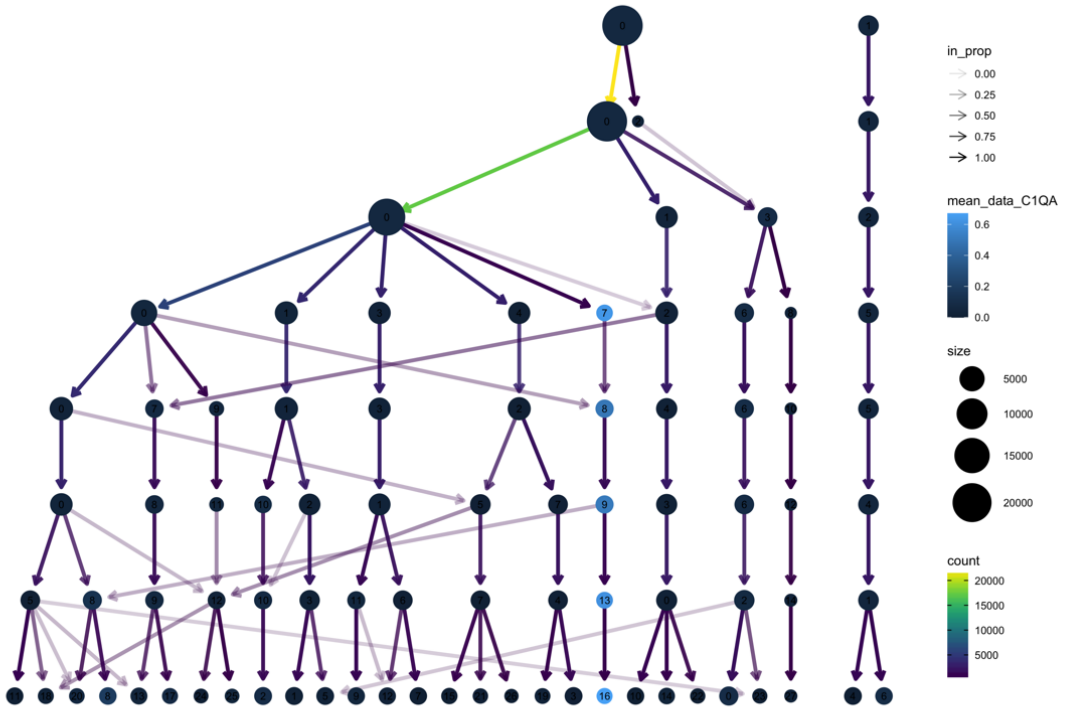




Appendix O: *Clustree* plots to show *CD14* and *SELL* expression

Clustree plot to show mean expression of classical monocyte marker CD14 (top figure) and SELL (bottom figure) across monocyte subset at varying clustering resolutions.





Appendix P: *Clustree* plot to show *CDKN1C* and *C1QA* expression

Clustree plots to show mean expression of non-classical monocyte markers CDKN1C and C1QA at varying clustering resolutions which are particularly highly expression within subset 7.

Classical monocytes	
S100A8	*RP11-522I20.3*
S100A12	*ITGB3*
S100A9	*CACNA1D*
VCAN	*CLU*
LYZ	*NSUN7*
CD14	*FAM131B*
GPX1	*TRAPPC3L*
GAPDH	*CD69*
FCMR	*CD99*
CTD-2231E14.8	*TENM4*
RP11-16K12.1	*RGS2*
RP11-72I8.1	*MCF2*
RP4-549L20.3	*SYCP2L*
ANXA1	*DNM3*
NCF1	*CD36*
LGMN	*CH507-513H4.3*
AP000866.1	*CDKN2B-AS1*
ASIC3	*P3H4*
MGST1	*HIPK1-AS1*
CEBPD	*DACH1*

CD16+ monocytes	
FCGR3A	*HES4*
COTL1	*EVL*
RPS19	*FTH1*
HLA-DPA1	*CDKN1C*
LST1	*IFITM2*
ABI3	*GSTO2*
IFITM3	*HELLS*
SAT1	*KLF2*
HLA-DPB1	*FCER1G*
TMSB15B	*APITD1*
TCF7L2	*YBX1*
RHOB	*SPN*
ZNF502	*RHOC*
FGF14-AS2	*HLA-DRB1*
PKIB	*NIPSNAP3B*
MS4A7	*CSF1R*
HLA-DQA1	*AIF1*
TUBA1B	*DOC2A*
MTSS1	*HLA-DRA*
RNF32	*RP11-492E3.51*

Appendix Q: Classical and CD16^{+} cluster gene markers

Gene sets derived from the top most significant differentially expressed genes (highly expressed genes) defining classical and CD16^{+} monocyte subsets.

Classical subset 0	
MT-ND4L	LRRC34
TMEM176B	RP11-863K10.7
FAM26F	ZNF883
TMEM176A	LHFP
MT-ATP8	MED7-TICAM2
EEF1A1	AP000866.1
JUP	FAM225B
RBM3	FAM95B1
LYZ	PLVAP
PER1	NAV1
DDIT4	CTD-2231E14.8
KANSL1-AS1	LGALS2
HOXB-AS1	DTNA
CD1E	LBH
AREG	EAF1-AS1
GRASP	RMI2
C1QC	PDE4B
DZIP1L	FPR3
SEMA4C	FMNL2
SP2-AS1	GGT5

Classical subset 1	
XIST	ENPP2
RPS26	KLF6
NEAT1	MYBL1
DUSP6	FCMR
ANGPTL6	PRSS21
THBS1	RPL39
RETN	CALM2
RHOC	CTD-2017F17.2
PAPSS2	ZNF783
AC010642.1	EAF1
NEBL	FOS
MCL1	SLAMF7
VCAN	FCGBP
MAF	MMP9
AHNAK	ANXA1
JUND	ZDHHC1
CDC42EP1	RGS2
CCL5	AGBL3
AC090498.1	BEGAIN
RP11-585P4.5	MT-ND2

Classical subset 3	
FOLR3	VIM
DUSP6	TRIM58
FOSB	MT-ND1
SGK1	LILRB2
MYADM	RPS28
EGR1	ADGRE1
RP11-1143G9.4	CTB-61M7.2
KLF6	DUSP2
RP11-539G18.3	JUND
PLA2G7	DFNA5
EGR2	HP
JUP	KIF26B
CEP72	RP11-295G20.2
RRP12	PLAC8
AHNAK	AC090498.1
RPS29	HIVEP3
NRG1	CX3CR1
ARID5B	CD180
PTGDR2	JAK34
AK5	WDR76

Classical subset 2	
MTRNR2L8	RPL36
RPL41	TSPAN5
RPS28	CYP1B1
BRI3	RPS10
RPL36A	RPS17
MTRNR2L12	RP5-1157M23.2
AC090498.1	HIST1H4F
STXBP2	RPL7
UQCR11	C4orf48
OLFM1	CEBPD
CD79A	LYZ
RPLP2	SEL1L3
RPS15A	RP11-108M9.4
LTA4H	LGALS2
RPS21	EPHA1-AS1
RPS29	IGF1R
GPR82	CASC15
RPL34	ANO7
RP11-180M15.6	FFAR4
RPL39	RPL8

Appendix R: Classical subsets 0,1, 2 and 3 gene markers

Top 40 most significant differentially expressed genes (with high expression) defining classical monocyte subsets 0, 1, 2 and 3. Genes were identified with the FindAllMarkers function.

Classical subset 4	
EGR1	GSTM1
FOS	CD14
DUSP1	NDUFA11
THBS1	RGS2
TMEM176A	NRG1
S100A9	VSTM1
S100A8	CDA
FAM118A	HP
BLOC1S1	SOCS3
RPS10	CEBPB
AC004556.1	LYZ
IER2	ALOX5AP
TMEM176B	NDUFC2
JUNB	CKLF
ZFP36	PLBD1
SELL	SERPINA1
BTG2	BCL2A1
HSPA9	CEBPD
S100A12	PALLD
CYP1B1	CD99

Classical subset 5	
MTRNR2L8	SLPI
MTRNR2L12	CD52
PSME2	UQCR11
CEBPD	NCF1
CFD	IL1R2
VAMP5	FPR1
FKBP5	FAU
UBE2L6	CD74
PSMB9	CST3
HLA-DRB5	PPDPF
LY6E	RPS15A
HLA-DQB1	MYL12A
IFITM2	RPL34
S100A8	ENPP4
MT2A	ITLN1
TXNIP	RPL10
AIF1	RPS12
GPX1	SULT1A2
S100A9	CLEC4E
C20orf27	DLC1

Classical subset 6	
IFITM1	IFIT3
IFI44L	RNU12
IFITM3	FOS
JUN	CXCL8
IFI6	SAMD4A
FOSB	WARS
ISG15	APBB2
MT2A	XIST
LY6E	EPSTI1
LINC00936	PATL2
KLF4	CCL3
KLF6	RP5-1024G6.2
RP11-160E2.6	RPL39
CH17-373J23.1	ETV5
SIGLEC1	MX1
IFI27	SERPING1
EFCAB2	MARCKS
RP11-707G18.1	HIPK1-AS1
OTOF	ASB2
RPL41	SNED1

Classical subset 8	
GFI1B	LINC00989
PPBP	GNG11
FO538757.3	SEPT5
SDPR	ZNF682
CTD-2260A17.3	TMEM40
TUBB1	HIST1H2AC
PF4	NRGN
HIST1H2BJ	PDE5A
RP3-329E20.2	PLK4
ASB2	GRAP2
CALD1	PPP1R17
CCL5	NEO1
GP9	KIFC3
RP11-501J20.5	RP11-560J1.2
EAF1-AS1	HIST1H2AG
SPARC	HIST1H2BF
TUBA8	CD9
ENHO	CCL23
CMTM5	CECR2
SNED1	C6orf25

Appendix S: Classical subset 4,5,6 and 8 gene markers

Top 40 most significant differentially expressed genes (with high expression) defining classical monocyte subsets 4, 5, 6 and 8. Genes were identified with the FindAllMarkers function.

CD16+ cluster 1	
HLA-DRA	*SPATS2L*
HLA-DPB1	*FIGNL1*
CD74	*MRC1*
HLA-DPA1	*COLEC12*
HLA-DRB1	*RPL13*
KLRF1	*RNASE1*
UBTD2	*TMPRSS5*
RPLP0	*TVP23A*
CST3	*HLA-DQA2*
HLA-DQA1	*ZNF883*
HLA-DQB1	*RPSA*
C6orf25	*CTNNAL1*
ID3	*TXK*
HLA-DMA	*CCL24*
VPRBP	*BCL7A*
KDELC2	*TUBA8*
TACSTD2	*GRIP1*
GSTO2	*DTNA*
NAV1	*KIAA0226L*
RP5-892K4.1	*ADGRG5*

CD16+ cluster 2	
S100A12	*RP11-127B20.2*
S100A8	*CCR2*
S100A9	*QPCT*
VCAN	*AC004381.6*
AK5	*PADI4*
ALOX5AP	*RP3-329E20.2*
SH3BGRL2	*SLPI*
RP11-1275H24.3	*TIAF1*
CFAP53	*RP11-560J1.2*
LYZ	*TPST1*
SEPT5	*VNN2*
EMB	*DHRS3*
DCBLD2	*GNG11*
CMTM5	*RP11-357P18.2*
LGALS8-AS1	*SPAG16*
SERPINB2	*LMNB1*
CLEC1B	*RP11-315D16.2*
MGST1	*PDE7B*
DLGAP1-AS2	*NFE2*
PPP1R14A	*NBEA*

CD16+ cluster 3	
FCGR3A	*LINC01272*
CEP72	*RFPL4A*
TNFAIP6	*CBX2*
TMEM40	*PTCH2*
FGFR1	*CECR7*
LST1	*XKR7*
CCL24	*RP11-563J2.3*
UST	*FLT1*
UPK3B	*TMPRSS5*
AC010642.1	*HES4*
COTL1	*HIST1H2AC*
PPP4R1-AS1	*ADA*
LINC00926	*TPRG1*
KLRF1	*LINC00299*
GPR137C	*RP11-739L10.1*
FCER1G	*SH3RF3*
RPS19	*PPARG*
INTU	*MS4A4A*
IFITM2	*LILRB2*
SSX2IP	*SELENBP1*

Appendix T: CD16^{+} cluster 1, 2 and 3 gene markers

Top 40 most significant differentially expressed genes (with high expression) defining the CD16^{+} cluster 1, CD16^{+} cluster 2 and CD16^{+} cluster 3. Genes were identified with the FindAllMarkers function.

A

Classical

	Gene	p-value	r
1	*S100A9*	0	0.3540962
2	*S100A8*	0	0.2962894
3	*VCAN*	0	0.1948986
4	*S100A12*	0	0.1897861
5	*LYZ*	0	0.1555977
6	*NCF1*	0	0.1462546
7	*THBS1*	0	0.1402439
8	*FPR1*	0	0.1293146
9	*RGS2*	0	0.1277944
10	*ANXA1*	0	0.1248676
11	*SELL*	0	0.1215175
12	*CLEC4E*	0	0.1175598
13	*CYP1B1*	0	0.1090549
14	*TREM1*	0	0.1061166
15	*ALOX5AP*	0	0.09531519
16	*CYP27A1*	0	0.09341996
17	*NEAT1*	0	0.09335838
18	*VNN2*	0	0.08992161
19	*VSTM1*	0	0.07936851
20	*AQP9*	0	0.07832116
21	*MCEMP1*	0	0.07662363
22	*SLC2A3*	0	0.07102139
23	*PROK2*	0	0.06998234
24	*CALM2*	0	0.06820256
25	*CDA*	0	0.06798507

	Gene	p-value	r
26	*CAPG*	0	0.0677527
27	*GSTM1*	0	0.06660224
28	*EGR1*	0	0.06606579
29	*CD63*	0	0.06449385
30	*CDC42EP3*	0	0.06373847
31	*C5AR1*	0	0.06288581
32	*HP*	0	0.05998982
33	*QPCT*	0	0.05964737
34	*CCR1*	0	0.05950729
35	*CLEC4D*	0	0.05825484
36	*RNASE2*	0	0.05739062
37	*BCL2A1*	0	0.0571917
38	*SGK1*	0	0.05643918
39	*CD36*	0	0.05545276
40	*PADI4*	0	0.05503422
41	*PLBD1*	0	0.05352179
42	*HSPA1A*	0	0.05275409
43	*EMP3*	0.000118423	0.02461185
44	*SLC11A1*	0.000129899	0.02446664
45	*SERPINB2*	0.000165562	0.02408208
46	*IER3*	0.000335019	0.02293223
47	*CRIP1*	0.000530655	0.02215357
48	*GPX7*	0.001006998	0.02102708
49	*LINC01001*	0.001037049	0.02097411
50	*ADM*	0.001056867	0.02093994

B

Chemokine

	Gene	p-value	r
1	*LYZ*	0	0.1186921
2	*GPX1*	0	0.1134797
3	*VCAN*	0	0.1042436
4	*S100A9*	0	0.09425999
5	*S100A12*	0	0.08845004
6	*HLA-DRB5*	0	0.08823959
7	*F13A1*	0	0.08392409
8	*S100A8*	0	0.0767672
9	*CD36*	0	0.07558954
10	*RPL41*	0	0.072654
11	*CD52*	0	0.07175413
12	*S100A10*	0	0.0713684
13	*THBS1*	0	0.07067187
14	*MTRNR2L12*	0	0.07051816
15	*FOS*	0	0.0704789
16	*CEBPD*	0	0.0701979
17	*"MARCH1"*	0	0.06991663
18	*CD99*	0	0.0681737
19	*CD163*	0	0.06524373
20	*EGR1*	0	0.0606945
21	*CYP1B1*	0	0.06042583
22	*GIMAP7*	0	0.05951076
23	*HLA-DQA2*	0	0.05756887
24	*NFE2*	0	0.05631981
25	*SOCS3*	0	0.05627747

	Gene	p-value	r
26	*MGST1*	0	0.05588065
27	*MTRNR2L8*	0	0.055739
28	*GIMAP4*	0	0.05557214
29	*DYNLL1*	0	0.0547041
30	*LTA4H*	0	0.05441986
31	*NRG1*	0	0.05335468
32	*RAB3IP*	0.000111164	0.02471071
33	*NDUFC2*	0.0001126	0.02469068
34	*CTSB*	0.000112668	0.02468974
35	*ST20*	0.000118479	0.02461111
36	*RASSF4*	0.000120374	0.02458625
37	*PLD3*	0.000130289	0.02446193
38	*FCER1G*	0.000131591	0.02444627
39	*ABHD5*	0.000132226	0.02443868
40	*SOX4*	0.00013636	0.02439012
41	*SYTL3*	0.000139434	0.02435491
42	*HLA-DQB2*	0.000146472	0.02427697
43	*WSB1*	0.000147964	0.0242609
44	*LBX2-AS1*	0.000162738	0.02410953
45	*LINC01094*	0.000167119	0.02406712
46	*NCL*	0.00016939	0.02404555
47	*TLR2*	0.000173496	0.02400723
48	*PRKCB*	0.000175659	0.02398739
49	*SET*	0.000183581	0.02391662
50	*CTC-490G23.2*	0.000210532	0.02369566

Appendix U: Classical and chemokine expression module gene sets

Gene lists comprising genes most significantly correlated with transcripts A) CD14 and B) CCR2. To identify these gene sets HMisc was used to perform correlation analysis across the scRNA-seq dataset.

A

Immaturity

	Gene	p-value	r
1	*VCAN*	0	0.2230525
2	*LYZ*	0	0.2204229
3	*VIM*	0	0.1656648
4	*S100A12*	0	0.1586557
5	*CYP1B1*	0	0.1516236
6	*S100A9*	0	0.1454055
7	*S100A8*	0	0.1420065
8	*PLBD1*	0	0.1250456
9	*ANXA1*	0	0.1235247
10	*FOS*	0	0.1177827
11	*SELL*	0	0.1147687
12	*LTA4H*	0	0.1134296
13	*ACTG1*	0	0.1120995
14	*EMB*	0	0.1053814
15	*QPCT*	0	0.1052462
16	*STEAP4*	0	0.1051506
17	*VNN2*	0	0.1050617
18	*CES1*	0	0.1035903
19	*MGST1*	0	0.1034717
20	*AHNAK*	0	0.1025288
21	*FAM198B*	0	0.1005864
22	*RP11-1143G9.4*	0	0.1003934
23	*SLC2A3*	0	0.09942823
24	*F13A1*	0	0.09547894
25	*RGS2*	0	0.09479519

	Gene	p-value	r
26	*STXBP2*	0	0.09455706
27	*CDC42EP3*	0	0.09312927
28	*PADI4*	0	0.09301619
29	*MCL1*	0	0.08973848
30	*NRG1*	0	0.08764174
31	*CAPG*	0	0.08763074
32	*ZFAND5*	0	0.08749113
33	*MYADM*	0	0.08720294
34	*FOLR3*	0	0.08546614
35	*ANPEP*	0	0.08298945
36	*PCBP1*	0	0.08289927
37	*F5*	0	0.082163
38	*SGK1*	0	0.08173036
39	*CKAP4*	0	0.08160657
40	*KLF6*	0	0.08159841
41	*CD93*	0	0.08158575
42	*ZFP36L1*	0	0.08131214
43	*MXD1*	0	0.08119176
44	*CLU*	0	0.08080538
45	*DMXL2*	0	0.080356
46	*NFE2*	0	0.07959984
47	*TLR2*	0	0.07957237
48	*HK3*	0	0.078885
49	*STAB1*	0	0.07882184
50	*EEF1A1*	0	0.07644538

B

Non-classical

	Gene	p-value	r
	CDKN1C	0	0.4363149
2	*HES4*	0	0.3710887
3	*RHOC*	0	0.3327486
4	*C1QA*	0	0.3322669
5	*IFITM3*	0	0.3062901
6	*FCGR3B*	0	0.2957327
7	*HLA-DPA1*	0	0.2935616
8	*TCF7L2*	0	0.2876543
9	*MTSS1*	0	0.2858853
10	*NR4A1*	0	0.2815628
11	*EVL*	0	0.2790972
12	*HLA-DPB1*	0	0.2757661
13	*ABI3*	0	0.2749675
14	*IFITM1*	0	0.2734287
15	*FAM26F*	0	0.2697689
16	*C1QB*	0	0.2677782
17	*LST1*	0	0.2415983
18	*RHOB*	0	0.2398474
19	*TUBA1B*	0	0.239239
20	*SAT1*	0	0.2340617
21	*CD79B*	0	0.2293679
22	*ADA*	0	0.2278917
23	*IFITM2*	0	0.2258535
24	*RPS19*	0	0.2230399
25	*KLF2*	0	0.2203017

	Gene	p-value	r
26	*MS4A7*	0	0.2164829
27	*PLAC8*	0	0.2123733
28	*LY6E*	0	0.2118218
29	*HLA-DQA1*	0	0.2107844
30	*INSIG1*	0	0.2049673
31	*APOBEC3A*	0	0.2043203
32	*WARS*	0	0.2027012
33	*CSF1R*	0	0.2011799
34	*MGLL*	0	0.2001259
35	*C1QC*	0	0.1957528
36	*HLA-DRA*	0	0.195164
37	*CTSL*	0	0.1942969
38	*AIF1*	0	0.1935414
39	*PSME2*	0	0.1881185
40	*BATF3*	0	0.1872406
41	*COTL1*	0	0.1872272
42	*NPC2*	0	0.1846693
43	*HLA-DRB1*	0	0.1838751
44	*MARCKSL1*	0	0.1838687
45	*VAMP5*	0	0.1837177
46	*LYPD2*	0	0.1836016
47	*SPN*	0	0.1825899
48	*MT2A*	0	0.1811472
49	*FMNL2*	0	0.1803491
50	*NAP1L1*	0	0.1789464

Appendix V: Immaturity and non-classical expression module gene sets

Gene lists comprising genes most significantly correlated with transcripts A) CD36 and B) FCGR3A. To identify these gene sets HMisc was used to perform correlation analysis across the scRNA-seq dataset.

Antigen-presentation

A

	Gene	p-value	r
1	*CD74*	0	0.652918
2	*HLA-DPA1*	0	0.633314
3	*HLA-DPB1*	0	0.631697
4	*HLA-DRB1*	0	0.599275
5	*HLA-DQB1*	0	0.47021
6	*HLA-DQA1*	0	0.465916
7	*HLA-DMA*	0	0.449291
8	*HLA-DRB5*	0	0.391931
9	*HLA-DQA2*	0	0.34704
10	*HLA-DMB*	0	0.31154
11	*NPC2*	0	0.309258
12	*CST3*	0	0.307152
13	*PSME2*	0	0.303159
14	*CPVL*	0	0.296534
15	*FTH1*	0	0.290179
16	*FAM26F*	0	0.286123
17	*IFITM3*	0	0.257475
18	*YBX1*	0	0.253899
19	*LY6E*	0	0.232801
20	*ABI3*	0	0.231283
21	*LGALS2*	0	0.228076
22	*HSPA8*	0	0.224607
23	*CLEC10A*	0	0.224514
24	*PTMA*	0	0.220667
25	*RPL10*	0	0.217985

	Gene	p-value	r
26	*SLC25A5*	0	0.21717
27	*PPIA*	0	0.215627
28	*PSMB9*	0	0.213649
29	*TUBA1B*	0	0.207112
30	*PID1*	0	0.206784
31	*LAP3*	0	0.202518
32	*ITGA4*	0	0.199419
33	*RPS5*	0	0.198481
34	*AIF1*	0	0.197201
35	*FCGR3A*	0	0.195164
36	*HSP90AB1*	0	0.194039
37	*MT2A*	0	0.193137
38	*CSF1R*	0	0.193125
39	*SAT1*	0	0.193088
40	*UBE2L6*	0	0.19264
41	*NPM1*	0	0.190349
42	*HSP90AA1*	0	0.189596
43	*GBP1*	0	0.189574
44	*RPL12*	0	0.189257
45	*RPL5*	0	0.187842
46	*RPS4X*	0	0.187413
47	*RPLP0*	0	0.18647
48	*RARRES3*	0	0.186242
49	*EEF1B2*	0	0.18612
50	*MTSS1*	0	0.185691

Complement

B

	Gene	p value	r
1	*C1QB*	0	0.7309149
2	*C1QC*	0	0.6509985
3	*FCGR3A*	0	0.3322669
4	*HES4*	0	0.2101896
5	*RHOC*	0	0.1975716
6	*HLA-DPA1*	0	0.1938667
7	*CDKN1C*	0	0.1891902
8	*HLA-DPB1*	0	0.1891486
9	*HLA-DQA1*	0	0.1846262
10	*TUBA1B*	0	0.1710147
11	*NR4A1*	0	0.1703159
12	*ABI3*	0	0.1701838
13	*TCF7L2*	0	0.1646684
14	*IFITM3*	0	0.1637057
15	*RHOB*	0	0.1566014
16	*FPR3*	0	0.1562388
17	*EVL*	0	0.1561612
18	*MTSS1*	0	0.1551419
19	*ENHO*	0	0.152196
20	*FMNL2*	0	0.1505719
21	*FAM26F*	0	0.1487458
22	*CLEC4F*	0	0.1443035
23	*FCGR3B*	0	0.1441915
24	*RPS19*	0	0.1425803
25	*INSIG1*	0	0.1424591

	Gene	p value	r
26	*CD74*	0	0.1392413
27	*MGLL*	0	0.1352063
28	*HLA-DRA*	0	0.1352058
29	*HLA-DRB1*	0	0.13342
30	*CD79B*	0	0.1306625
31	*KLF2*	0	0.1272809
32	*CST3*	0	0.126971
33	*LYPD2*	0	0.126566
34	*GFRA2*	0	0.1249393
35	*IFITM1*	0	0.1240068
36	*HLA-DQA2*	0	0.1229851
37	*HLA-DOA*	0	0.1226096
38	*PLAC8*	0	0.1209204
39	*SAT1*	0	0.1205138
40	*CTSL*	0	0.118702
41	*CAMK1*	0	0.1182639
42	*HSPB1*	0	0.1173917
43	*YBX1*	0	0.1173083
44	*MARCKSL1*	0	0.1172543
45	*CSF1R*	0	0.1170002
46	*ADA*	0	0.1167382
47	*GPBAR1*	0	0.1156611
48	*HLA-DQB1*	0	0.1149744
49	*LY6E*	0	0.1121802
50	*HSPA8*	0	0.1119899

Appendix W: Antigen presentation and complement expression module gene sets

Gene lists comprising genes most significantly correlated with transcripts A) HLA-DRA and B) C1QA. To identify these gene sets HMisc was used to perform correlation analysis across the scRNA-seq dataset.

IFN-signalling

A

	Gene	P-value	r
1	*GBP1*	0	0.2721542
2	*LAP3*	0	0.2480748
3	*GBP5*	0	0.2356551
4	*MT2A*	0	0.2176687
5	*TNFSF10*	0	0.2130676
6	*GBP4*	0	0.2029185
7	*LY6E*	0	0.2012912
8	*EPSTI1*	0	0.1991645
9	*RNF213*	0	0.1966924
10	*UBE2L6*	0	0.1960587
11	*GBP2*	0	0.1946112
12	*TYMP*	0	0.1904422
13	*WARS*	0	0.1897429
14	*PSMB9*	0	0.1867021
15	*PARP14*	0	0.1850424
16	*ISG15*	0	0.1762283
17	*IFI44L*	0	0.1757887
18	*PSME2*	0	0.1736286
19	*SAT1*	0	0.17355
20	*APOL6*	0	0.171881
21	*VAMP5*	0	0.1704598
22	*FAM26F*	0	0.16639
23	*MX1*	0	0.1623132
24	*XAF1*	0	0.1619949
25	*SCO2*	0	0.1598178

	Gene	P-value	r
26	*SAMD9L*	0	0.1570261
27	*IFIT3*	0	0.1563043
28	*PARP9*	0	0.1541184
29	*IFI6*	0	0.1535874
30	*STAT2*	0	0.1533287
31	*IRF1*	0	0.1516237
32	*TAP1*	0	0.1499642
33	*IFIT2*	0	0.1456199
34	*SERPING1*	0	0.1424546
35	*IFITM3*	0	0.141452
36	*PSTPIP2*	0	0.1390403
37	*IFITM1*	0	0.1377443
38	*TRIM22*	0	0.1359278
39	*HLA-DPB1*	0	0.1346789
40	*IFI44*	0	0.1332055
41	*HLA-DPA1*	0	0.1322464
42	*MARCKS*	0	0.131148
43	*OAS2*	0	0.1307654
44	*HERC5*	0	0.1271127
45	*HLA-DQB1*	0	0.1237313
46	*SOD2*	0	0.1214268
47	*HLA-DRA*	0	0.1207624
48	*FYB*	0	0.1200286
49	*SECTM1*	0	0.1187663
50	*APOBEC3A*	0	0.1183734

Pro-inflammatory

B

Gene	p-value	r
PSME2	0	0.1050807
MT2A	0	0.1009552
HLA-DPA1	0	0.09726773
HLA-DPB1	0	0.09477335
HLA-DQB1	0	0.09162297
CD74	0	0.09015225
LY6E	0	0.08890579
UBE2L6	0	0.08446459
VAMP5	0	0.08403667
FAM26F	0	0.08231295
HLA-DRA	0	0.08113001
IFITM1	0	0.0797831
GBP1	0	0.07973408
RARRES3	0	0.07677827
HLA-DRB1	0	0.07391884
SERPING1	0	0.07175031
WARS	0	0.07091867
LAP3	0	0.07055506
GBP5	0	0.0702672
IFITM3	0	0.06834608
GBP4	0	0.06801156
PSMB9	0	0.06796467
IFI6	0	0.06690971
HLA-DQA1	0	0.06387227
HLA-DMA	0	0.06086987

	Gene	p-value	r
26	*STAT1*	0	0.06016243
27	*MRC1*	0	0.0577529
28	*SAT1*	0	0.0566649
29	*GBP2*	0	0.05663586
30	*CALR*	0	0.05657529
31	*IFI44L*	0	0.05657143
32	*FCGR3A*	0	0.05615893
33	*NPC2*	0	0.05550643
34	*RPLP0*	0	0.05523197
35	*APOL3*	0	0.05454983
36	*SAMD9L*	0	0.05449761
37	*EPSTI1*	0	0.05367257
38	*NDUFA12*	0	0.0532751
39	*ISG15*	0	0.05325264
40	*APOL6*	0	0.0529865
41	*CXCL16*	0.000101096	0.02485844
42	*GNLY*	0.000105661	0.02478982
43	*CH507-513H4.4*	0.000105684	0.02478948
44	*MS4A7*	0.000122716	0.02455604
45	*GFRA2*	0.000131482	0.02444757
46	*TBC1D8*	0.000131703	0.02444492
47	*ETV7*	0.00013711	0.02438146
48	*PID1*	0.000149017	0.02424965
49	*MAP3K13*	0.000150689	0.02423194
50	*GCH1*	0.000150782	0.02423096

Appendix X: Interferon-signalling and pro-inflammatory module expression gene sets

Gene lists comprising genes most significantly correlated with transcripts A) STAT1 and B) CD38. To identify these gene sets HMisc was used to perform correlation analysis across the scRNA-seq dataset.

Macrophage differentiation

A

	Gene	p-value	r
	HLA-DPB1	0	0.2299861
2	*HLA-DPA1*	0	0.2283967
3	*FCGR3A*	0	0.2011799
4	*CD74*	0	0.1944191
5	*HLA-DRA*	0	0.1931246
6	*HLA-DRB1*	0	0.1900243
7	*HLA-DQA1*	0	0.1838944
8	*RHOC*	0	0.16978
9	*YBX1*	0	0.1571732
10	*HLA-DQB1*	0	0.1491038
11	*TUBA1B*	0	0.1489573
12	*NR4A1*	0	0.1454522
13	*JUNB*	0	0.141415
14	*COTL1*	0	0.1410522
15	*EEF1A1*	0	0.1403839
16	*ABI3*	0	0.1403008
17	*FTH1*	0	0.1398849
18	*HLA-DMA*	0	0.1389
19	*ZFP36*	0	0.1382152
20	*MTSS1*	0	0.1369963
21	*CDKN1C*	0	0.1302754
22	*IFITM3*	0	0.1294739
23	*HES4*	0	0.1280335
24	*KLF2*	0	0.1279753
25	*HSPA8*	0	0.127408

	Gene	p-value	r
26	*RPS19*	0	0.1234545
27	*EVL*	0	0.1221682
28	*"SEPT9"*	0	0.1217586
29	*NME2*	0	0.1209105
30	*CAMK1*	0	0.1196463
31	*FAM26F*	0	0.1194509
32	*C1QA*	0	0.1170002
33	*DOK2*	0	0.1152695
34	*IER2*	0	0.1147282
35	*RHOB*	0	0.1141968
36	*ITGA4*	0	0.1140823
37	*CLEC10A*	0	0.1136303
38	*MARCKS*	0	0.1134178
39	*CST3*	0	0.1128197
40	*CPVL*	0	0.1123041
41	*HSP90AB1*	0	0.111527
42	*NAP1L1*	0	0.1108062
43	*MGLL*	0	0.1101987
44	*LIPA*	0	0.1094306
45	*HLA-DQA2*	0	0.1093063
46	*HNRNPA1*	0	0.1091491
47	*PEA15*	0	0.1083687
48	*NPC2*	0	0.1079507
49	*TCF7L2*	0	0.10635
50	*ID2*	0	0.1061262

Immunosuppression

B

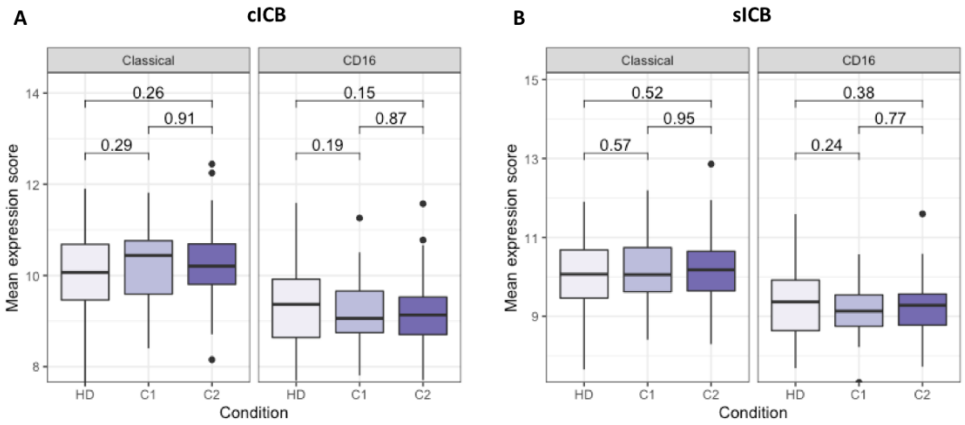
	Gene	p-value	r
1	*IFI30*	0.000131021	0.0244531
2	*CFAP53*	0.000166392	0.02407408
3	*EEF1G*	0.000206836	0.02372433
4	*CD300E*	0.000217299	0.02364436
5	*EGR3*	0.000235693	0.02351216
6	*JUNB*	0.000271954	0.02327772
7	*EGR1*	0.000326654	0.02297437
8	*MGLL*	0.000330536	0.02295469
9	*THBD*	0.000365355	0.02278728
10	*HIST2H2AA4*	0.000370515	0.02276374
11	*RP11-707G18.1*	0.000484111	0.02231089
12	*PHLDA2*	0.000494568	0.02227436
13	*RAPH1*	0.000522764	0.02217931
14	*IRF1*	0.000606508	0.02192281
15	*METRNL*	0.000650873	0.02180002
16	*PHTF2*	0.000654387	0.02179063
17	*SLC11A1*	0.000679726	0.02172427
18	*ZNRF3*	0.00069825	0.0216772
19	*CCL2*	0.000759142	0.02153026
20	*SETBP1*	0.000760045	0.02152816
21	*PTGS2*	0.000855056	0.02131963
22	*NBPF26*	0.000869045	0.02129076
23	*STEAP4*	0.000892042	0.02124422
24	*DZIP1L*	0.000994007	0.02105043
25	*CLEC4F*	0.001053122	0.02094635

	Gene	p-value	r
26	*NLRP3*	0.001089525	0.02088492
27	*DRAP1*	0.001091998	0.02088082
28	*NAMPT*	0.001200352	0.02070893
29	*SMC2*	0.001292058	0.02057431
30	*CST3*	0.001445693	0.02036739
31	*C15orf39*	0.001587049	0.02019419
32	*HLA-DMA*	0.001790928	0.01996789
33	*LILRB1*	0.002047157	0.01971491
34	*ADA*	0.002093375	0.0196724
35	*DZIP3*	0.002345508	0.01945464
36	*IRF7*	0.002410225	0.01940221
37	*PHF19*	0.002410331	0.01940212
38	*NFKBID*	0.002460069	0.0193627
39	*RPL17*	0.002482173	0.01934541
40	*DOK2*	0.00260559	0.01925141
41	*HBEGF*	0.002792219	0.01911671
42	*TPM4*	0.003065615	0.01893355
43	*TMEM158*	0.003165535	0.0188703
44	*WDR60*	0.003174837	0.01886451
45	*DUSP6*	0.003259182	0.01881266
46	*APCDD1*	0.0032766	0.01880211
47	*NAB2*	0.003533887	0.01865189
48	*VSIG4*	0.003570112	0.01863154
49	*NRROS*	0.003709089	0.01855519
50	*PNPLA6*	0.003778133	0.01851822

Appendix Y: Macrophage differentiation and immunosuppression module gene sets

Gene lists comprising genes most significantly correlated with transcripts A) CSF1R and B) IL10. To identify these gene sets HMisc was used to perform correlation analysis across the scRNA-seq dataset.

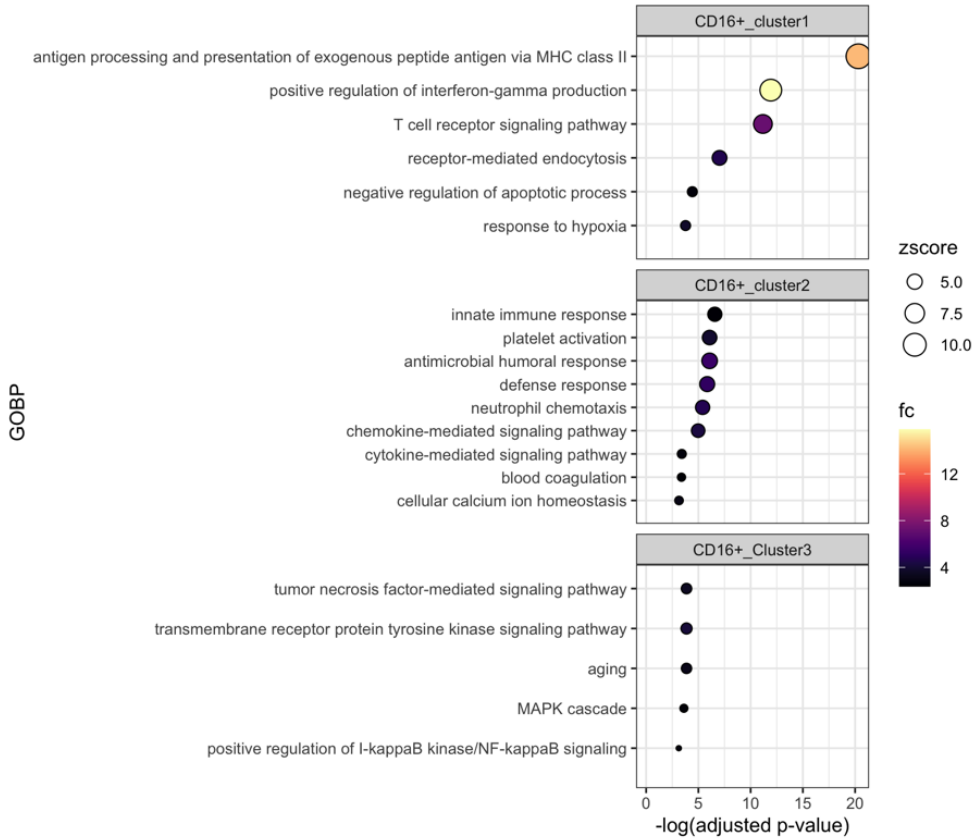
Classical and CD16^{+} cluster expression scores in bulk RNA-seq data by treatment



Appendix Z: Boxplots showing classical and CD16^{+} expression scores by treatment

Boxplots to show classical and CD16^{+} expression scores in bulk RNA-seq data across Controls (45 'HD'), C1 and C2 patient samples by treatment: A) cICB, 56 C1 patients, 51 C2 patients; B) sICB, 58 C1 patients, 47 C2 patients (Wilcoxon signed-rank test). Classical and CD16^{+} expression scores were calculated by taking the top 40 most significantly highly expressed gene markers and finding the geometric mean expression of these genes within the normalised RNA-seq data.

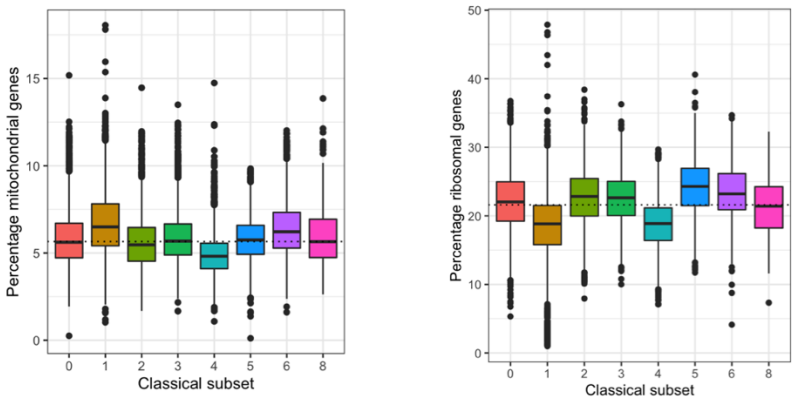
CD16^{+} clusters 1-3 unqiuely enriched pathways



Appendix AA: CD16^{+} cluster-wise enriched pathways

Plot to show key GOBP pathways which show cluster-wise enrichment across CD16^{+} monocyte clusters 1, 2 and 3.

Percentage mitochondrial and ribosomal genes across classical subsets



Appendix BB: Percentage mitochondrial and ribosomal genes in classical monocyte subsets
Boxplots to show percentage mitochondrial (left) and ribosomal genes (right) per cell by classical monocyte subset 0-6 and 8. Dotted line indicated median score across all cells.

Sample	ID	Batch	Cycle	Agent	Age	Sex	6MP	AI_SE	5hmC:C	5hmC:5mC
86a	86	2	C1	cICB	65	F	NO	YES	3	16.5
88a	88	2	C1	sICB	69	M	NO	NO	4.9	49
146a	146	2	C1	cICB	34	F	YES	YES	44	H
153a	153	2	C1	cICB	58	F	NO	YES	3.7	13.67
157a	157	2	C1	sICB	72	F	NO	NO	3	12
160a	160	2	C1	sICB	96	F	YES	NO	3.9	9.5
71a	71	2	C1	sICB	84	F	NO	NO	3.12	H
110a	110	2	C1	sICB	79	F	NO	NO	4.1	20.5
124a	124	2	C1	cICB	61	M	NO	YES	3.9	H
137a	137	2	C1	sICB	95	M	NO	NO	2.9	43
144a	144	2	C1	cICB	55	F	NO	YES	2.4	H
99a	99	2	C1	cICB	62	F	YES	YES	4.5	H
101a	101	2	C1	sICB	71	F	NO	NO	4.9	54
156b	156	2	C2	cICB	71	M	NO	NA	2.5	H
157b	157	2	C2	sICB	72	F	NO	NO	2.4	5.7
71b	71	2	C2	sICB	84	F	NO	NO	2.5	H
110b	110	2	C2	sICB	79	F	NO	NO	4.5	13.5
101b	101	2	C2	sICB	71	F	NO	NO	8	16
804	804	2	HD	NA	46	F	NA	NA	3	2.4
811	811	2	HD	NA	51	F	NA	NA	2.8	H
10b	64	1	C2	sICB	72	M	NO	YES	6.8	H
11a	74	1	C1	sICB	71	M	NO	NO	2.1	7
11b	74	1	C2	sICB	71	M	NO	NO	8.25	H
12a	75	1	C1	sICB	67	F	NO	NO	4.2	4.2
12b	75	1	C2	sICB	67	F	NO	NO	2.6	6
13b	76	1	C2	cICB	70	M	NO	YES	6.9	H
14b	77	1	C2	cICB	57	M	NO	YES	7.5	19
15b	78	1	C2	cICB	32	F	NO	YES	3.7	5.9
16a	79	1	C1	sICB	81	F	YES	NO	3.8	8
17a	80	1	C1	sICB	82	M	NO	YES	8.9	20.8
17b	80	1	C2	sICB	82	M	NO	YES	33.6	33.8
18a	81	1	C1	sICB	72	F	NO	NO	12	24.5
18b	81	1	C2	sICB	72	F	NO	NO	31.6	56.5
19a	83	1	C1	cICB	45	F	NO	YES	3.9	11.9
19b	83	1	C2	cICB	45	F	NO	YES	15.6	25.5
20b	84	1	C2	sICB	43	F	NO	YES	2.8	12
22b	85	1	C2	cICB	43	F	YES	YES	10.3	5
23a	73	1	C1	sICB	70	F	YES	NO	2.43	3
23b	73	1	C2	sICB	70	F	YES	NO	5.25	4.8
2a	54	1	C1	sICB	86	M	YES	NO	2.7	6
2b	54	1	C2	sICB	86	M	YES	NO	3.1	H
3b	55	1	C2	cICB	67	M	YES	YES	3.8	3.25
4a	57	1	C1	cICB	47	M	NO	YES	6.1	7.7
4b	57	1	C2	cICB	47	M	NO	YES	4	4
5b	58	1	C2	cICB	67	M	YES	YES	6.5	16
6a	59	1	C1	cICB	65	M	NO	YES	23.2	22.5
6b	59	1	C2	cICB	65	M	NO	YES	10.6	29.5
T4	T4	1	HD	NA	22	F	NA	NA	10.8	H
T5	T5	1	HD	NA	31	F	NA	NA	9.7	18
T6	T6	1	HD	NA	22	F	NA	NA	4.5	H

Appendix CC: 5hmC cohort details

Table to show 5hmC cohort details. Batch indicated sequencing batch, 6MP indicates progression by six months, AI_SEs indicates irAEs prior to cycle 5. '5hmC:C' indicates ratio of 5hmC spike-in oligo reads to control spike-in oligo read to indicate relative enrichment. '5hmC:5mC indicates ratio of 5hmC spike-in oligo reads to 5mC control spike-in oligo read to indicate relative enrichment. H indicates 5hmC to 5mC control ratio is versus a zero 5mC enrichment value.

5hmC pooled non-enriched samples

Sample	ID	Batch	Cycle	Agent	Age	Sex	6MP	irAE
110a	110	2	C1	sICB	79	F	NO	NO
122a	122	2	C1	cICB	71	M	YES	YES
124a	124	2	C1	cICB	61	M	NO	YES
137a	137	2	C1	sICB	95	M	NO	NO
138a	138	2	C1	cICB	56	F	YES	NO
110b	110	2	C2	sICB	79	F	NO	NO
122b	122	2	C2	cICB	71	M	YES	YES
124b	124	2	C2	cICB	61	M	NO	YES
137b	137	2	C2	sICB	95	M	NO	NO
138b	138	2	C2	cICB	56	F	YES	NO

Appendix DD: Non-5hmC enriched samples

Input cfDNA from these samples were pooled and adapter ligated and amplified, but not 5hmC enriched. Untreated (C1) and treated (C2) samples were pooled together and adapter ligated and sequenced together. Following sequencing, .BAM files (from both C1 and C2 pooled samples) were merged to create a merged non-5hmC enriched .BAM file.

5mC cohort details

Sample	ID	Cycle	Agent	Batch	Sex	Age	6MP	AI_c5	Specificity (%)
58a	58	C1	cICB	Two	1	67	YES	YES	91
58b	58	C2	cICB	Two	1	67	YES	YES	97
62a	62	C1	sICB	Two	2	67	YES	NO	94
62b	62	C2	sICB	Two	2	67	YES	NO	96
88a	88	C1	sICB	Two	1	69	NO	NO	95
88b	88	C2	sICB	Two	1	69	NO	NO	92
99a	99	C1	cICB	Two	2	62	YES	YES	94
99b	99	C2	cICB	Two	2	62	YES	YES	94
55a	55	C1	cICB	Two	1	67	YES	YES	96
55b	55	C2	cICB	Two	1	67	YES	YES	97
71a	71	C1	sICB	Two	2	84	NO	NO	94
71b	71	C2	sICB	Two	2	84	NO	NO	97
122a	122	C1	cICB	Two	1	71	YES	YES	92
122b	122	C2	cICB	Two	1	71	YES	YES	92
805	805	C	None	Two	2	50	NA	NA	97
812	812	C	None	Two	2	56	NA	NA	97
813	813	C	None	Two	2	64	NA	NA	97
HD17	HD17	C	None	Two			NA	NA	95
HD18	HD18	C	None	Two			NA	NA	96
HD19	HD19	C	None	Two			NA	NA	97
60a	60	C1	sICB	One	1	69	YES	NO	94
60b	60	C2	sICB	One	1	69	YES	NO	93
138a	138	C1	cICB	One	2	56	YES	NO	97
138b	138	C2	cICB	One	2	56	YES	NO	94
153a	153	C1	cICB	One	2	58	NO	YES	95
153b	153	C2	cICB	One	2	58	NO	YES	96
137a	137	C1	sICB	One	1	95	NO	NO	94
137b	137	C2	sICB	One	1	95	NO	NO	95
157a	157	C1	sICB	One	2	72	NO	NO	91
157b	157	C2	sICB	One	2	72	NO	NO	98
158a	158	C1	cICB	One	1	68	YES	YES	96
158b	158	C2	cICB	One	1	68	YES	YES	94
160a	160	C1	sICB	One	2	96	YES	NO	93
160b	160	C2	sICB	One	2	96	YES	NO	95
811	811	C	None	One	2	51	NA	NA	95
HD16	HD16	C	None	One	1	20	NA	NA	92
81a	81	C1	sICB	Three	2	72	NO	NO	94
81b	81	C2	sICB	Three	2	72	NO	NO	93
83a	83	C1	cICB	Three	2	45	NO	YES	93
83b	83	C2	cICB	Three	2	45	NO	YES	94
84a	84	C1	sICB	Three	2	43	NO	YES	94
84b	84	C2	sICB	Three	2	43	NO	YES	97
85a	85	C1	cICB	Three	1	56	YES	YES	92
85b	85	C2	cICB	Three	1	56	YES	YES	94
87a	87	C1	cICB	Three	1	68	YES	YES	92
87b	87	C2	cICB	Three	1	68	YES	YES	89
90a	90	C1	cICB	Three	1	66	NO	YES	98
90b	90	C2	cICB	Three	1	66	NO	YES	98
104a	104	C1	cICB	Three	1	55	NO	YES	93
104b	104	C2	cICB	Three	1	55	NO	YES	91
108a	108	C1	cICB	Three	1	63	NO	YES	97
108b	108	C2	cICB	Three	1	63	NO	YES	95
132a	132	C1	cICB	Three	1	48	NO	YES	92
132b	132	C2	cICB	Three	1	48	NO	YES	91
174a	174	C1	sICB	Three	1	81	NO	NO	96
174b	174	C2	sICB	Three	1	81	NO	NO	87

Appendix EE: 5mC cohort details

Table to show 5hmC cohort details. Batch indicates sequencing batch, agent denotes single or combination ICB therapy. 6MP indicates progression by six months, AI_c5 indicated irAEs prior to cycle 5. Specificity (%) indicates 5mC pull-down specificity. Non-5mC immunoprecipitated samples from samples 137a, 137b, 153a, 153b, 138a, 138b, 60a and 60b were pooled in equal quantities by condition (untreated/treated) and sequenced as a background non-5mC-enriched control.

A

Chromosome	Start	End	Conc_C1	Conc_C2	FDR	Genomic annotation	Gene
chr22	48916131	48917502	6.74	7.4	0.000121	Distal Intergenic	*LINC01310*
chr2	106465724	106468626	7.8	8.26	0.000121	Promoter (<=1kb)	*RGPD3*
chr8	1350053	1352429	7.18	7.84	0.000121	Intron	*LOC286083*
chr17	79305808	79308666	5.52	6.3	0.000121	Promoter (<=1kb)	*RBFOX3*
chr5	2490265	2491674	6.19	6.81	0.000166	Distal Intergenic	*LSINCT5*
chr10	346285	350076	6.81	7.35	0.00039	Exon	*DIP2C*
chr2	232897480	232902759	7.36	7.89	0.000551	Promoter (<=1kb)	*NGEF*
chr12	130531719	130533911	7.67	8.18	0.000626	Intron	*RIMBP2*
chr15	23439894	23442122	7.67	8.16	0.000702	5' UTR	*GOLGA6L2*
chr1	3210785	3215284	7	7.44	0.000702	Intron	*PRDM16*
chr2	10008316	10017959	8.58	8.99	0.000785	Distal Intergenic	*GRHL1*
chr8	55449225	55455417	6.92	7.4	0.000852	Promoter (<=1kb)	*SBF1P1*
chr9	34370738	34373802	7.19	7.66	0.000917	5' UTR	*MYORG*
chr8	143070760	143074047	7.17	7.6	0.000984	Distal Intergenic	*LY6L*
chr4	189845865	189847033	5.98	6.58	0.000984	Intron	*LINC01596*
chr6	166897812	166905459	8.84	9.22	0.000984	Promoter (<=1kb)	*RPS6KA2-AS1*
chr3	194943867	194946715	5.32	6.1	0.000984	Distal Intergenic	*XXYLT1*
chr2	106342181	106343718	5.57	6.32	0.000984	Distal Intergenic	*LOC402096*
chr20	64268306	64270679	6.54	7.15	0.000984	Promoter (<=1kb)	*PCMTD2*
chr5	1287968	1294504	8.42	8.85	0.000984	Promoter (<=1kb)	*TERT*

B

Chromosome	Start	End	Conc_C1	Conc_C2	FDR	Genomic annotation	Gene
chr10	107809343	107809665	0.47	2.61	0.000263	Intron	*LINC01435*
chr18	49842783	49843471	1.03	3.06	0.000412	Exon	*MYO5B*
chr5	10538521	10539085	1.73	3.58	0.000449	Distal Intergenic	*LINC02213*
chr19	32828798	32829873	2.7	4.09	0.000471	Promoter (2-3kb)	*TDRD12*
chr13	107678560	107680491	2.79	4.27	0.000755	Intron	*MIR1267*
chr18	62986938	62988405	3.97	4.85	0.000766	Distal Intergenic	*PHLPP1*
chr10	118407930	118409184	2.85	4.24	0.000971	Distal Intergenic	*LINC00867*
chr17	13096009	13097085	1.19	2.97	0.00105	Exon	*ELAC2*
chr8	20823460	20824226	2.3	3.91	0.00138	Distal Intergenic	*LOC102467222*
chr19	58165470	58165936	1.14	3.28	0.00141	Intron	*ZNF329*
chr21	17261186	17261445	2.26	3.52	0.00159	Distal Intergenic	*LINC01549*
chr12	1848151	1849035	2.58	3.84	0.00209	Intron	*CACNA2D4*
chr15	69380477	69380949	2.48	3.66	0.00283	Intron	*PAQR5*
chr8	80091163	80091551	1.74	3.28	0.00359	Intron	*TPD52*
chr14	64764034	64764382	1.67	3	0.00407	Promoter (<=1kb)	*SPTB*
chr18	57118335	57120119	3.3	4.53	0.00447	Distal Intergenic	*BOD1L2*
chr10	114247516	114248387	1.92	3.39	0.00489	Intron	*VWA2*
chr15	70604221	70604928	2.27	3.84	0.00591	Distal Intergenic	*SALRNA3*
chr17	82533177	82533997	2.6	3.94	0.00621	Intron	*FOXK2*
chr17	47049185	47050111	2.66	3.74	0.0067	Exon	*RPRML*

C

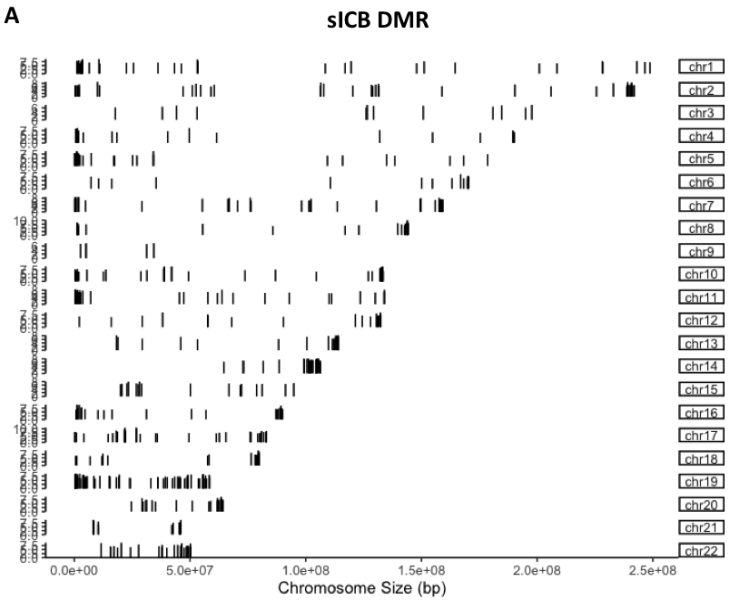
Chromosome	Start	End	Conc_C1	Conc_C2	FDR	annotation	SYMBOL
chr9	120887138	120888358	4.45	1.05	1.06E-08	Intron	*PHF19*
chr9	123270211	123271700	3.54	0.19	1.06E-08	Promoter (1-2kb)	*STRBP*
chr9	120894308	120895438	5.65	2.17	1.86E-07	Promoter (<=1kb)	*PHF19*
chr9	126414286	126416956	4.14	0.66	1.86E-07	Intron	*MVB12B*
chr9	122202160	122203352	3.44	0.47	1.86E-07	3' UTR	*LHX6*
chr9	123390415	123392514	4.83	1.68	1.86E-07	Intron	*MIR601*
chr9	120151363	120152494	3.7	0.7	1.86E-07	Distal Intergenic	*MIR147A*
chr9	121982741	121983895	3.22	-0.04	1.86E-07	Intron	*TTLL11*
chr9	122710880	122711819	4.45	1.11	1.86E-07	Distal Intergenic	*OR1L4*
chr9	123508080	123510154	4.11	0.87	1.86E-07	Intron	*MIR7150*
chr9	122052417	122054323	4.45	1.13	1.86E-07	Intron	*TTLL11*
chr9	127121614	127122068	3.19	0.13	3.11E-07	Promoter (<=1kb)	*ANGPTL2*
chr9	123645117	123647498	3.7	0.08	3.14E-07	Intron	*DENND1A*
chr9	121781852	121786158	5.51	2.4	3.19E-07	3' UTR	*DAB2IP*
chr9	124049665	124050797	3.48	0.24	3.19E-07	Distal Intergenic	*LHX2*
chr9	121788315	121791710	5.68	2.29	3.65E-07	Distal Intergenic	*DAB2IP*
chr9	122851408	122853001	4.76	1.86	4.69E-07	Promoter (<=1kb)	*RC3H2*
chr9	121794690	121796407	4.39	1.05	4.93E-07	Distal Intergenic	*DAB2IP*
chr9	123418742	123421460	4.75	1.57	4.93E-07	Intron	*DENND1A*
chr9	121755009	121756646	3.82	0.55	5.05E-07	Intron	*DAB2IP*

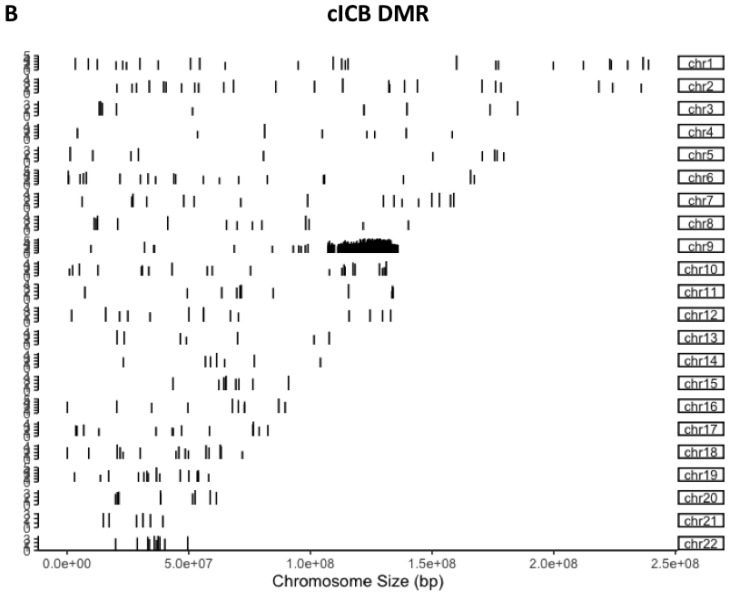
Appendix FF: Top 20 most significantly differentially methylated regions (DMRs) with immune checkpoint blockade (ICB)

Top 20 most significant DMR annotated with with pairwise C1 v C2 differential enrichment analysis in: A) regions enriched with sICB (10 paired samples); B) DMR enriched with cICB (14 paired samples; C) DMRs depleted with cICB (14 paired samples). DMRs were annotated with ChipSeeker. Conc_C1 and Conc_C2 denotes average read across conditions.

Distribution of DMR across the genome

.





Appendix GG: Distribution of DMRs across the genome

Genome coverage plot to show significant DMR identified with pairwise C1 versus C2 differential methylation analysis in A) patients receiving sICB (10 paired samples) and; B) patients receiving cICB (14 paired samples).

Samples used for integration of bulk RNA-seq expression profiles and epiegentic cfDNA data

ID	Cell_type	Cycle	RNA-seq_batch	Treatment	5mC_batch	Sex	Age
104	CD14	C1	One	cICB	Three	1	55
108	CD14	C1	Two	cICB	Three	1	63
122	CD14	C1	Two	cICB	Two	1	71
132	CD14	C1	Two	cICB	Three	1	48
137	CD14	C1	Two	sICB	One	1	95
138	CD14	C1	Two	cICB	One	2	56
153	CD14	C1	Three	cICB	One	2	58
157	CD14	C1	Three	sICB	One	2	72
158	CD14	C1	Three	cICB	One	1	68
160	CD14	C1	Three	sICB	One	2	96
174	CD14	C1	Three	sICB	Three	1	81
71	CD14	C1	One	sICB	Two	2	84
83	CD14	C1	One	cICB	Three	2	45
84	CD14	C1	One	sICB	Three	2	43
85	CD14	C1	One	cICB	Three	1	56
87	CD14	C1	One	cICB	Three	1	68
88	CD14	C1	One	sICB	Two	1	69
90	CD14	C1	One	cICB	Three	1	66
99	CD14	C1	One	cICB	Two	2	62
104	CD14	C2	One	cICB	Three	1	55
108	CD14	C2	Two	cICB	Three	1	63
122	CD14	C2	Two	cICB	Two	1	71
132	CD14	C2	Two	cICB	Three	1	48
137	CD14	C2	Two	sICB	One	1	95
138	CD14	C2	Two	cICB	One	2	56
153	CD14	C2	Three	cICB	One	2	58
157	CD14	C2	Three	sICB	One	2	72
158	CD14	C2	Three	cICB	One	1	68
160	CD14	C2	Three	sICB	Two	2	96
174	CD14	C2	Three	sICB	Three	1	81
71	CD14	C2	One	sICB	Two	2	84
81	CD14	C2	One	sICB	Three	2	72
83	CD14	C2	One	cICB	Three	2	45
84	CD14	C2	One	sICB	Three	2	43
85	CD14	C2	One	cICB	Three	1	56
87	CD14	C2	One	cICB	Three	1	68
88	CD14	C2	One	sICB	Two	1	69
90	CD14	C2	One	cICB	Three	1	66
99	CD14	C2	One	cICB	Two	2	62

Appendix HH: Cohort utilised for integration of monocyte expression and 5mC cfDNA profiles
Table to show samples used for integration analysis of monocyte expression profiles in bulk RNA-seq data with 5mC cfDNA profiles. ID indicates sample ID, CD14 infers monocytes, Cycle indicates ICB treatment cycle (C1 – untreated, C2 – treated), 'RNA-seq_batch' indicates sequencing batch (of three), Treatment infers ICB agent received (sICB – single agent ICB, cICB – combination ICB therapy), '5mC_batch' denotes 5mC cfDNA sequencing batch (of three). For Sex, 1 denotes male, 2 denotes female.

ID	Cell_type	Cycle	RNA-sq_batch	Treatment	5mC_batch	Sex	Age
104	CD8	C1	One	cICB	Three	1	55
108	CD8	C1	One	cICB	Three	1	63
122	CD8	C1	Two	cICB	Two	1	71
132	CD8	C1	Two	cICB	Three	1	48
137	CD8	C1	Two	sICB	One	1	95
138	CD8	C1	Two	cICB	One	2	56
157	CD8	C1	Three	sICB	One	2	72
158	CD8	C1	Three	cICB	One	1	68
160	CD8	C1	Three	sICB	One	2	96
174	CD8	C1	Three	sICB	Three	1	81
55	CD8	C1	One	cICB	Two	1	67
58	CD8	C1	One	cICB	Two	1	67
60	CD8	C1	One	sICB	One	1	60
62	CD8	C1	One	sICB	Two	2	67
71	CD8	C1	One	sICB	Two	2	84
83	CD8	C1	One	cICB	Three	2	45
84	CD8	C1	One	sICB	Three	2	43
85	CD8	C1	One	cICB	Three	1	56
87	CD8	C1	One	cICB	Three	1	68
88	CD8	C1	One	sICB	Two	1	69
90	CD8	C1	One	cICB	Three	1	66
99	CD8	C1	One	cICB	Two	2	62
104	CD8	C2	One	cICB	Three	1	55
108	CD8	C2	One	cICB	Three	1	63
122	CD8	C2	Two	cICB	Two	1	71
132	CD8	C2	Two	cICB	Three	1	48
137	CD8	C2	Two	sICB	One	1	95
138	CD8	C2	Two	cICB	One	2	56
157	CD8	C2	Three	sICB	One	2	72
158	CD8	C2	Three	cICB	One	1	68
160	CD8	C2	Three	sICB	One	2	96
174	CD8	C2	Three	sICB	Three	1	81
55	CD8	C2	One	cICB	Two	1	67
58	CD8	C2	One	cICB	Two	1	67
60	CD8	C2	One	sICB	One	1	60
62	CD8	C2	One	sICB	Two	2	67
71	CD8	C2	One	sICB	Two	2	84
81	CD8	C2	One	sICB	Three	2	72
83	CD8	C2	One	cICB	Three	2	45
84	CD8	C2	One	sICB	Three	2	43
85	CD8	C2	One	cICB	Three	1	56
87	CD8	C2	One	cICB	Three	1	68
88	CD8	C2	One	sICB	Two	1	69
90	CD8	C2	One	cICB	Three	1	66
99	CD8	C2	One	cICB	Two	2	62

Appendix II: Cohort utilised for integration of CD8$^+$ T cell expression and 5mC cfDNA profiles

Table to show samples used for integration analysis of monocyte expression profiles in bulk RNA-seq data with 5mC cfDNA profiles. ID indicates sample ID, CD14 infers monocytes, Cycle indicates ICB treatment cycle (C1 – untreated, C2 – treated), 'RNA-seq_batch' indicates sequencing batch (of three), Treatment infers ICB agent received (sICB – single agent ICB, cICB – combination ICB therapy), '5mC_batch' denotes 5mC cfDNA sequencing batch (of three). For Sex, 1 denotes male, 2 denotes female.

References

1. Melanoma skin cancer statistics | Cancer Research UK. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/melanoma-skin-cancer#heading-One.

2. Pak, B. J. *et al.* Radiation resistance of human melanoma analysed by retroviral insertional mutagenesis reveals a possible role for dopachrome tautomerase. *Oncogene* **23**, 30–38 (2004).

3. Shahbazian, D., Bindra, R. S., Kluger, H. M. & Glazer, P. M. Radiation sensitivity and sensitization in melanoma. *Pigment Cell Melanoma Res.* **26**, 928–930 (2013).

4. Luke, J. J., Flaherty, K. T., Ribas, A. & Long, G. V. Targeted agents and immunotherapies: Optimizing outcomes in melanoma. *Nature Reviews Clinical Oncology* vol. 14 463–482 (2017).

5. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).

6. Steuer, C. E. & Ramalingam, S. S. Tumor Mutation Burden: Leading Immunotherapy to the Era of Precision Medicine? *J. Clin. Oncol.* **36**, 631–632 (2018).

7. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science (80-.).* **348**, 69–74 (2015).

8. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252 (2012).

9. Passarelli, A., Mannavola, F., Stucci, L. S., Tucci, M. & Silvestris, F. Immune system and melanoma biology: a balance between immunosurveillance and immune escape. *Oncotarget* **8**, 106132 (2017).

10. He, X. & Xu, C. Immune checkpoint signaling and cancer immunotherapy. *Cell Research* vol. 30 660–669 (2020).

11. Brunet, J.-F. *et al.* A new member of the immunoglobulin superfamily—CTLA-4.

Nat. 1987 3286127 **328**, 267–270 (1987).

12. Walker, L. S. K. & Sansom, D. M. The emerging role of CTLA4 as a cell-extrinsic regulator of T cell responses. *Nature Reviews Immunology* vol. 11 852–863 (2011).

13. Teft, W. A., Kirchhof, M. G. & Madrenas, J. A molecular perspective of CTLA-4 function. *Annual Review of Immunology* vol. 24 65–97 (2006).

14. Collins, A. V. *et al.* The interaction properties of costimulatory molecules revisited. *Immunity* **17**, 201–210 (2002).

15. Robert, C. A decade of immune-checkpoint inhibitors in cancer therapy. *Nature Communications* vol. 11 1–3 (2020).

16. Hodi, F. S. *et al.* Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *N. Engl. J. Med.* **363**, 711–723 (2010).

17. Ascierto, P. A. *et al.* Overall survival (OS) and safety results from a phase 3 trial of ipilimumab (IPI) at 3 mg/kg vs 10 mg/kg in patients with metastatic melanoma (MEL). *Ann. Oncol.* **27**, vi379 (2016).

18. Ishida, Y., Agata, Y., Shibahara, K. & Honjo, T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **11**, 3887–3895 (1992).

19. Nishimura, H., Nose, M., Hiai, H., Minato, N. & Honjo, T. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity* **11**, 141–151 (1999).

20. 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.* **192**, 1027–1034 (2000).

21. 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.* **192**, 1027–1034 (2000).

22. Fares, J., Fares, M. Y. & Fares, Y. Immune checkpoint inhibitors: Advances and

impact in neuro-oncology. *Surgical Neurology International* vol. 10 (2019).

23. Topalian, S. L. *et al.* Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454 (2012).

24. Hamid, O. *et al.* Safety and Tumor Responses with Lambrolizumab (Anti–PD-1) in Melanoma. *N. Engl. J. Med.* **369**, 134–144 (2013).

25. Robert, C. *et al.* Nivolumab in Previously Untreated Melanoma without BRAF Mutation . *N. Engl. J. Med.* **372**, 320–330 (2015).

26. Hamid, O. *et al.* Five-year survival outcomes for patients with advanced melanoma treated with pembrolizumab in KEYNOTE-001. *Ann. Oncol.* **30**, 582–588 (2019).

27. Garon, E. B. *et al.* Pembrolizumab for the Treatment of Non–Small-Cell Lung Cancer. *http://dx.doi.org/10.1056/NEJMoa1501824* **372**, 2018–2028 (2015).

28. Motzer, R. J. *et al.* Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. *N. Engl. J. Med.* **378**, 1277–1290 (2018).

29. Powles, T. *et al.* Avelumab Maintenance Therapy for Advanced or Metastatic Urothelial Carcinoma. *https://doi.org/10.1056/NEJMoa2002788* **383**, 1218–1230 (2020).

30. Ferris, R. L. *et al.* Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *https://doi.org/10.1056/NEJMoa1602252* **375**, 1856–1867 (2016).

31. Wolchok, J. D. *et al.* Nivolumab plus Ipilimumab in advanced melanoma. *N. Engl. J. Med.* **369**, 122–133 (2013).

32. Postow, M. A. *et al.* Nivolumab and ipilimumab versus ipilimumab in untreated melanoma. *N. Engl. J. Med.* **372**, 2006–2017 (2015).

33. Hodi, F. S. *et al.* Nivolumab plus ipilimumab or nivolumab alone versus ipilimumab alone in advanced melanoma (CheckMate 067): 4-year outcomes of a multicentre, randomised, phase 3 trial. *Lancet Oncol.* **19**, 1480–1492 (2018).

34. Larkin, J. *et al.* Five-Year Survival with Combined Nivolumab and Ipilimumab in

Advanced Melanoma. *N. Engl. J. Med.* **381**, 1535–1546 (2019).

35. Braun, D. A. *et al.* Beyond conventional immune-checkpoint inhibition — novel immunotherapies for renal cell carcinoma. *Nature Reviews Clinical Oncology* vol. 18 199–214 (2021).

36. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N. Engl. J. Med.* **372**, 2509–2520 (2015).

37. Darvin, P., Toor, S. M., Sasidharan Nair, V. & Elkord, E. Immune checkpoint inhibitors: recent progress and potential biomarkers. *Experimental and Molecular Medicine* vol. 50 (2018).

38. Boyiadzis, M. M. *et al.* Significance and implications of FDA approval of pembrolizumab for biomarker-defined disease. *Journal for ImmunoTherapy of Cancer* vol. 6 35 (2018).

39. Rausch, M. P. & Hastings, K. T. Immune Checkpoint Inhibitors in the Treatment of Melanoma: From Basic Science to Clinical Application. in *Cutaneous Melanoma: Etiology and Therapy* 121–142 (Codon Publications, 2017). doi:10.15586/codon.cutaneousmelanoma.2017.ch9.

40. Ye, W. *et al.* Checkpoint-blocker-induced autoimmunity is associated with favourable outcome in metastatic melanoma and distinct T-cell expression profiles. *Br. J. Cancer* 1–9 (2021) doi:10.1038/s41416-021-01310-3.

41. Attia, P. *et al.* Autoimmunity correlates with tumor regression in patients with metastatic melanoma treated with anti-cytotoxic T-lymphocyte antigen-4. *J. Clin. Oncol.* **23**, 6043–6053 (2005).

42. Downey, S. G. *et al.* Prognostic factors related to clinical response in patients with metastatic melanoma treated by CTL-associated antigen-4 blockade. *Clin. Cancer Res.* **13**, 6681–6688 (2007).

43. Weber, J. S. *et al.* Safety profile of nivolumab monotherapy: A pooled analysis of patients with advanced melanoma. *J. Clin. Oncol.* **35**, 785–792 (2017).

44. Liu, D. *et al.* Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat. Med.* **25**, 1916–1927

(2019).

45. Luke, J. J. & Ascierto, P. A. Biology confirmed but biomarkers elusive in melanoma immunotherapy. *Nature Reviews Clinical Oncology* 1–2 (2020) doi:10.1038/s41571-020-0328-8.

46. Thomas, N. E. *et al.* Tumor-infiltrating lymphocyte grade in primary melanomas is independently associated with melanoma-specific survival in the population-based genes, environment and melanoma study. *J. Clin. Oncol.* **31**, 4252–4259 (2013).

47. Idos, G. E. *et al.* The Prognostic Implications of Tumor Infiltrating Lymphocytes in Colorectal Cancer: A Systematic Review and Meta-Analysis. *Sci. Rep.* **10**, 3360 (2020).

48. Chen, B. *et al.* Prognostic value of the common tumour-infiltrating lymphocyte subtypes for patients with non-small cell lung cancer: A meta-analysis. *PLoS One* **15**, e0242173 (2020).

49. *Standards and datasets for reporting cancers Dataset for histopathological reporting of primary cutaneous malignant melanoma and regional lymph nodes Unique document number G125 Document name Dataset for histopathological reporting of primary cutaneous malignant melanoma and regional lymph nodes*. (2019).

50. Tumeh, P. C. *et al.* PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature* **515**, 568–571 (2014).

51. Havel, J. J., Chowell, D. & Chan, T. A. The evolving landscape of biomarkers for checkpoint inhibitor immunotherapy. *Nature Reviews Cancer* vol. 19 133–150 (2019).

52. Fridman, W. H., Pagès, F., Saụ̀ts-Fridman, C. & Galon, J. The immune contexture in human tumours: Impact on clinical outcome. *Nature Reviews Cancer* vol. 12 298–306 (2012).

53. Wu, Y., Chen, W., Xu, Z. P. & Gu, W. PD-L1 distribution and perspective for cancer immunotherapy— blockade, knockdown, or inhibition. *Frontiers in*

Immunology vol. 10 (2019).

54. Hino, R. *et al.* Tumor cell expression of programmed cell death-1 ligand 1 is a prognostic factor for malignant melanoma. *Cancer* **116**, 1757–1766 (2010).

55. Davis, A. A. & Patel, V. G. The role of PD-L1 expression as a predictive biomarker: An analysis of all US food and drug administration (FDA) approvals of immune checkpoint inhibitors. *J. Immunother. Cancer* **7**, 1–8 (2019).

56. Zhang, B. *et al.* Predictive effect of PD-L1 expression for immune checkpoint inhibitor (PD-1/PD-L1 inhibitors) treatment for non-small cell lung cancer: A meta-analysis. *Int. Immunopharmacol.* **80**, (2020).

57. Pai-Scherf, L. *et al.* FDA Approval Summary: Pembrolizumab for Treatment of Metastatic Non-Small Cell Lung Cancer: First-Line Therapy and Beyond. *Oncologist* **22**, 1392–1399 (2017).

58. 1 Recommendation | Pembrolizumab for untreated PD-L1-positive metastatic non-small-cell lung cancer | Guidance | NICE.

59. Gibney, G. T., Weiner, L. M. & Atkins, M. B. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *The Lancet Oncology* vol. 17 e542–e551 (2016).

60. Kim, H., Kwon, H. J., Park, S. Y., Park, E. & Chung, J. H. PD-L1 immunohistochemical assays for assessment of therapeutic strategies involving immune checkpoint inhibitors in non-small cell lung cancer: A comparative study. *Oncotarget* **8**, 98524–98532 (2017).

61. Mansfield, A. S. *et al.* Heterogeneity of Programmed Cell Death Ligand 1 Expression in Multifocal Lung Cancer. *Clin. Cancer Res.* **22**, 2177–2182 (2016).

62. Wang, M., Wang, S., Trapani, J. A. & Neeson, P. J. Challenges of PD-L1 testing in non-small cell lung cancer and beyond. *J. Thorac. Dis.* **12**, 4541–4548 (2020).

63. Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med.* **9**, 1–14 (2017).

64. Hellmann, M. D. *et al.* Nivolumab plus Ipilimumab in Lung Cancer with a High Tumor Mutational Burden. *N. Engl. J. Med.* **378**, 2093–2104 (2018).

65. Rosenberg, J. E. *et al.* Atezolizumab in patients with locally advanced and metastatic urothelial carcinoma who have progressed following treatment with platinum-based chemotherapy: A single-arm, multicentre, phase 2 trial. *Lancet* **387**, 1909–1920 (2016).

66. Hellmann, M. D. *et al.* Tumor Mutational Burden and Efficacy of Nivolumab Monotherapy and in Combination with Ipilimumab in Small-Cell Lung Cancer. *Cancer Cell* **33**, 853-861.e4 (2018).

67. Bai, R., Lv, Z., Xu, D. & Cui, J. Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors. *Biomarker Research* vol. 8 34 (2020).

68. André, T. *et al.* Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *N. Engl. J. Med.* **383**, 2207–2218 (2020).

69. Marabelle, A. *et al.* Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/ mismatch repair–deficient cancer: Results from the phase II KEYNOTE-158 study. *J. Clin. Oncol.* **38**, 1–10 (2020).

70. Bedard, P. L., Hansen, A. R., Ratain, M. J. & Siu, L. L. Tumour heterogeneity in the clinic. *Nature* vol. 501 355–364 (2013).

71. Fairfax, B. P. *et al.* Peripheral CD8+ T cell characteristics associated with durable responses to immune checkpoint blockade in patients with metastatic melanoma. *Nature Medicine* vol. 26 193–199 (2020).

72. Strauss, L. *et al. Targeted deletion of PD-1 in myeloid cells induces antitumor immunity*. *Sci. Immunol* vol. 5 http://immunology.sciencemag.org/ (2020).

73. Krieg, C. *et al.* High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nat. Med.* **24**, 144–153 (2018).

74. Ginhoux, F. & Jung, S. Monocytes and macrophages: developmental pathways and tissue homeostasis. *Nat. Rev. Immunol. 2014 146* **14**, 392–404 (2014).

75. Ginhoux, F. & Jung, S. Monocytes and macrophages: Developmental pathways and tissue homeostasis. *Nature Reviews Immunology* vol. 14 392–404 (2014).

76. Ma, W. T., Gao, F., Gu, K. & Chen, D. K. The role of monocytes and macrophages

in autoimmune diseases: A comprehensive review. *Frontiers in Immunology* vol. 10 1140 (2019).

77. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat. Immunol. 2010 115* **11**, 373–384 (2010).

78. Ahn, J. & Barber, G. N. STING signaling and host defense against microbial infection. *Exp. Mol. Med. 2019 5112* **51**, 1–10 (2019).

79. Jakubzick, C. V., Randolph, G. J. & Henson, P. M. Monocyte differentiation and antigen-presenting functions. *Nat. Rev. Immunol. 2017 176* **17**, 349–362 (2017).

80. Lubbers, R., Essen, M. F. van, Kooten, C. van & Trouw, L. A. Production of complement components by cells of the immune system. *Clin. Exp. Immunol.* **188**, 183 (2017).

81. Dale, D. C., Boxer, L. & Conrad Liles, W. The phagocytes: Neutrophils and monocytes. *Blood* **112**, 935–945 (2008).

82. Mkaddem, S. Ben, Benhamou, M. & Monteiro, R. C. Understanding Fc receptor involvement in inflammatory diseases: From mechanisms to new therapeutic tools. *Front. Immunol.* **10**, 811 (2019).

83. Bournazos, S., Gupta, A. & Ravetch, J. V. The role of IgG Fc receptors in antibody-dependent enhancement. *Nat. Rev. Immunol. 2020 2010* **20**, 633–643 (2020).

84. Vogelpoel, L. T. C., Baeten, D. L. P., de Jong, E. C. & den Dunnen, J. Control of cytokine production by human Fc gamma receptors: Implications for pathogen defense and autoimmunity. *Front. Immunol.* **6**, 79 (2015).

85. Ziegler-Heitbrock, L. *et al.* Nomenclature of monocytes and dendritic cells in blood. (2010) doi:10.1182/blood-2010-02-258558.

86. Belge, K.-U. *et al.* The Proinflammatory CD14 + CD16 + DR ++ Monocytes Are a Major Source of TNF . *J. Immunol.* **168**, 3536–3542 (2002).

87. Szaflarska, A. *et al.* Antitumor response of CD14 +/CD16 + monocyte

subpopulation. *Exp. Hematol.* **32**, 748–755 (2004).

88. Ziegler-Heitbrock, L. *et al.* Nomenclature of monocytes and dendritic cells in blood. *Blood* vol. 116 (2010).

89. Boyette, L. B. *et al.* Phenotype, function, and differentiation potential of human monocyte subsets. *PLoS One* **12**, e0176460 (2017).

90. Wong, K. L. *et al.* Gene expression profiling reveals the defining features of the classical, intermediate,and nonclassical human monocyte subsets. *Blood* **118**, e16–e31 (2011).

91. Kapellos, T. S. *et al.* Human monocyte subsets and phenotypes in major chronic inflammatory diseases. *Frontiers in Immunology* vol. 10 2035 (2019).

92. Weber, C. *et al.* Differential chemokine receptor expression and function in human monocyte subpopulations. *J. Leukoc. Biol.* **67**, 699–704 (2000).

93. Wong, K. L. *et al.* Gene expression profiling reveals the defining features of the classical, intermediate,and nonclassical human monocyte subsets. *Blood* **118**, e16–e31 (2011).

94. Xia, C., Braunstein, Z., Toomey, A. C., Zhong, J. & Rao, X. S100 proteins as an important regulator of macrophage inflammation. *Frontiers in Immunology* vol. 8 (2018).

95. Schmidl, C. *et al.* Transcription and enhancer profiling in human monocyte subsets. *Blood* **123**, e90–e99 (2014).

96. Kanamori-Katayama, M. *et al.* Unamplified cap analysis of gene expression on a single-molecule sequencer. *Genome Res.* **21**, 1150–1159 (2011).

97. Ravenhill, B. J., Soday, L., Houghton, J., Antrobus, R. & Weekes, M. P. Comprehensive cell surface proteomics defines markers of classical, intermediate and non-classical monocytes. *bioRxiv* 2020.02.20.958009 (2020) doi:10.1101/2020.02.20.958009.

98. Boyette, L. B. *et al.* Phenotype, function, and differentiation potential of human monocyte subsets. *PLoS One* **12**, (2017).

99. Zawada, A. M. *et al.* SuperSAGE evidence for CD14 ++CD16 + monocytes as a third monocyte subset. *Blood* **118**, e50–e61 (2011).

100. Gren, S. T. *et al.* A Single-Cell Gene-Expression Profile Reveals Inter-Cellular Heterogeneity within Human Monocyte Subsets. *PLoS One* **10**, e0144351 (2015).

101. Ingersoll, M. A. *et al.* Comparison of gene expression profiles between human and mouse monocyte subsets. *Blood* **115**, e10 (2010).

102. Cros, J. *et al.* Human CD14dim Monocytes Patrol and Sense Nucleic Acids and Viruses via TLR7 and TLR8 Receptors. *Immunity* **33**, 375–386 (2010).

103. Guilliams, M., Mildner, A. & Yona, S. Developmental and Functional Heterogeneity of Monocytes. *Immunity* vol. 49 595–613 (2018).

104. Villani, A. C. *et al.* Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science (80-.).* **356**, (2017).

105. Chimen, M. *et al.* Monocyte Subsets Coregulate Inflammatory Responses by Integrated Signaling through TNF and IL-6 at the Endothelial Cell Interface. *J. Immunol.* **198**, 2834–2843 (2017).

106. Ong, S. M. *et al.* The pro-inflammatory phenotype of the human non-classical monocyte subset is attributed to senescence article. *Cell Death Dis.* **9**, (2018).

107. Ancuta, P. *et al.* Transcriptional profiling reveals developmental relationship and distinct biological functions of CD16+ and CD16- monocyte subsets. *BMC Genomics* **10**, 403 (2009).

108. Carlin, L. M. *et al.* Nr4a1-dependent Ly6Clow monocytes monitor endothelial cells and orchestrate their disposal. *Cell* **153**, 362–375 (2013).

109. Yeap, W. H. *et al.* CD16 is indispensable for antibodydependent cellular cytotoxicity by human monocytes. *Sci. Rep.* **6**, 1–22 (2016).

110. Olikowsky, T. *et al.* Two distinct pathways of human macrophage differentiation are mediated by interferon-γ and interleukin-10. *Immunology* **91**, 104–108 (1997).

111. Welch, G. R., Wong, H. L. & Wahl, S. M. Selective induction of Fc gamma RIII on

human monocytes by transforming growth factor-beta. *J. Immunol.* **144**, (1990).

112. Schmid, I. *et al.* Alterations in phenotype and cell-surface antigen expression levels of human monocytes: Differential response to in vivo administration of rhm-csf or rhgm-csf. *Cytometry* **22**, 103–110 (1995).

113. Pugin, J. *et al.* CD14 Is a pattern recognition receptor. *Immunity* **1**, 509–516 (1994).

114. Tsou, C. L. *et al.* Critical roles for CCR2 and MCP-3 in monocyte mobilization from bone marrow and recruitment to inflammatory sites. *J. Clin. Invest.* **117**, 902–909 (2007).

115. SELL Gene - GeneCards | LYAM1 Protein | LYAM1 Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=SELL.

116. CD36 Gene - GeneCards | CD36 Protein | CD36 Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=CD36.

117. HLA-DRA - HLA class II histocompatibility antigen, DR alpha chain precursor - Homo sapiens (Human) - HLA-DRA gene & protein. https://www.uniprot.org/uniprot/P01903.

118. CDKN1C Gene - GeneCards | CDN1C Protein | CDN1C Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=CDKN1C.

119. Liu, K. *et al.* MTSS1 inhibits metastatic potential and induces G2/M phase cell cycle arrest in gastric cancer. *Onco. Targets. Ther.* **12**, 5143 (2019).

120. Thomas, G., Tacke, R., Hedrick, C. C. & Hanna, R. N. Nonclassical Patrolling Monocyte Function in the Vasculature. *Arterioscler. Thromb. Vasc. Biol.* **35**, 1306–1316 (2015).

121. Mortensen, S. A. *et al.* Structure and activation of C1, the complex initiating the classical pathway of the complement cascade. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 986–991 (2017).

122. Feeley, E. M. *et al.* IFITM3 Inhibits Influenza A Virus Infection by Preventing Cytosolic Entry. *PLOS Pathog.* **7**, e1002337 (2011).

123. Gadalla, R. *et al.* Validation of CyTOF Against Flow Cytometry for Immunological Studies and Monitoring of Human Cancer Clinical Trials. *Front. Oncol.* **9**, 415 (2019).

124. Spitzer, M. H. & Nolan, G. P. Mass Cytometry: Single Cells, Many Features. *Cell* vol. 165 780–791 (2016).

125. Roussel, M. *et al.* Mass cytometry deep phenotyping of human mononuclear phagocytes and myeloid-derived suppressor cells from human blood and bone marrow. *J. Leukoc. Biol.* **102**, 437–447 (2017).

126. Ihanus, E., Uotila, L. M., Toivanen, A., Varis, M. & Gahmberg, C. G. Red-cell ICAM-4 is a ligand for the monocyte/macrophage integrin CD11c/CD18: characterization of the binding sites on ICAM-4. *Blood* **109**, 802–810 (2007).

127. Keizer, G. D., Te Velde, A. A., Schwarting, R., Figdor, C. G. & De Vries, J. E. Role of p150,95 in adhesion, migration, chemotaxis and phagocytosis of human monocytes. *Eur. J. Immunol.* **17**, 1317–1322 (1987).

128. Hofer, T. P., van de Loosdrecht, A. A., Stahl-Hennig, C., Cassatella, M. A. & Ziegler-Heitbrock, L. 6-Sulfo LacNAc (Slan) as a Marker for Non-classical Monocytes. *Frontiers in Immunology* vol. 10 (2019).

129. Xia, L. *et al.* P-selectin glycoprotein ligand-1–deficient mice have impaired leukocyte tethering to E-selectin under flow. *J. Clin. Invest.* **109**, 939–950 (2002).

130. Ong, S.-M. *et al.* A Novel, Five-Marker Alternative to CD16–CD14 Gating to Identify the Three Human Monocyte Subsets. *Front. Immunol.* **10**, 1761 (2019).

131. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc. 2013 91* **9**, 171–181 (2014).

132. Günther, P. *et al.* A rule-based data-informed cellular consensus map of the human mononuclear phagocyte cell space. doi:10.1101/658179.

133. Cormican, S. & Griffin, M. D. Human Monocyte Subset Distinctions and Function: Insights From Gene Expression Analysis. *Front. Immunol.* **0**, 1070 (2020).

134. Zilionis, R. *et al.* Single-Cell Transcriptomics of Human and Mouse Lung Cancers

Reveals Conserved Myeloid Populations across Individuals and Species. *Immunity* **50**, 1317-1334.e10 (2019).

135. Vishwakarma, A. *et al.* Title: Mapping the Immune Landscape of Clear Cell Renal Cell Carcinoma by Single-Cell RNA-seq. doi:10.1101/824482.

136. Goudot, C. *et al.* Aryl Hydrocarbon Receptor Controls Monocyte Differentiation into Dendritic Cells versus Macrophages. *Immunity* **47**, 582-596.e6 (2017).

137. Ghaedi, M. *et al.* Common-Lymphoid-Progenitor-Independent Pathways of Innate and T Lymphocyte Development. *Cell Rep.* **15**, 471–480 (2016).

138. Akashi, K., Traver, D., Miyamoto, T. & Weissman, I. L. A clonogenic common myeloid progenitor that gives rise to all myeloid lineages. *Nature* **404**, 193–197 (2000).

139. Lee, J. *et al.* Restricted dendritic cell and monocyte progenitors in human cord blood and bone marrow. *J. Exp. Med.* **212**, 385–399 (2015).

140. Kawamura, S. *et al.* Identification of a Human Clonogenic Progenitor with Strict Monocyte Differentiation Potential: A Counterpart of Mouse cMoPs. *Immunity* **46**, 835-848.e4 (2017).

141. Serbina, N. V. & Pamer, E. G. Monocyte emigration from bone marrow during bacterial infection requires signals mediated by chemokine receptor CCR2. *Nat. Immunol.* **7**, 311–317 (2006).

142. Tsou, C. L. *et al.* Critical roles for CCR2 and MCP-3 in monocyte mobilization from bone marrow and recruitment to inflammatory sites. *J. Clin. Invest.* **117**, 902–909 (2007).

143. Bennett, F. C. *et al.* A Combination of Ontogeny and CNS Environment Establishes Microglial Identity. *Neuron* **98**, 1170-1183.e8 (2018).

144. Jakubzick, C. *et al.* Minimal differentiation of classical monocytes as they survey steady-state tissues and transport antigen to lymph nodes. *Immunity* **39**, 599–610 (2013).

145. Davies, L. C., Jenkins, S. J., Allen, J. E. & Taylor, P. R. Tissue-resident

macrophages. *Nature Immunology* vol. 14 986–995 (2013).

146. Bain, C. C. *et al.* Constant replenishment from circulating monocytes maintains the macrophage pool in the intestine of adult mice. *Nat. Immunol.* **15**, 929–937 (2014).

147. Epelman, S. *et al.* Embryonic and adult-derived resident cardiac macrophages are maintained through distinct mechanisms at steady state and during inflammation. *Immunity* **40**, 91–104 (2014).

148. Tamoutounour, S. *et al.* Origins and functional specialization of macrophages and of conventional and monocyte-derived dendritic cells in mouse skin. *Immunity* **39**, 925–938 (2013).

149. Patel, A. A. *et al.* The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. *J. Exp. Med.* **214**, 1913–1923 (2017).

150. Korkosz, M., Bukowska-Strakova, K., Sadis, S., Grodzicki, T. & Siedlar, M. Monoclonal antibodies against macrophage colony-stimulating factor diminish the number of circulating intermediate and nonclassical (CD14 ++CD16 +/CD14 +CD16 ++) monocytes in rheumatoid arthritis patient. *Blood* vol. 119 5329–5330 (2012).

151. Selimoglu-Buet, D. *et al.* A miR-150/TET3 pathway regulates the generation of mouse and human non-classical monocyte subset. *Nat. Commun.* **9**, 1–17 (2018).

152. Rojo, R. *et al.* Deletion of a Csf1r enhancer selectively impacts CSF1R expression and development of tissue macrophage populations. *Nat. Commun.* **10**, 14 (2019).

153. Mildner, A. *et al.* Genomic Characterization of Murine Monocytes Reveals C/EBPβ Transcription Factor Dependence of Ly6C− Cells. *Immunity* **46**, 849-862.e7 (2017).

154. Laurenti, E. & Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. *Nature* vol. 553 418–426 (2018).

155. Kiss, M., Caro, A. A., Raes, G. & Laoui, D. Systemic Reprogramming of Monocytes

in Cancer. *Frontiers in Oncology* vol. 10 1399 (2020).

156. Lawrence, T. & Natoli, G. Transcriptional regulation of macrophage polarization: Enabling diversity with identity. *Nature Reviews Immunology* vol. 11 750–761 (2011).

157. Lin, Y., Xu, J. & Lan, H. Tumor-associated macrophages in tumor metastasis: Biological roles and clinical therapeutic applications. *Journal of Hematology and Oncology* vol. 12 1–16 (2019).

158. Biswas, S. K. *et al.* A distinct and unique transcriptional program expressed by tumor-associated macrophages (defective NF-κB and enhanced IRF-3/STAT1 activation). *Blood* **107**, 2112–2122 (2006).

159. Ojalvo, L. S., King, W., Cox, D. & Pollard, J. W. High-density gene expression analysis of tumor-associated macrophages from mouse mammary tumors. *Am. J. Pathol.* **174**, 1048–1064 (2009).

160. Mantovani, A., Marchesi, F., Malesci, A., Laghi, L. & Allavena, P. Tumour-associated macrophages as treatment targets in oncology. *Nature Reviews Clinical Oncology* vol. 14 399–416 (2017).

161. Ostuni, R., Kratochvill, F., Murray, P. J. & Natoli, G. Macrophages and cancer: From mechanisms to therapeutic implications. *Trends in Immunology* vol. 36 229–239 (2015).

162. Kiss, M., Caro, A. A., Raes, G. & Laoui, D. Systemic Reprogramming of Monocytes in Cancer. *Frontiers in Oncology* vol. 10 1399 (2020).

163. Gebhardt, C. *et al.* Myeloid cells and related chronic inflammatory factors as novel predictive markers in melanoma treatment with ipilimumab. *Clin. Cancer Res.* **21**, 5453–5459 (2015).

164. Wen, J. *et al.* Prognostic significance of preoperative circulating monocyte count in patients with breast cancer based on a large cohort study. *Med. (United States)* **94**, (2015).

165. Wen, S. *et al.* Peripheral monocyte counts predict the clinical outcome for patients with colorectal cancer: a systematic review and meta-analysis. *Eur. J.*

Gastroenterol. Hepatol. **31**, 1313–1321 (2019).

166. Iacono, D. *et al.* Prognostic role of disease extent and lymphocyte–monocyte ratio in advanced melanoma. *Melanoma Res.* **29**, 510–515 (2019).

167. Li, K. J. *et al.* Predictive value of lymphocyte-to-monocyte ratio (LMR) and neutrophil-to-lymphocyte ratio (NLR) in patients with oesophageal cancer undergoing concurrent chemoradiotherapy. *BMC Cancer* **19**, 1–9 (2019).

168. Lee, S. F. & Luque-Fernandez, M. A. Prognostic value of lymphocyte-To-monocyte ratio and neutrophil-To-lymphocyte ratio in follicular lymphoma: A retrospective cohort study. *BMJ Open* **7**, e017904 (2017).

169. Chan, J. C. Y. *et al.* The lymphocyte-to-monocyte ratio is a superior predictor of overall survival in comparison to established biomarkers of resectable colorectal cancer. *Ann. Surg.* **265**, 539–546 (2017).

170. Li, W., Tao, L., Zhang, L. & Xiu, D. Prognostic role of lymphocyte to monocyte ratio for patients with pancreatic cancer: A systematic review and meta-analysis. *OncoTargets and Therapy* vol. 10 3391–3397 (2017).

171. Volodko, N., Reiner, A., Rudas, M. & Jakesz, R. Tumour-associated macrophages in breast cancer and their prognostic correlations. *Breast* **7**, 99–105 (1998).

172. Koukourakis, M. I. *et al.* Different patterns of stromal and cancer cell thymidine phosphorylase reactivity in non small-cell lung cancer: Impact on tumour neoangiogenesis and survival. *Br. J. Cancer* **77**, 1696–1703 (1998).

173. Hanada, T. *et al.* Prognostic value of tumor-associated macrophage count in human bladder cancer. *Int. J. Urol.* **7**, 263–269 (2000).

174. Bingle, L., Brown, N. J. & Lewis, C. E. The role of tumour-associated macrophages in tumour progression: implications for new anticancer therapies. *J. Pathol.* **196**, 254–265 (2002).

175. Wu, W. C. *et al.* Circulating hematopoietic stem and progenitor cells are myeloid-biased in cancer patients. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4221–4226 (2014).

176. Cassetta, L. *et al.* Human Tumor-Associated Macrophage and Monocyte Transcriptional Landscapes Reveal Cancer-Specific Reprogramming, Biomarkers, and Therapeutic Targets. *Cancer Cell* **35**, 588-602.e10 (2019).

177. Ramos, R. N. *et al.* CD163 $^{+}$ tumor-associated macrophage accumulation in breast cancer patients reflects both local differentiation signals and systemic skewing of monocytes. *Clin. Transl. Immunol.* **9**, (2020).

178. Qian, B. Z. *et al.* CCL2 recruits inflammatory monocytes to facilitate breast-tumour metastasis. *Nature* **475**, 222–225 (2011).

179. Ugurel, S. *et al.* Down-regulation of HLA II and costimulatory CD86/B7-2 on circulating monocytes from melanoma patients. *Cancer Immunol. Immunother.* **53**, 551–559 (2004).

180. Gordon, I. O. & Freedman, R. S. Defective antitumor function of monocyte-derived macrophages from epithelial ovarian cancer patients. *Clin. Cancer Res.* **12**, 1515–1524 (2006).

181. Almand, B. *et al.* Increased Production of Immature Myeloid Cells in Cancer Patients: A Mechanism of Immunosuppression in Cancer. *J. Immunol.* **166**, 678–689 (2001).

182. Trivedi, R. & Mishra, D. P. Trailing TRAIL Resistance: Novel Targets for TRAIL Sensitization in Cancer Cells. *Front. Oncol.* **5**, (2015).

183. Jaiswal, S. *et al.* CD47 Is Upregulated on Circulating Hematopoietic Stem Cells and Leukemia Cells to Avoid Phagocytosis. *Cell* **138**, 271–285 (2009).

184. Veglia, F., Perego, M. & Gabrilovich, D. Myeloid-derived suppressor cells coming of age review-article. *Nature Immunology* vol. 19 108–119 (2018).

185. Bronte, V. *et al.* Recommendations for myeloid-derived suppressor cell nomenclature and characterization standards. *Nature Communications* vol. 7 (2016).

186. Mengos, A. E., Gastineau, D. A. & Gustafson, M. P. The CD14+HLA-DrlO/NEG monocyte: An immunosuppressive phenotype that restrains responses to cancer immunotherapy. *Frontiers in Immunology* vol. 10 (2019).

187. Gustafson, M. P. *et al.* Intratumoral CD14+ cells and circulating CD14+HLA-DRlo/neg monocytes correlate with decreased survival in patients with clear cell renal cell carcinoma. *Clin. Cancer Res.* **21**, 4224–4233 (2015).

188. Gonda, K. *et al.* Myeloid-derived suppressor cells are increased and correlated with type 2 immune responses, malnutrition, inflammation, and poor prognosis in patients with breast cancer. *Oncol. Lett.* **14**, 1766–1774 (2017).

189. Sun, H. L. *et al.* Increased frequency and clinical significance of myeloidderived suppressor cells in human colorectal carcinoma. *World J. Gastroenterol.* **18**, 3303–3309 (2012).

190. Zhang, S. *et al.* The role of myeloid-derived suppressor cells in patients with solid tumors: A meta-analysis. *PLoS One* **11**, (2016).

191. Meyer, C. *et al.* Frequencies of circulating MDSC correlate with clinical outcome of melanoma patients treated with ipilimumab. *Cancer Immunol. Immunother.* **63**, 247–257 (2014).

192. Chavan, R. *et al.* Untreated stage IV Melanoma patients exhibit abnormal monocyte phenotypes and decreased functional capacity. in *Cancer Immunology Research* vol. 2 241–248 (American Association for Cancer Research Inc., 2014).

193. Hamm, A. *et al.* Tumour-educated circulating monocytes are powerful candidate biomarkers for diagnosis and disease follow-up of colorectal cancer. *Gut* **65**, 990–1000 (2016).

194. Sun, H. L. *et al.* Increased frequency and clinical significance of myeloidderived suppressor cells in human colorectal carcinoma. *World J. Gastroenterol.* **18**, 3303–3309 (2012).

195. Peranzoni, E., Ingangi, V., Masetto, E., Pinton, L. & Marigo, I. Myeloid Cells as Clinical Biomarkers for Immune Checkpoint Blockade. *Frontiers in Immunology* vol. 11 1590 (2020).

196. Chasseuil, E. *et al.* Blood predictive biomarkers for nivolumab in advanced melanoma. *Acta Derm. Venereol.* **98**, 406–410 (2018).

197. Soyano, A. E. *et al.* Peripheral blood biomarkers correlate with outcomes in

advanced non-small cell lung Cancer patients treated with anti-PD-1 antibodies. *J. Immunother. Cancer* **6**, 129 (2018).

198. Strauss, L. *et al.* Targeted deletion of PD-1 in myeloid cells induces antitumor immunity. *Sci. Immunol.* **5**, (2020).

199. Weber, J. *et al.* Phase I/II study of metastatic melanoma patients treated with nivolumab who had progressed after ipilimumab. *Cancer Immunol. Res.* **4**, 345–353 (2016).

200. Gibney, G. T. *et al.* Safety, correlative markers, and clinical results of adjuvant nivolumab in combination with vaccine in resected high-risk metastatic melanoma. *Clin. Cancer Res.* **21**, 712–720 (2015).

201. Retseck, J. *et al.* Long term impact of CTLA4 blockade immunotherapy on regulatory and effector immune responses in patients with melanoma. *J. Transl. Med.* **16**, 184 (2018).

202. de Coaña, Y. P. *et al.* Ipilimumab treatment decreases monocytic MDSCs and increases CD8 effector memory T cells in long-term survivors with advanced melanoma. *Oncotarget* **8**, 21539–21553 (2017).

203. Romano, E. *et al.* Ipilimumab-dependent cell-mediated cytotoxicity of regulatory T cells ex vivo by nonclassical monocytes in melanoma patients. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 6140–6145 (2015).

204. Gubin, M. M. *et al.* High-Dimensional Analysis Delineates Myeloid and Lymphoid Compartment Remodeling during Successful Immune-Checkpoint Cancer Therapy. *Cell* **175**, 1014-1030.e19 (2018).

205. Gubin, M. M. *et al.* High-Dimensional Analysis Delineates Myeloid and Lymphoid Compartment Remodeling during Successful Immune-Checkpoint Cancer Therapy. *Cell* **175**, 1014-1030.e19 (2018).

206. Cabel, L. *et al.* Clinical potential of circulating tumour DNA in patients receiving anticancer immunotherapy. *Nature Reviews Clinical Oncology* vol. 15 639–650 (2018).

207. Ma, X. *et al.* Cell-free DNA provides a good representation of the tumor genome

despite its biased fragmentation patterns. *PLoS One* **12**, (2017).

208. Moss, J. *et al.* Comprehensive human cell-type methylation atlas reveals origins of circulating cell-free DNA in health and disease. *Nat. Commun.* **9**, 1–12 (2018).

209. Mouliere, F. *et al.* Enhanced detection of circulating tumor DNA by fragment size analysis. *Sci. Transl. Med.* **10**, (2018).

210. Alcaide, M. *et al.* Evaluating the quantity, quality and size distribution of cell-free DNA by multiplex droplet digital PCR. *Sci. Reports 2020 101* **10**, 1–10 (2020).

211. Henikoff, S. & Church, G. M. Simultaneous Discovery of Cell-Free DNA and the Nucleosome Ladder. *Genetics* **209**, 27–29 (2018).

212. Kustanovich, A., Schwartz, R., Peretz, T. & Grinshpun, A. Life and death of circulating cell-free DNA. *Cancer Biology and Therapy* vol. 20 1057–1067 (2019).

213. Shen, S. Y. *et al.* Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* vol. 563 579–583 (2018).

214. Openshaw, M. R. *et al.* Longitudinal monitoring of circulating tumour DNA improves prognostication and relapse detection in gastroesophageal adenocarcinoma. *Br. J. Cancer* **123**, 1271–1279 (2020).

215. Marczynski, G. T., Laus, A. C., dos Reis, M. B., Reis, R. M. & Vazquez, V. de L. Circulating tumor DNA (ctDNA) detection is associated with shorter progression-free survival in advanced melanoma patients. *Sci. Rep.* **10**, 1–11 (2020).

216. Zhang, E. W. *et al.* Association between circulating tumor DNA burden and disease burden in patients with ALK-positive lung cancer. *Cancer* **126**, 4473–4484 (2020).

217. Tie, J. *et al.* Circulating tumor DNA analysis detects minimal residual disease and predicts recurrence in patients with stage II colon cancer. *Sci. Transl. Med.* **8**, (2016).

218. Bratman, S. V. *et al.* Personalized circulating tumor DNA analysis as a predictive biomarker in solid tumor patients treated with pembrolizumab. *Nat. Cancer* **1**, 873–881 (2020).

219. Escudero, L. *et al.* Circulating tumour DNA from the cerebrospinal fluid allows the characterisation and monitoring of medulloblastoma. *Nat. Commun.* **11**, 1–11 (2020).

220. Ignatiadis, M., Sledge, G. W. & Jeffrey, S. S. Liquid biopsy enters the clinic — implementation issues and future challenges. *Nature Reviews Clinical Oncology* 1–16 (2021) doi:10.1038/s41571-020-00457-x.

221. Russano, M. *et al.* Liquid biopsy and tumor heterogeneity in metastatic solid tumors: The potentiality of blood samples. *Journal of Experimental and Clinical Cancer Research* vol. 39 95 (2020).

222. McEvoy, A. C. *et al.* Monitoring melanoma recurrence with circulating tumor DNA: a proof of concept from three case studies. *Oncotarget* **10**, (2019).

223. Murtaza, M. *et al.* Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **497**, 108–112 (2013).

224. Cutts, A. *et al.* Characterisation of the changing genomic landscape of metastatic melanoma using cell free DNA. *npj Genomic Med.* **2**, (2017).

225. Gardiner-Garden, M. & Frommer, M. CpG Islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282 (1987).

226. Shi, D. Q., Ali, I., Tang, J. & Yang, W. C. New insights into 5hmC DNA modification: Generation, distribution and function. *Frontiers in Genetics* vol. 8 (2017).

227. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).

228. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**, 386–389 (1998).

229. Parry, L. & Clarke, A. R. The Roles of the Methyl-CpG Binding Proteins in Cancer. *Genes Cancer* **2**, 618–630 (2011).

230. Mahmood, N. & Rabbani, S. A. DNA methylation readers and cancer: Mechanistic and therapeutic applications. *Frontiers in Oncology* vol. 9 489

(2019).

231. Michalak, E. M., Burr, M. L., Bannister, A. J. & Dawson, M. A. The roles of DNA, RNA and histone methylation in ageing and cancer. *Nature Reviews Molecular Cell Biology* vol. 20 573–589 (2019).

232. Greenberg, M. V. C. & Bourc'his, D. The diverse roles of DNA methylation in mammalian development and disease. *Nature Reviews Molecular Cell Biology* vol. 20 590–607 (2019).

233. Yang, X. *et al.* Gene body methylation can alter gene expression and is a therapeutic target in cancer. *Cancer Cell* **26**, 577–590 (2014).

234. Robertson, K. D. DNA methylation and human disease. *Nature Reviews Genetics* vol. 6 597–610 (2005).

235. Farlik, M. *et al.* DNA Methylation Dynamics of Human Hematopoietic Stem Cell Differentiation. *Cell Stem Cell* **19**, 808–822 (2016).

236. Morales-Nebreda, L., McLafferty, F. S. & Singer, B. D. DNA methylation as a transcriptional regulator of the immune system. *Translational Research* vol. 204 1–18 (2019).

237. Baron, U. *et al.* DNA demethylation in the humanFOXP3 locus discriminates regulatory T cells from activated FOXP3+ conventional T cells. *Eur. J. Immunol.* **37**, 2378–2389 (2007).

238. You, J. S. & Jones, P. A. Cancer Genetics and Epigenetics: Two Sides of the Same Coin? *Cancer Cell* vol. 22 9–20 (2012).

239. Kinney, S. R. M. & Pradhan, S. Regulation of expression and activity of DNA (Cytosine-5) methyltransferases in mammalian cells. in *Progress in Molecular Biology and Translational Science* vol. 101 311–333 (Elsevier B.V., 2011).

240. Locke, W. J. *et al.* DNA Methylation Cancer Biomarkers: Translation to the Clinic. *Frontiers in Genetics* vol. 10 1150 (2019).

241. Herman, J. G. & Baylin, S. B. Gene Silencing in Cancer in Association with Promoter Hypermethylation. *N. Engl. J. Med.* **349**, 2042–2054 (2003).

242. Watts, G. S. *et al.* Methylation of discrete regions of the O6-methylguanine DNA methyltransferase (MGMT) CpG island is associated with heterochromatinization of the MGMT transcription start site and silencing of the gene. *Mol. Cell. Biol.* **17**, 5612–5619 (1997).

243. Xia, L., Zhang, W. & Gao, L. Clinical and prognostic effects of CDKN2A, CDKN2B and CDH13 promoter methylation in ovarian cancer: a study using meta-analysis and TCGA data. *Biomarkers* **24**, 700–711 (2019).

244. 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.* **57**, (1997).

245. Kanaya, T. *et al.* Frequent hypermethylation of MLH1 promoter in normal endometrium of patients with endometrial cancers. *Oncogene* **22**, 2352–2360 (2003).

246. Zhang, W. & Xu, J. DNA methyltransferases and their roles in tumorigenesis. *Biomarker Research* vol. 5 (2017).

247. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes and Development* vol. 30 733–750 (2016).

248. Yang, H., Ye, D., Guan, K. L. & Xiong, Y. IDH1 and IDH2 mutations in tumorigenesis: Mechanistic insights and clinical perspectives. *Clinical Cancer Research* vol. 18 5562–5571 (2012).

249. Kanai, Y., Ushijima, S., Nakanishi, Y., Sakamoto, M. & Hirohashi, S. Mutation of the DNA methyltransferase (DNMT) 1 gene in human colorectal cancers. *Cancer Lett.* **192**, 75–82 (2003).

250. Turcan, S. *et al.* IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* **483**, 479–483 (2012).

251. Huang, L. E. Friend or foe-IDH1 mutations in glioma 10 years on. *Carcinogenesis* **40**, 1299–1307 (2019).

252. Koch, A. *et al.* Analysis of DNA methylation in cancer: Location revisited. *Nature Reviews Clinical Oncology* vol. 15 459–466 (2018).

253. Cao, F. *et al.* Integrated epigenetic biomarkers in circulating cell-free DNA as a robust classifier for pancreatic cancer. *Clin. Epigenetics* **12**, 1–14 (2020).

254. Nassiri, F. *et al.* Detection and discrimination of intracranial tumors using plasma cell-free DNA methylomes. *Nat. Med.* **26**, 1044–1047 (2020).

255. Nuzzo, P. V. *et al.* Detection of renal cell carcinoma using plasma and urine cell-free DNA methylomes. *Nat. Med.* **26**, 1041–1043 (2020).

256. Lasseter, K. *et al.* Plasma cell-free DNA variant analysis compared with methylated DNA analysis in renal cell carcinoma. *Genet. Med.* **22**, 1366–1373 (2020).

257. Holmila, R. *et al.* Targeted deep sequencing of plasma circulating cell-free DNA reveals Vimentin and Fibulin 1 as potential epigenetic biomarkers for hepatocellular carcinoma. *PLoS One* **12**, e0174265 (2017).

258. Widschwendter, M. *et al.* The potential of circulating tumor DNA methylation analysis for the early detection and management of ovarian cancer. *Genome Med.* **9**, 116 (2017).

259. Hulbert, A. *et al.* Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. in *Clinical Cancer Research* vol. 23 1998–2005 (American Association for Cancer Research Inc., 2017).

260. Darst, R. P., Pardo, C. E., Ai, L., Brown, K. D. & Kladde, M. P. Bisulfite sequencing of DNA. *Current Protocols in Molecular Biology* vol. CHAPTER Unit (2010).

261. Shen, S. Y., Burgener, J. M., Bratman, S. V. & De Carvalho, D. D. Preparation of cfMeDIP-seq libraries for methylome profiling of plasma cell-free DNA. *Nat. Protoc.* **14**, 2749–2780 (2019).

262. Kristensen, L. S. *et al.* Aberrant methylation of cell-free circulating DNA in plasma predicts poor outcome in diffuse large B cell lymphoma. *Clin. Epigenetics* **8**, 95 (2016).

263. Wyatt, G. R. & Cohen, S. S. A new pyrimidine base from bacteriophage nucleic acids [7]. *Nature* vol. 170 1072–1073 (1952).

264. Kriaucionis, S. & Heintz, N. The nuclear DNA base 5-hydroxymethylcytosine is present in purkinje neurons and the brain. *Science (80-.).* **324**, 929–930 (2009).

265. Tahiliani, M. *et al.* Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science (80-.).* **324**, 930–935 (2009).

266. Song, C. X., Yi, C. & He, C. Mapping recently identified nucleotide variants in the genome and transcriptome. *Nature Biotechnology* vol. 30 1107–1116 (2012).

267. He, Y. F. *et al.* Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science (80-.).* **333**, 1303–1307 (2011).

268. Hahn, M. A., Szabó, P. E. & Pfeifer, G. P. 5-Hydroxymethylcytosine: A stable or transient DNA modification? *Genomics* **104**, 314–323 (2014).

269. Pfeifer, G. P., Kadam, S. & Jin, S. G. 5-hydroxymethylcytosine and its potential roles in development and cancer. *Epigenetics and Chromatin* vol. 6 10 (2013).

270. Bachman, M. *et al.* 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **6**, 1049–1055 (2014).

271. Thomson, J. P. & Meehan, R. R. The application of genome-wide 5-hydroxymethylcytosine studies in cancer research. *Epigenomics* **9**, 77–91 (2017).

272. Globisch, D. *et al.* Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. *PLoS One* **5**, e15367 (2010).

273. Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **29**, 68–75 (2011).

274. Zeng, C., Stroup, E. K., Zhang, Z., Chiu, B. C. H. & Zhang, W. Towards precision medicine: Advances in 5-hydroxymethylcytosine cancer biomarker discovery in liquid biopsy. *Cancer Communications* vol. 39 1–9 (2019).

275. Nestor, C. E. *et al.* Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res.* **22**, 467–477 (2012).

276. Lin, I.-H., Chen, Y.-F. & Hsu, M.-T. Correlated 5-Hydroxymethylcytosine (5hmC) and Gene Expression Profiles Underpin Gene and Organ-Specific Epigenetic

Regulation in Adult Mouse Brain and Liver. *PLoS One* **12**, e0170779 (2017).

277. Li, W. & Liu, M. Distribution of 5-Hydroxymethylcytosine in Different Human Tissues. *J. Nucleic Acids* **2011**, (2011).

278. Wu, H. *et al.* Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev.* **25**, 679–684 (2011).

279. Stroud, H., Feng, S., Morey Kinney, S., Pradhan, S. & Jacobsen, S. E. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol.* **12**, R54 (2011).

280. Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S. & Heintz, N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012).

281. Ficz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–404 (2011).

282. Han, D. *et al.* A Highly Sensitive and Robust Method for Genome-wide 5hmC Profiling of Rare Cell Populations Molecular Cell Technology A Highly Sensitive and Robust Method for Genome-wide 5hmC Profiling of Rare Cell Populations. *Mol. Cell* **63**, 711–719 (2016).

283. Xu, Y. *et al.* Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Mol. Cell* **42**, 451–464 (2011).

284. Sérandour, A. A. *et al.* Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res.* **40**, 8255–65 (2012).

285. Williams, K. *et al.* TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473**, 343–349 (2011).

286. Pronier, E. *et al.* Inhibition of TET2-mediated conversion of 5-methylcytosine to 5-hydroxymethylcytosine disturbs erythroid and granulomonocytic differentiation of human hematopoietic progenitors. *Blood* **118**, 2551–2555 (2011).

287. Tsagaratou, A. *et al.* Dissecting the dynamic changes of 5-hydroxymethylcytosine in T-cell development and differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E3306–E3315 (2014).

288. Ichiyama, K. *et al.* The Methylcytosine Dioxygenase Tet2 Promotes DNA Demethylation and Activation of Cytokine Gene Expression in T Cells. *Immunity* **42**, 613–626 (2015).

289. Laird, A., Thomson, J. P., Harrison, D. J. & Meehan, R. R. 5-hydroxymethylcytosine profiling as an indicator of cellular state. *Epigenomics* **5**, 655–669 (2013).

290. Thomson, J. P. *et al.* Dynamic changes in 5-hydroxymethylation signatures underpin early and late events in drug exposed liver. doi:10.1093/nar/gkt232.

291. Thomson, J. P. *et al.* Non-genotoxic carcinogen exposure induces defined changes in the 5-hydroxymethylome. *Genome Biol.* **13**, 1–17 (2012).

292. Jin, S. G. *et al.* 5-hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer Res.* **71**, 7360–7365 (2011).

293. Chen, X. *et al.* Long non-coding RNA XIST promotes osteoporosis through inhibiting bone marrow mesenchymal stem cell differentiation. *Exp. Ther. Med.* (2018) doi:10.3892/etm.2018.7033.

294. Lian, C. G. *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of Melanoma. *Cell* **150**, 1135–1146 (2012).

295. Lian, C. G. *et al.* Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of Melanoma. *Cell* **150**, 1135–1146 (2012).

296. Song, C. X. *et al.* 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages. *Cell Res.* **27**, 1231–1242 (2017).

297. Guler, G. D. *et al.* Detection of early stage pancreatic cancer using 5-hydroxymethylcytosine signatures in circulating cell free DNA. *Nat. Commun.* **11**, 1–12 (2020).

298. Li, W. *et al.* 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers. *Cell Res.* **27**, 1243–1257 (2017).

299. Gao, P. *et al.* 5-Hydroxymethylcytosine profiling from genomic and cell-free DNA for colorectal cancers patients. *J. Cell. Mol. Med.* **23**, 3530–3537 (2019).

300. Tian, X. *et al.* Circulating tumor DNA 5-hydroxymethylcytosine as a novel diagnostic biomarker for esophageal cancer. *Cell Research* vol. 28 597–600 (2018).

301. Zhang, J. *et al.* 5-Hydroxymethylome in Circulating Cell-free DNA as A Potential Biomarker for Non-small-cell Lung Cancer. *Genomics, Proteomics Bioinforma.* **16**, 187–199 (2018).

302. Cai, J. *et al.* Genome-wide mapping of 5-hydroxymethylcytosines in circulating cell-free DNA as a non-invasive approach for early detection of hepatocellular carcinoma. *Gut* **68**, 2195–2205 (2019).

303. Chen, H. Y. *et al.* 5-Hydroxymethylcytosine profiles of cfDNA are highly predictive of R-CHOP treatment response in diffuse large B cell lymphoma patients. *Clin. Epigenetics* **13**, 33 (2021).

304. Chiu, B. C. H. *et al.* Prognostic implications of 5-hydroxymethylcytosines from circulating cell-free DNA in diffuse large B-cell lymphoma. *Blood Adv.* **3**, 2790–2799 (2019).

305. Barault, L. *et al.* Discovery of methylated circulating DNA biomarkers for comprehensive non-invasive monitoring of treatment response in metastatic colorectal cancer. *Gut* **67**, 1995–2005 (2018).

306. Gallardo-Gómez, M. *et al.* A new approach to epigenome-wide discovery of non-invasive methylation biomarkers for colorectal cancer screening in circulating cell-free DNA using pooled samples. *Clin. Epigenetics* **10**, 53 (2018).

307. Takahashi, H. *et al.* Correlation of Methylated Circulating Tumor DNA With Response to Neoadjuvant Chemotherapy in Breast Cancer Patients. *Clin. Breast Cancer* **17**, 61-69.e3 (2017).

308. Methylation Sequencing | Sequence bisulfite-converted DNA.

https://emea.illumina.com/techniques/sequencing/methylation-sequencing.html.

309. oxBS-Seq. https://emea.illumina.com/science/sequencing-method-explorer/kits-and-arrays/oxbs-seq.html.

310. Grunau, C., Clark, S. J. & Rosenthal, A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res.* **29**, 65–65 (2001).

311. Hansen, K. D., Langmead, B. & Irizarry, R. A. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* **13**, 1–10 (2012).

312. Thu, K. L. *et al.* Methylated DNA immunoprecipitation. *J. Vis. Exp.* (2009) doi:10.3791/935.

313. MeDIP-Seq/DIP-Seq. https://emea.illumina.com/science/sequencing-method-explorer/kits-and-arrays/medip-seq-dip-seq.html.

314. Shen, S. Y. *et al.* Sensitive tumour detection and classification using plasma cell-free DNA methylomes. *Nature* **563**, 579–583 (2018).

315. Chan, R. F. *et al.* Enrichment methods provide a feasible approach to comprehensive and adequately powered investigations of the brain methylome. *Nucleic Acids Res.* **45**, e97–e97 (2017).

316. Thomson, J. P. *et al.* Comparative analysis of affinity-based 5-hydroxymethylation enrichment techniques. doi:10.1093/nar/gkt1080.

317. Common Terminology Criteria for Adverse Events (CTCAE) | Protocol Development | CTEP. https://ctep.cancer.gov/protocoldevelopment/electronic_applications/ctc.htm#ctc_40.

318. AllPrep DNA/RNA/miRNA Universal Handbook - QIAGEN. https://www.qiagen.com/gb/resources/resourcedetail?id=612cdf65-324c-4710-b0e7-186d9868b0bb&lang=en.

319. Kim, D., Langmead, B. & Salzberg, S. L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

320. Home · pezmaster31/bamtools Wiki · GitHub. https://github.com/pezmaster31/bamtools/wiki.

321. Picard Tools - By Broad Institute. https://broadinstitute.github.io/picard/.

322. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166 (2015).

323. Love, M. I., Huber, W. & Anders, S. *DESeq2 Manual*. *bioRxiv* (Cold Spring Harbor Laboratory, 2014). doi:10.1101/002832.

324. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, (2014).

325. Zhang, Y., Parmigiani, G. & Johnson, W. E. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genomics Bioinforma.* **2**, (2020).

326. Fang, H., Knezevic, B., Burnham, K. L. & Knight, J. C. XGR software for enhanced interpretation of genomic summary data, illustrated by application to immunological traits. *Genome Med.* **8**, 129 (2016).

327. Venables WN, R. B. *Modern Applied Statistics in S*. (Springer, 2002).

328. Chromium Single Cell V(D)J Reagent Kits User Guide (v1.1 Chemistry) with Feature Barcoding technology for Cell Surface Protein -User Guide -Index -Single Cell Immune Profiling -Official 10x Genomics Support. https://support.10xgenomics.com/single-cell-vdj/index/doc/user-guide-chromium-single-cell-vdj-reagent-kits-user-guide-v11-chemistry-with-feature-barcoding-technology-for-cell-surface-protein.

329. Release Notes for Cell Ranger 3.0 Gene Expression & Feature Barcoding - Software -Single Cell Gene Expression -Official 10x Genomics Support. https://support.10xgenomics.com/single-cell-gene-expression/software/release-notes/3-0.

330. Package ‘scater’ Title Single-Cell Analysis Toolkit for Gene Expression Data in R

Description A collection of tools for doing various analyses of single-cell RNA-seq gene expression data, with a focus on quality control and visualization. (2021).

331. Bioconductor - scran. https://bioconductor.org/packages/release/bioc/html/scran.html.

332. Butler, A. & Satija, R. Integrated analysis of single cell transcriptomic data across conditions, technologies, and species. *bioRxiv* 164889 (2017) doi:10.1101/164889.

333. Zappia, L. & Oshlack, A. Clustering trees: a visualization for evaluating clusterings at multiple resolutions. *Gigascience* **7**, 1–9 (2018).

334. Aran, D. *et al.* Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* **20**, 163–172 (2019).

335. Mabbott, N. A., Baillie, J. K., Brown, H., Freeman, T. C. & Hume, D. A. An expression atlas of human primary cells: Inference of gene function from coexpression networks. *BMC Genomics* **14**, 632 (2013).

336. Monaco, G. *et al.* RNA-Seq Signatures Normalized by mRNA Abundance Allow Absolute Deconvolution of Human Immune Cell Types. *Cell Rep.* **26**, 1627-1640.e7 (2019).

337. Schmiedel, B. J. *et al.* Impact of Genetic Polymorphisms on Human Immune Cell Gene Expression. *Cell* **175**, 1701-1715.e16 (2018).

338. Trapnell, C. *et al.* Pseudo-temporal ordering of individual cells reveals dynamics and regulators of cell fate decisions HHS Public Access Author manuscript. *Nat Biotechnol* **32**, 381–386 (2014).

339. Package 'survminer' Type Package Title Drawing Survival Curves using 'ggplot2'. (2021).

340. Package 'survival' Title Survival Analysis Priority recommended. (2021).

341. QIAamp Circulating Nucleic Acid Handbook - QIAGEN. https://www.qiagen.com/gb/resources/resourcedetail?id=0c4b31ab-f4fb-425f-99bf-10ab9538c061&lang=en.

342. Fisher Scientific, T. *User Guide: Qubit dsDNA HS Assay Kits*.

343. Taiwo, O. *et al.* Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat. Protoc.* **7**, 617–636 (2012).

344. Technologies, A. *Agilent Technologies Agilent 2200 TapeStation System User Manual Agilent 2200 TapeStation System-User Manual Notices*.

345. HISAT2. http://daehwankimlab.github.io/hisat2/.

346. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

347. Peak calling with MACS2 | Introduction to ChIP-Seq using high-performance computing. https://hbctraining.github.io/Intro-to-ChIPseq/lessons/05_peak_calling_macs.html.

348. Shen, L., Shao, N., Liu, X. & Nestler, E. Ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics* **15**, 284 (2014).

349. Tools for the Efficient Analysis of High-Resolution Genomics Data • BRGenomics. https://mdeber.github.io/index.html.

350. Yu, G., Wang, L.-G. & He, Q.-Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. doi:10.1093/bioinformatics/btv145.

351. Gel, B. *et al.* regioneR: an R/Bioconductor package for the association analysis of genomic regions based on permutation tests. *Bioinformatics* **32**, btv562 (2015).

352. Roadmap Epigenomics Project - Home. http://www.roadmapepigenomics.org/.

353. Cavalcante, R. G. & Sartor, M. A. Annotatr: Genomic regions in context. *Bioinformatics* **33**, 2381–2383 (2017).

354. Stark, R. & Brown, G. *DiffBind: Differential binding analysis of ChIP-Seq peak data*.

355. Leek, R. D. *et al.* Association of Macrophage Infiltration with Angiogenesis and Prognosis in Invasive Breast Carcinoma. *Cancer Res.* **56**, (1996).

356. Piaggio, F. *et al.* A novel liposomal Clodronate depletes tumor-associated macrophages in primary and metastatic melanoma: Anti-angiogenic and anti-tumor effects. *J. Control. Release* **223**, 165–177 (2016).

357. Mantovani, A., Sozzani, S., Locati, M., Allavena, P. & Sica, A. Macrophage polarization: Tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends in Immunology* vol. 23 549–555 (2002).

358. Parikh, K. *et al.* Peripheral monocytes and neutrophils predict response to immune checkpoint inhibitors in patients with metastatic non-small cell lung cancer. *Cancer Immunol. Immunother.* **67**, 1365–1370 (2018).

359. Sekine, K. *et al.* Change in the lymphocyte-to-monocyte ratio is an early surrogate marker of the efficacy of nivolumab monotherapy in advanced non-small-cell lung cancer. *Lung Cancer* **124**, 179–188 (2018).

360. Chittezhath, M. *et al.* Molecular Profiling Reveals a Tumor-Promoting Phenotype of Monocytes and Macrophages in Human Cancer Progression. *Immunity* **41**, 815–829 (2014).

361. House, I. G. *et al.* Macrophage-derived CXCL9 and CXCL10 are required for antitumor immune responses following immune checkpoint blockade. *Clin. Cancer Res.* **26**, 487–504 (2020).

362. Schmidt, H. *et al.* Elevated neutrophil and monocyte counts in peripheral blood are associated with poor survival in patients with metastatic melanoma: A prognostic model. *Br. J. Cancer* **93**, 273–278 (2005).

363. Lin, G. N., Peng, J. W., Xiao, J. J., Liu, D. Y. & Xia, Z. J. Prognostic impact of circulating monocytes and lymphocyte-to-monocyte ratio on previously untreated metastatic non-small cell lung cancer patients receiving platinum-based doublet. *Med. Oncol.* **31**, (2014).

364. Failing, J. J., Yan, Y., Porrata, L. F. & Markovic, S. N. Lymphocyte-to-monocyte ratio is associated with survival in pembrolizumab-treated metastatic melanoma patients. *Melanoma Res.* **27**, 596–600 (2017).

365. Cloughesy, T. F. *et al.* Neoadjuvant anti-PD-1 immunotherapy promotes a

survival benefit with intratumoral and systemic immune responses in recurrent glioblastoma. *Nat. Med.* **25**, 477–486 (2019).

366. Khunger, M. *et al.* Post-treatment changes in hematological parameters predict response to nivolumab monotherapy in non-small cell lung cancer patients. *PLoS One* **13**, (2018).

367. Jeong, H., Hwang, I., Kang, S. H., Shin, H. C. & Kwon, S. Y. Tumor-associated macrophages as potential prognostic biomarkers of invasive breast cancer. *J. Breast Cancer* **22**, 38–51 (2019).

368. Hwang, I. *et al.* Tumor-associated macrophage, angiogenesis and lymphangiogenesis markers predict prognosis of non-small cell lung cancer patients. *J. Transl. Med.* **18**, 443 (2020).

369. Jiang, P. *et al.* Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **24**, 1550–1558 (2018).

370. Bergenfelz, C. *et al.* Systemic monocytic-MDSCs are generated from monocytes and correlate with disease progression in breast cancer patients. *PLoS One* **10**, (2015).

371. Schalper, K. A. *et al.* Elevated serum interleukin-8 is associated with enhanced intratumor neutrophils and reduced clinical benefit of immune-checkpoint inhibitors. *Nat. Med.* **26**, 688–692 (2020).

372. Iyer, S. S. & Cheng, G. Role of interleukin 10 transcriptional regulation in inflammation and autoimmune disease. *Crit. Rev. Immunol.* **32**, 23–63 (2012).

373. Ashburner, M. *et al.* Gene ontology: Tool for the unification of biology. *Nature Genetics* vol. 25 25–29 (2000).

374. Hao, Q., Vadgama, J. V. & Wang, P. CCL2/CCR2 signaling in cancer pathogenesis. *Cell Communication and Signaling* vol. 18 1–13 (2020).

375. Liu, T., Zhang, L., Joo, D. & Sun, S. C. NF-κB signaling in inflammation. *Signal Transduction and Targeted Therapy* vol. 2 1–9 (2017).

376. David, J., Dominguez, C., Hamilton, D. & Palena, C. The IL-8/IL-8R Axis: A Double

Agent in Tumor Immune Resistance. *Vaccines* **4**, 22 (2016).

377. Steele, C. W. *et al.* CXCR2 Inhibition Profoundly Suppresses Metastases and Augments Immunotherapy in Pancreatic Ductal Adenocarcinoma. *Cancer Cell* **29**, 832–845 (2016).

378. Richards, D. M., Hettinger, J. & Feuerer, M. Monocytes and macrophages in cancer: Development and functions. *Cancer Microenviron.* **6**, 179–191 (2013).

379. Trinchieri, G. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nature Reviews Immunology* vol. 3 133–146 (2003).

380. Metzemaekers, M., Vanheule, V., Janssens, R., Struyf, S. & Proost, P. Overview of the mechanisms that may contribute to the non-redundant activities of interferon-inducible CXC chemokine receptor 3 ligands. *Frontiers in Immunology* vol. 8 1970 (2018).

381. Tokunaga, R. *et al.* CXCL9, CXCL10, CXCL11/CXCR3 axis for immune activation – A target for novel cancer therapy. *Cancer Treatment Reviews* vol. 63 40–47 (2018).

382. Khan, S. *et al.* Immune dysregulation in cancer patients developing immune-related adverse events. *Br. J. Cancer* **120**, 63–68 (2019).

383. Amici, S. A. *et al.* CD38 is robustly induced in human macrophages and monocytes in inflammatory conditions. *Front. Immunol.* **9**, 1593 (2018).

384. Chen, S. *et al.* Mechanisms regulating PD-L1 expression on tumor and immune cells. *J. Immunother. Cancer* **7**, 1–12 (2019).

385. Zerdes, I., Matikas, A., Bergh, J., Rassidakis, G. Z. & Foukakis, T. Genetic, transcriptional and post-translational regulation of the programmed death protein ligand 1 in cancer: biology and clinical correlations. *Oncogene* vol. 37 4639–4661 (2018).

386. Daubon, T. *et al.* Deciphering the complex role of thrombospondin-1 in glioblastoma development. *Nat. Commun.* **10**, 1–15 (2019).

387. PPARGC1A Gene - GeneCards | PRGC1 Protein | PRGC1 Antibody.

https://www.genecards.org/cgi-bin/carddisp.pl?gene=PPARGC1A.

388. Lee, A. J. & Ashkar, A. A. The dual nature of type I and type II interferons. *Frontiers in Immunology* vol. 9 2061 (2018).

389. Sakatsume, M. *et al.* The Jak Kinases Differentially Associate with the α and β (Accessory Factor) Chains of the Interferon γ Receptor to Form a Functional Receptor Unit Capable of Activating STAT Transcription Factors *. *J. Biol. Chem.* **270**, 17528–17534 (1995).

390. Krause, C. D. *et al.* Preassembly and ligand-induced restructuring of the chains of the IFN-γ receptor complex: the roles of Jak kinases, Stat1 and the receptor chains. *Cell Res. 2006 161* **16**, 55–69 (2006).

391. Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-γ: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**, 163–189 (2004).

392. Bauvois, B., Nguyen, J., Tang, R., Billard, C. & Kolb, J. P. Types I and II interferons upregulate the costimulatory CD80 molecule in monocytes via interferon regulatory factor-1. *Biochem. Pharmacol.* **78**, 514–522 (2009).

393. Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-γ: an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**, 163–189 (2004).

394. Drevets, D. A., Leenen, P. J. M. & Campbell, P. A. Complement receptor type 3 mediates phagocytosis and killing of Listeria monocytogenes by a TNF-α- and IFN-γ-stimulated macrophage precursor hybrid. *Cell. Immunol.* **169**, 1–6 (1996).

395. Wang, F. *et al.* Interferon Gamma Induces Reversible Metabolic Reprogramming of M1 Macrophages to Sustain Cell Viability and Pro-Inflammatory Activity. *EBioMedicine* **30**, 303–316 (2018).

396. Medina-Echeverz, J. *et al.* IFN-γ regulates survival and function of tumor-induced CD11b+Gr-1high myeloid derived suppressor cells by modulating the anti-apoptotic molecule Bcl2a1. *Eur. J. Immunol.* **44**, 2457–2467 (2014).

397. Chung, B., De, M., De, A. & Bankey, P. E. Interferon-gamma induces PD-L1 expression in macrophages and neutrophils. *J. Am. Coll. Surg.* **213**, S63 (2011).

398. Ramos, J. M. P. *et al.* Type I IFNs are required to promote central nervous system immune surveillance through the recruitment of inflammatory monocytes upon systemic inflammation. *Front. Immunol.* **8**, 1666 (2017).

399. Montoya, M. *et al.* Type I interferons produced by dendritic cells promote their phenotypic and functional activation. *Blood* **99**, 3263–3271 (2002).

400. Simmons, D. P. *et al.* Type I IFN Drives a Distinctive Dendritic Cell Maturation Phenotype That Allows Continued Class II MHC Synthesis and Antigen Processing. *J. Immunol.* **188**, 3116–3126 (2012).

401. Wilson, E. B. *et al.* Emergence of distinct multiarmed immunoregulatory antigen-presenting cells during persistent viral infection. *Cell Host Microbe* **11**, 481–491 (2012).

402. Ng, C. T. & Oldstone, M. B. A. Infected CD8α- dendritic cells are the predominant source of IL-10 during establishment of persistent viral infection. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 14116–14121 (2012).

403. Wang, L. *et al.* Breast cancer induces systemic immune changes on cytokine signaling in peripheral blood monocytes and lymphocytes. *EBioMedicine* **52**, 102631 (2020).

404. Stroncek, D. F., Caruccio, L. & Bettinotti, M. CD177: A member of the Ly-6 gene superfamily involved with neutrophil proliferation and polycythemia vera. *Journal of Translational Medicine* vol. 2 8 (2004).

405. Al-Mossawi, H. *et al.* Context-specific regulation of surface and soluble IL7R expression by an autoimmune risk allele. *Nat. Commun.* **10**, 1–11 (2019).

406. Arend, W. P., Palmer, G. & Gabay, C. IL-1, IL-18, and IL-33 families of cytokines. *Immunological Reviews* vol. 223 20–38 (2008).

407. Wen, G. *et al.* A novel role of matrix metalloproteinase-8 in macrophage differentiation and polarization. *J. Biol. Chem.* **290**, 19158–19172 (2015).

408. Watson, R. *et al.* Immune Checkpoint Blockade induces peripheral cytotoxicity and persistence of large effector CD8+ T cell clones. *bioRxiv* 2020.11.15.383786 (2020) doi:10.1101/2020.11.15.383786.

409. Thomas, G. D. *et al.* Human Blood Monocyte Subsets: A New Gating Strategy Defned Using Cell Surface Markers Identifed by Mass Cytometry. *Arterioscler. Thromb. Vasc. Biol.* **37**, 1548–1558 (2017).

410. Seurat - Guided Clustering Tutorial • Seurat. https://satijalab.org/seurat/articles/pbmc3k_tutorial.html.

411. Boussiotis, V. A. Molecular and biochemical aspects of the PD-1 checkpoint pathway. *New England Journal of Medicine* vol. 375 1767–1778 (2016).

412. Wang, S. *et al.* S100A8/A9 in inflammation. *Frontiers in Immunology* vol. 9 1298 (2018).

413. CLEC4E Gene - GeneCards | CLC4E Protein | CLC4E Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=CLEC4E.

414. Wight, T. N. *et al.* Versican—A Critical Extracellular Matrix Regulator of Immunity and Inflammation. *Front. Immunol.* **0**, 512 (2020).

415. Masuda, A. *et al.* Versican is upregulated in circulating monocytes in patients with systemic sclerosis and amplifies a CCL2-mediated pathogenic loop. *Arthritis Res. Ther.* **15**, R74 (2013).

416. Colell, A., Green, D. R. & Ricci, J.-E. Novel roles for GAPDH in cell death and carcinogenesis. *Cell Death Differ. 2009 1612* **16**, 1573–1581 (2009).

417. Sobczak, M. *et al.* Downregulation of microsomal glutathione-S-transferase 1 modulates protective mechanisms in differentiated PC12 cells. *J. Physiol. Biochem. 2014 702* **70**, 375–383 (2014).

418. Reid, K. B. M. Complement Component C1q: Historical Perspective of a Functionally Versatile, and Structurally Unusual, Serum Protein. *Front. Immunol.* **0**, 764 (2018).

419. Roumenina, L. T. *et al.* Functional Complement C1q Abnormality Leads to Impaired Immune Complexes and Apoptotic Cell Clearance. *J. Immunol.* **187**, 4369–4373 (2011).

420. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma

by single-cell RNA-seq. *Science (80-.).* **352**, 189–196 (2016).

421. Mann, E. R. *et al.* Longitudinal immune profiling reveals key myeloid signatures associated with COVID-19. *Sci. Immunol.* **5**, (2020).

422. Segovia, M. *et al.* Targeting TMEM176B Enhances Antitumor Immunity and Augments the Efficacy of Immune Checkpoint Blockers by Unleashing Inflammasome Activation. *Cancer Cell* **35**, 767-781.e6 (2019).

423. Bodzioch, M. *et al.* Evidence for potential functionality of nuclearly-encoded humanin isoforms. *Genomics* **94**, 247–256 (2009).

424. MTRNR2L12 - Humanin-like 12 - Homo sapiens (Human) - MTRNR2L12 gene & protein. https://www.uniprot.org/uniprot/P0DMP1.

425. Cell type atlas - MTRNR2L8 - The Human Protein Atlas. https://www.proteinatlas.org/ENSG00000255823-MTRNR2L8/celltype.

426. Shen, Y. *et al.* Epigenome-Wide Association Study Indicates Hypomethylation of MTRNR2L8 in Large-Artery Atherosclerosis Stroke. *Stroke* **50**, 1330–1338 (2019).

427. PPBP Gene - GeneCards | CXCL7 Protein | CXCL7 Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=PPBP.

428. GP9 Gene - GeneCards | GPIX Protein | GPIX Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=GP9.

429. TUBB1 Gene - GeneCards | TBB1 Protein | TBB1 Antibody. https://www.genecards.org/cgi-bin/carddisp.pl?gene=TUBB1.

430. Allen, N. *et al.* Circulating monocyte-platelet aggregates are a robust marker of platelet activity in cardiovascular disease. *Atherosclerosis* **282**, 11–18 (2019).

431. Hottz, E. D. *et al.* Platelet activation and platelet-monocyte aggregate formation trigger tissue factor expression in patients with severe COVID-19. *Blood* **136**, 1330–1341 (2020).

432. Lee, J. *et al.* The MHC class II antigen presentation pathway in human monocytes differs by subset and is regulated by cytokines. *PLoS One* **12**, (2017).

433. Wolf, A. A., Yáñez, A., Barman, P. K. & Goodridge, H. S. The Ontogeny of

Monocyte Subsets. *Front. Immunol.* **10**, 1642 (2019).

434. Patel, A. A. *et al.* The fate and lifespan of human monocyte subsets in steady state and systemic inflammation. *J. Exp. Med.* **214**, 1913–1923 (2017).

435. Ingersoll, M. A. *et al.* Comparison of gene expression profiles between human and mouse monocyte subsets. *Blood* **115**, e10 (2010).

436. Hui, L., Maron, J. L. & Gahan, P. B. Other body fluids as non-invasive sources of cell-free DNA/RNA. *Adv. Predict. Prev. Pers. Med.* **5**, 295–323 (2015).

437. Snyder, M. W., Kircher, M., Hill, A. J., Daza, R. M. & Shendure, J. Cell-free DNA Comprises an in Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. *Cell* **164**, 57–68 (2016).

438. Sun, K. *et al.* Plasma DNA tissue mapping by genome-wide methylation sequencing for noninvasive prenatal, cancer, and transplantation assessments. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5503–E5512 (2015).

439. Stroun, M. *et al.* Neoplastic Characteristics of the DNA Found in the Plasma of Cancer Patients. *Oncology* **46**, 318–322 (1989).

440. McEvoy, A. C. *et al.* Correlation between circulating tumour DNA and metabolic tumour burden in metastatic melanoma patients. *BMC Cancer* **18**, 726 (2018).

441. Lipson, E. J. *et al.* Circulating tumor DNA analysis as a real-time method for monitoring tumor burden in melanoma patients undergoing treatment with immune checkpoint blockade. *J. Immunother. Cancer* **2**, 42 (2014).

442. Chae, Y. K. *et al.* Clinical Implications of Circulating Tumor DNA Tumor Mutational Burden (ctDNA TMB) in Non-Small Cell Lung Cancer. *Oncologist* **24**, 820–828 (2019).

443. Cohen, J. D. *et al.* Detection and localization of surgically resectable cancers with a multi-analyte blood test. *Science (80-.).* **359**, 926–930 (2018).

444. Daniotti, M. *et al.* Detection of mutated BRAFV600E variant in circulating DNA of stage III–IV melanoma patients. *Int. J. Cancer* **120**, 2439–2444 (2007).

445. Forthun, R. B. *et al.* ctDNA detected by ddPCR reveals changes in tumour load in

metastatic malignant melanoma treated with bevacizumab. *Sci. Rep.* **9**, 1–15 (2019).

446. van der Velden, V. H. J. *et al.* Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: Principles, approaches, and laboratory aspects. *Leukemia* vol. 17 1013–1034 (2003).

447. Chaudhuri, A. A. *et al.* Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. *Cancer Discov.* **7**, 1394–1403 (2017).

448. Cullinane, C. *et al.* Association of Circulating Tumor DNA with Disease-Free Survival in Breast Cancer: A Systematic Review and Meta-analysis. *JAMA Netw. Open* **3**, e2026921–e2026921 (2020).

449. Garcia-Murillas, I. *et al.* Mutation tracking in circulating tumor DNA predicts relapse in early breast cancer. *Sci. Transl. Med.* **7**, 302ra133-302ra133 (2015).

450. Esteller, M. Epigenetics in Cancer. *N. Engl. J. Med.* **358**, 1148–1159 (2008).

451. Eden, A. & Waghmare, A. Chromosomal Instability and Tumors Promoted by DNA Hypomethylation The role of the neural crest specific transcription factor AP-2 View project Understanding reprogramming of iPS View project. (2003) doi:10.1126/science.1083557.

452. Cui, H. *et al.* Loss of IGF2 imprinting: A potential marker of colorectal cancer risk. *Science (80-.).* **299**, 1753–1755 (2003).

453. Feinberg, A. P. Imprinting of a Genomic Domain of 11p15 and Loss of Imprinting in Cancer: An Introduction. *Cancer Res.* **59**, (1999).

454. Luo, H., Wei, W., Ye, Z., Zheng, J. & Xu, R. hua. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. *Trends in Molecular Medicine* vol. 0 (2021).

455. Lan, J. *et al.* Functional role of Tet-mediated RNA hydroxymethylcytosine in mouse ES cells and during differentiation. *Nat. Commun.* **11**, 1–15 (2020).

456. Cui, X. L. *et al.* A human tissue map of 5-hydroxymethylcytosines exhibits tissue specificity through gene and enhancer modulation. *Nat. Commun.* **11**, 1–11

(2020).

457. Chen, S. *et al.* Prognostic value of downregulated 5-hydroxymethyl-cytosine expression in renal cell carcinoma: A 10 year follow-up retrospective study. *J. Cancer* **11**, 1212–1222 (2020).

458. Zhao, F. *et al.* Loss of 5-Hydroxymethylcytosine as an Epigenetic Signature That Correlates With Poor Outcomes in Patients With Medulloblastoma. *Front. Oncol.* **11**, (2021).

459. Zhang, F. *et al.* 5-hydroxymethylcytosine loss is associated with poor prognosis for patients with WHO grade II diffuse astrocytomas. *Sci. Rep.* **6**, 1–14 (2016).

460. Storebjerg, T. M. *et al.* Dysregulation and prognostic potential of 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC) levels in prostate cancer. *Clin. Epigenetics* **10**, 105 (2018).

461. Yang, Q. *et al.* Decreased 5-hydroxymethylcytosine (5-hmC) is an independent poor prognostic factor in gastric cancer patients. *J. Biomed. Nanotechnol.* **9**, 1607–1616 (2013).

462. Saldanha, G. *et al.* 5-Hydroxymethylcytosine is an independent predictor of survival in malignant melanoma. *Mod. Pathol.* **30**, 60–68 (2017).

463. Dong, C. *et al.* 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic and predictive biomarkers for coronary artery disease. *Clin. Epigenetics* **12**, 1–13 (2020).

464. Cao, F. *et al.* Integrated epigenetic biomarkers in circulating cell-free DNA as a robust classifier for pancreatic cancer. *Clin. Epigenetics* **12**, 1–14 (2020).

465. Moran, S. *et al.* Epigenetic profiling to classify cancer of unknown primary: a multicentre, retrospective analysis. *Lancet Oncol.* **17**, 1386–1395 (2016).

466. Zhang, Q. & Cao, X. Epigenetic regulation of the innate immune response to infection. *Nature Reviews Immunology* vol. 19 417–432 (2019).

467. Kondilis-Mangum, H. D. & Wade, P. A. Epigenetics and the adaptive immune

response. *Molecular Aspects of Medicine* vol. 34 813–825 (2013).

468. Lio, C. W. J. & Rao, A. TET enzymes and 5hMC in adaptive and innate immune systems. *Frontiers in Immunology* vol. 10 210 (2019).

469. Peng, D. *et al.* Epigenetic silencing of TH1-type chemokines shapes tumour immunity and immunotherapy. *Nature* **527**, 249–253 (2015).

470. Jung, H. *et al.* DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat. Commun.* **10**, (2019).

471. Song, C. X. *et al.* Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **29**, 68–75 (2011).

472. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).

473. Pastor, W. A. *et al.* Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473**, 394–397 (2011).

474. Li, X., Liu, Y., Salz, T., Hansen, K. & Feinberg, A. Whole genome analysis of the methylome and hydroxymethylome in normal and malignant lung and liver. *bioRxiv* 062588 (2016) doi:10.1101/062588.

475. ChIP-seq Peak Annotation and Functional Analysis | Introduction to ChIP-Seq using high-performance computing. https://hbctraining.github.io/Intro-to-ChIPseq/lessons/12_functional_analysis.html.

476. Kent, W. J. *et al.* The Human Genome Browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).

477. Song, C. X., Yi, C. & He, C. Mapping recently identified nucleotide variants in the genome and transcriptome. *Nature Biotechnology* vol. 30 1107–1116 (2012).

478. Saxonov, S., Berg, P. & Brutlag, D. L. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1412–1417 (2006).

479. Deaton, A. M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).

480. Doi, A. *et al.* Differential methylation of tissue-and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* **41**, 1350–1353 (2009).

481. Irizarry, R. A. *et al.* The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat. Genet.* **41**, 178–186 (2009).

482. McCarthy, N. CpG islands and 5-hydroxymethylcytosine. *Nat. Rev. Cancer* **12**, 374–374 (2012).

483. Booth, M. J. *et al.* Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science (80-.).* **336**, 934–937 (2012).

484. Valpione, S. *et al.* Immune-awakening revealed by peripheral T cell dynamics after onecycle of immunotherapy. *Nat. cancer* **1**, 210 (2020).

485. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Research* vol. 21 381–395 (2011).

486. Zhao, Z. & Shilatifard, A. Epigenetic modifications of histones in cancer. *Genome Biology* vol. 20 1–16 (2019).

487. Kimura, H. Histone modifications for human epigenome analysis. *Journal of Human Genetics* vol. 58 439–445 (2013).

488. Kubiura, M., Okano, M., Kimura, H., Kawamura, F. & Tada, M. Chromosome-wide regulation of euchromatin-specific 5mC to 5hmC conversion in mouse ES cells and female human somatic cells. *Chromosom. Res.* **20**, 837–848 (2012).

489. Ficz, G. *et al.* Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473**, 398–404 (2011).

490. Shi, D. Q., Ali, I., Tang, J. & Yang, W. C. New insights into 5hmC DNA modification: Generation, distribution and function. *Frontiers in Genetics* vol. 8 (2017).

491. Zhang, T., Zhang, Z., Dong, Q., Xiong, J. & Zhu, B. Histone H3K27 acetylation is

dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol.* **21**, 45 (2020).

492. Cai, Y. *et al.* H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions. *Nat. Commun.* **12**, 1–22 (2021).

493. Kouzarides, T. Chromatin Modifications and Their Function. *Cell* vol. 128 693–705 (2007).

494. Kolasinska-Zwierz, P. *et al.* Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* **41**, 376–381 (2009).

495. Local, A. *et al.* Identification of H3K4me1-associated proteins at mammalian enhancers. *Nat. Genet.* **50**, 73–82 (2018).

496. Wei, Y. *et al.* Loss of trimethylation at lysine 27 of histone H3 is a predictor of poor outcome in breast, ovarian, and pancreatic cancers. *Mol. Carcinog.* **47**, 701–706 (2008).

497. Haffner, M. C. *et al.* Tight correlation of 5-hydroxymethylcytosine and Polycomb marks in health and disease. *Cell Cycle* **12**, 1835–1841 (2013).

498. Haffner, M. C. *et al.* Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* **2**, 627–637 (2011).

499. Pellakuru, L. G. *et al.* Global levels of H3K27me3 track with differentiation in Vivo and are deregulated by MYC in prostate cancer. *Am. J. Pathol.* **181**, 560–569 (2012).

500. King, A. D. *et al.* Reversible Regulation of Promoter and Enhancer Histone Landscape by DNA Methylation in Mouse Embryonic Stem Cells. *Cell Rep.* **17**, 289–302 (2016).

501. Chiu, B. C. H. *et al.* Alterations of 5-hydroxymethylation in circulating cell-free DNA reflect molecular distinctions of subtypes of non-Hodgkin lymphoma. *npj Genomic Med.* **6**, 1–10 (2021).

502. Cao, H., Mok, A., Miskie, B. & Hegele, R. A. Single-nucleotide polymorphisms of

the proprotein convertase subtilisin/kexin type 5 (PCSK5) gene. *J. Hum. Genet. 2001 4612* **46**, 730–732 (2001).

503. Yue, Y. *et al.* VIRMA mediates preferential m 6 A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov. 2018 41* **4**, 1–17 (2018).

504. Wee, S. *et al.* PTEN-deficient cancers depend on PIK3CB. *Proc. Natl. Acad. Sci.* **105**, 13057–13062 (2008).

505. Waters, C. E., Saldivar, J. C., Hosseini, S. A. & Huebner, K. The FHIT gene product: tumor suppressor and genome 'caretaker'. *Cell. Mol. Life Sci.* **71**, 4577 (2014).

506. Miuma, S. *et al.* Fhit Deficiency-Induced Global Genome Instability Promotes Mutation and Clonal Expansion. *PLoS One* **8**, e80730 (2013).

507. Arnold, M. *et al.* The Endosome Localized Arf-GAP AGAP1 Modulates Dendritic Spine Morphology Downstream of the Neurodevelopmental Disorder Factor Dysbindin. *Front. Cell. Neurosci.* **0**, 218 (2016).

508. Zhu, S. *et al.* Genetic alterations of TRAF proteins in human cancers. *Frontiers in Immunology* vol. 9 2111 (2018).

509. Cassandri, M. *et al.* Zinc-finger proteins in health and disease. *Cell Death Discovery* vol. 3 1–12 (2017).

510. Florl, A. R. *et al.* DNA methylation and the mechanisms of CDKN2A inactivation in transitional cell carcinoma of the urinary bladder. *Lab. Investig.* **80**, 1513–1522 (2000).

511. Orntoft, T. F. & Wolf, H. Molecular alterations in bladder cancer. *Urological Research* vol. 26 223–233 (1998).

512. Coleman, A. *et al. Distinct Deletions of Chromosome 9p Associated with Melanoma versus Glioma, Lung Cancer, and Leukemia 1*. (1994).

513. Rollins, R. A. *et al.* Large-scale structure of genomic methylation patterns. *Genome Res.* **16**, 157–163 (2006).

514. Miller, O. J., Schnedl, W., Allen, J. & Erlanger, B. F. 5-Methylcytosine localised in

mammalian constitutive heterochromatin. *Nature* **251**, 636–637 (1974).

515. Zehir, A. *et al.* Mutational Landscape of Metastatic Cancer Revealed from Prospective Clinical Sequencing of 10,000 Patients. *Nat. Med.* **23**, 703 (2017).

516. Fratta, E. *et al.* The biology of cancer testis antigens: Putative function, regulation and therapeutic potential. *Molecular Oncology* vol. 5 164–182 (2011).

517. Gjerstorff, M. F., Andersen, M. H. & Ditzel, H. J. Oncogenic cancer/testis antigens: Prime candidates for immunotherapy. *Oncotarget* vol. 6 15772–15787 (2015).

518. Barrow, C. *et al.* Tumor antigen expression in melanoma varies according to antigen and stage. *Clin. Cancer Res.* **12**, 764–771 (2006).

519. Faramarzi, S. & Ghafouri-Fard, S. Melanoma: A prototype of cancer-testis antigen-expressing malignancies. *Immunotherapy* vol. 9 1103–1113 (2017).

520. Siebenkäs, C. *et al.* Inhibiting DNA methylation activates cancer testis antigens and expression of the antigen processing and presentation machinery in colon and ovarian cancer cells. *PLoS One* **12**, e0179501 (2017).

521. Roman-Gomez, J. *et al.* Epigenetic regulation of human cancer/testis antigen gene, HAGE, in chronic myeloid leukemia. *Haematologica* **92**, 153–162 (2007).

522. Almeida, L. G. *et al.* CTdatabase: A knowledge-base of high-throughput and curated data on cancer-testis antigens. *Nucleic Acids Res.* **37**, D816–D819 (2009).

523. Igarashi, K. I. *et al.* Interferon-γ induces tyrosine phosphorylation of interferon-γ receptor and regulated association of protein tyrosine kinases, Jak1 and Jak2, with its receptor. *J. Biol. Chem.* **269**, 14333–14336 (1994).

524. Zhang, X. M. & Xu, Y. H. The associated regulators and signal pathway in rIL-16/CD4 mediated growth regulation in Jurkat cells. *Cell Res.* **12**, 363–372 (2002).

525. Hata, K. *et al.* IL-17 stimulates inflammatory responses via NF-κB and MAP kinase pathways in human colonic myofibroblasts. *Am. J. Physiol. - Gastrointest. Liver Physiol.* **282**, (2002).

526. Qian, Y., Kang, Z., Liu, C. & Li, X. IL-17 signaling in host defense and inflammatory diseases. *Cellular and Molecular Immunology* vol. 7 328–333 (2010).

527. Rudolf, R. *et al.* Architecture and Expression of the Nfatc1 Gene in Lymphocytes. *Front. Immunol.* **5**, 21 (2014).

528. Willis, B. C., Johnson, G., Wang, J. & Cohen, C. SOX10. *Appl. Immunohistochem. Mol. Morphol.* **23**, 109–112 (2015).

529. Yuan, X., Larsson, C. & Xu, D. Mechanisms underlying the activation of TERT transcription and telomerase activity in human cancer: old actors and new players. *Oncogene* vol. 38 6172–6183 (2019).

530. Ren, M., Gao, L. & Wu, X. TLR4: the receptor bridging *Acanthamoeba* challenge and intracellular inflammatory responses in human corneal cell lines. *Immunol. Cell Biol.* **88**, 529–536 (2010).

531. Vaure, C. & Liu, Y. A comparative review of toll-like receptor 4 expression and functionality in different animal species. *Frontiers in Immunology* vol. 5 316 (2014).

532. Edilova, M. I., Abdul-Sater, A. A. & Watts, T. H. TRAF1 Signaling in Human Health and Disease. *Frontiers in immunology* vol. 9 2969 (2018).

533. Walport, M. J. Complement. *N. Engl. J. Med.* **344**, 1058–1066 (2001).

534. English, W. R. *et al.* Membrane Type 4 Matrix Metalloproteinase (MMP17) Has Tumor Necrosis Factor-α Convertase Activity but Does Not Activate Pro-MMP2 *. *J. Biol. Chem.* **275**, 14046–14055 (2000).

535. Sood, R., Kamikubo, Y. & Liu, P. Role of RUNX1 in hematological malignancies. *Blood* **129**, 2070–2082 (2017).

536. Reiter, V. *et al.* The CDK5 repressor CDK5RAP1 is a methylthiotransferase acting on nuclear and mitochondrial RNA. *Nucleic Acids Res.* **40**, 6235–6240 (2012).

537. Thwaites, D. T. & Anderson, C. M. H. The SLC36 family of proton-coupled amino acid transporters and their potential role in drug transport. *British Journal of Pharmacology* vol. 164 1802–1816 (2011).

538. Kanno, T., Lin, W. D., Fu, J. L., Matzke, A. J. M. & Matzke, M. A genetic screen implicates a CWC16/Yju2/CCDC130 protein and SMU1 in alternative splicing in Arabidopsis thaliana. *RNA* **23**, 1068–1079 (2017).

539. Chen, X. *et al.* SMURF1-mediated ubiquitination of ARHGAP26 promotes ovarian cancer cell invasion and migration. *Exp. Mol. Med.* **51**, 1–12 (2019).

540. Zemirli, N., Pourcelot, M., Dogan, N., Vazquez, A. & Arnoult, D. The E3 ubiquitin ligase RNF121 is a positive regulator of NF-κB activation. *Cell Commun. Signal.* **12**, 72 (2014).

541. Gingras, S., Pelletier, S., Boyd, K. & Ihle, J. N. Characterization of a Family of Novel Cysteine- Serine-Rich Nuclear Proteins (CSRNP). *PLoS One* **2**, e808 (2007).

542. Zhang, H., Qiu, X. & Yang, G. The CSRNP Gene Family Serves as a Prognostic Biomarker in Clear Cell Renal Cell Carcinoma. *Front. Oncol.* **11**, 620126 (2021).

543. Bellazzo, A., Di Minin, G. & Collavin, L. Block one, unleash a hundred. Mechanisms of DAB2IP inactivation in cancer. *Cell Death and Differentiation* vol. 24 15–25 (2017).

544. De Donato, M. *et al.* Identification and antitumor activity of a novel inhibitor of the NIMA-related kinase NEK6. *Sci. Rep.* **8**, 1–13 (2018).

545. Vizán, P. *et al.* The Polycomb-associated factor PHF19 controls hematopoietic stem cell state and differentiation. *Sci. Adv.* **6**, eabb2745 (2020).

546. Song, G. L. *et al.* Regulation of the RhoA/ROCK/AKT/catenin pathway by arginine-specific ADP-ribosytransferases 1 promotes migration and epithelial-mesenchymal transition in colon carcinoma. *Int. J. Oncol.* **49**, 646–656 (2016).

547. Nakagawa, Y. *et al.* Identification of Semaphorin 4B as a Negative Regulator of Basophil-Mediated Immune Responses. *J. Immunol.* **186**, 2881–2888 (2011).

548. Szlachta, K. *et al.* Topoisomerase II contributes to DNA secondary structure-mediated double-stranded breaks. *Nucleic Acids Res.* **48**, 6654–6671 (2020).

549. FeiFei, W. *et al.* FBX8 degrades GSTP1 through ubiquitination to suppress colorectal cancer progression. *Cell Death Dis. 2019 105* **10**, 1–12 (2019).

550. Groom, J. R. & Luster, A. D. CXCR3 in T cell function. *Exp. Cell Res.* **317**, 620 (2011).

551. Kleiman, F. E. & Manley, J. L. Functional Interaction of BRCA1-Associated BARD1 with Polyadenylation Factor CstF-50. *Science (80-.).* **285**, 1576–1579 (1999).

552. Thomas, R. *et al.* NY-ESO-1 based immunotherapy of cancer: Current perspectives. *Frontiers in Immunology* vol. 9 947 (2018).

553. Valković, T., Lučin, K., Krstulja, M., Dobi-Babić, R. & Jonjić, N. Expression of monocyte chemotactic protein-1 in human invasive ductal breast cancer. *Pathol. - Res. Pract.* **194**, 335–340 (1998).

554. Ueno, T. *et al.* Significance of macrophage chemoattractant protein-1 in macrophage recruitment, angiogenesis, and survival in human breast cancer. *Clin. Cancer Res.* **6**, 3282–9 (2000).

555. Weide, B. *et al.* Serum markers lactate dehydrogenase and S100B predict independently disease outcome in melanoma patients with distant metastasis. *Br. J. Cancer 2012 1073* **107**, 422–428 (2012).

556. Miltenyi, S., Müller, W., Weichel, W. & Radbruch, A. High gradient magnetic cell separation with MACS. *Cytometry* **11**, 231–238 (1990).