

Antibodies, B cells, and immune tolerance in immunotherapy treated oesophageal cancer

---

Joseph Chadwick

A thesis submitted for the degree of
Doctor of Philosophy

St. Cross College
Nuffield Department of Clinical Medicine
University of Oxford

Michaelmas Term 2022

Abstract

Antibodies, B cells and immune tolerance in immunotherapy treated oesophageal cancer

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY, MICHAELMAS TERM 2022, JOSEPH CHADWICK, ST. CROSS COLLEGE, UNIVERSITY OF OXFORD

Oesophageal cancer (OC) is a leading cause of cancer mortality that is rapidly rising in prevalence in many parts of the world. Immune checkpoint inhibitors (ICI) have revolutionised treatment of many cancer types and were recently approved by the FDA to treat advanced OC. However, there are still many patients who do not exhibit an improved outcome in response to ICI therapy, and predictive biomarkers are lacking. Studies of multiple cancer types have shown that development of immune related adverse events (irAEs) during ICI treatment is associated with good clinical outcome, suggesting that the baseline immune activation of an individual may influence their response. Similarly, the presence of tertiary lymphoid structures, which are sites in which B cells differentiate into antibody secreting cells, in tumours has recently been shown to predict improved clinical outcome to ICI treatment, suggesting that production of tumour-reactive antibodies by intra-tumour B cells may contribute to an effective anti-tumour immune response. Despite this, there is a lack of studies of ICI treated OC which either evaluate markers of baseline immune activation, such as autoantibodies (AAbs), as predictive biomarkers of clinical outcome, or characterise antibodies produced by intra-tumour B cells. In this thesis, I extensively profile the circulating AAbs and peripheral blood mononuclear cells (PBMCs) of ICI treated OC patients from the LUD2015-005 clinical trial to identify novel and distinct signatures predictive of good clinical outcome and development of irAEs during ICI treatment. Further, through transcriptional profiling of OC tumours I show evidence of fully mature B cell responses and identify B cells expressing tumour-reactive antibodies. In summary, this work characterises the role of baseline immune activation in ICI patient clinical outcomes, and expands our understanding of mechanisms that predispose to good clinical outcome or irAE development during

ICI treatment. The validation of these results in larger independent cohorts could inform patient treatment decisions based on biomarkers predictive of clinical outcome.

Declaration

This thesis is submitted to the Nuffield Department of Medicine, University of Oxford, in fulfilment of the requirements for the degree of Doctor of Philosophy. This thesis is entirely my own work and describes my own research, except where otherwise stated. Where information has been derived from other sources, I confirm that this has been indicated in the thesis. Any data that are not my own work are acknowledged and referenced appropriately. This work has not been submitted for another degree in this or any other University.

Joseph Chadwick

Word count: 48,641

Acknowledgments

I would first like to extend my gratitude to all those who participated in the clinical research underpinning this thesis, particularly the patients who enrolled in the LUD2015-005 clinical trial and their families. This work would also not have been possible without the contributions of the physicians, clinical staff, and administrators who oversaw the LUD2015-005 trial and TGU Biobank. These include Naomi McGregor, Sam Dobbie, and Lucy Griffith (OCTO), Luminita Puta and the TGU Research Nurses, Stewart Norris-Bulpitt and the Sample Handling Lab (Early Phase Clinical Trials Unit), Katy McCann (Southampton General Hospital), the oversight team from Ludwig Cancer Research (including Kristen Aufiero Ramirez and Aileen Ryan), Simon Lord (Department of Oncology), and particularly Mark Middleton, who served as Chief Investigator for the LUD2015-005 study. Additionally, this research was conducted with support from AstraZeneca UK limited. This support came in the form of provision of study drugs for the LUD2015-005 trial. I would also like to thank Richard Owen and Michael White, who helped design this trial and performed much of the initial sample collection and analysis, as well as to Thomas Carroll, who was a mentor to me during my whole PhD and remains a dear friend. I am also thankful for the many others who contributed to this working group over the years, including Benedikt Kessler, Georgina Berridge, Roman Fischer, Joseph Kaplinsky, Benjamin Schuster-Böckler, Robert Amess, Brittany-Amber Jacobs, Hantao Lou, Carlos Ruiz-Puig, Jaeho Chang, Phil Xie, Ayo Omiyale, Chan Phetsouphanh, Amanda Chong, Isabela Pedroza-Pacheco, Persephone Borrow, Hannah Fuchs, Margarida Rei, and Sabrina James.

I am grateful to Xin Lu and Lynn Dustin for their mentorship and guidance over the years. I am so glad to have worked with all the members of the Lu Lab, who could always be counted on for help, advice and good company when I needed it. I will always treasure the time we spent in the lab (and the pub) together, talking about science and life!

Finally, I feel so grateful for the love and support of my parents Helen and Richard, and my brother Matt. My family has been my support system throughout my life, but especially during the past few years. I could not finish this DPhil without their encouragement, and so I dedicate this thesis to them. I hope they are proud of what I have achieved and of the person that I have become. A special mention goes to my partner Lucija. By far the best part of my DPhil was meeting you! You have been the greatest source of joy, kindness, and support over the last few years. The balance you bring to my life makes me a better person, and a better scientist.

Abbreviations

AA	Amino acid
ACAA	Anti-cytokine autoantibody
ADCC	Antibody dependent cellular phagocytosis
ADCP	Antibody dependent cellular cytotoxicity
AF	Alexa Fluor
AICDA	Activation induced cytidine deaminase
ANA	Anti-nuclear autoantibodies
APC	Antigen presenting cell
ASC	Antibody secreting cell
ASPRV1	Aspartic Peptidase Retroviral Like 1
BAFF	B cell activating factor
BAFFR	B cell activating factor receptor
BCL	Binary base call
BCL11B	BCL11 Transcription Factor B
BCMA	B cell maturation antigen
BCR	B cell receptor
BLAST	Basic local alignment search tool
BM	Bone marrow
BO	Barrett's Oesophagus
BOP1	Ribosome biogenesis protein BOP1
BSA	Bovine serum albumin
BTK	Bruton's tyrosine kinase
BTLA	B and T lymphocyte attenuator
BUV	Brilliant Ultra Violet
BV	Brilliant Violet
C	Constant
C1D8	Cycle one day eight
C1W3	Cycle one week three
CADM1	Cell Adhesion Molecule 1
CAPRIN2	Caprin Family Member 2
CB	Clinical benefit
CCAR1	Cell Division Cycle And Apoptosis Regulator 1
CCP	Cyclic citrullinated peptide
CCR	Chemokine receptor
CD	Cluster of differentiation
CDR	Conplementarity-determining region
CI	Confidence interval
CLP	Common lymphoid progenitor
CMV	Cytomegalovirus
CPS	Combined Positive Score
CR	Complete response
CRC	Colorectal cancer
CT	Computerised tomography

CTL	Cytotoxic lymphocyte
CTLA4	Cytotoxic T-lymphocyte–associated antigen 4
CTX	Chemotherapy
CXCL	Chemokine (C-X-C motif) ligand
CXCR	Chemokine (C-X-C motif) receptor
D	Diversity
DAMP	Damage associated molecular pattern
DAPI	4’,6-diamidino-2-phenylindole
DC	Dendritic cell
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DN	Double negative
DNA	Deoxyribonucleic acid
DNAM	DNAX Accessory Molecule
DP	Double positive
EBNA1	Epstein–Barr nuclear antigen 1
EBV	Epstein–Barr virus
ECOG	Eastern Cooperative Oncology Group
EDTA	Ethylenediaminetetraacetic acid
EEF2	Eukaryotic elongation factor 2
ELISA	Enzyme-linked immunosorbent assay
EO	Eosinophillic oesophagitis
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
FACS	Fluorescence-activated cell sorting
FAS	Fas Cell Surface Death Receptor
FBS	Foetal bovine serum
FC	Fragment crystallizable
FCR	Fragment crystallizable receptor
FDA	U.S. Food and Drug Administration
FDC	Follicular dendritic cell
FDR	False discovery rate
FITC	Fluorescein isothiocyanate
FLOT	Fluorouracil, Leucovorin, Oxaliplatin and Docetaxel
FMM	Finite mixture modelling
FOXP3	Forkhead box P3
FR	Framework region
FSC	Forward scatter
GAC	Gastric adenocarcinoma
GALT	Gut-associated lymphoid tissue
GC	Gastric cancer
GEX	Gene expression
GI	Gastro intestinal
GORD	Gastro-oesophageal reflux disease
GRAB	Granzyme B
GST	Glutathione S-transferase
H	Heavy

HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HEK	Human embryonic kidney
HER2	Human epidermal growth factor receptor 2
HIV	Human immunodeficiency virus
HLA	Human leukocyte antigen
HPV	Human papillomavirus
HTLV	Human T-lymphotropic virus
HV	Healthy volunteer
ICI	Immune checkpoint inhibitor
ICOS	Inducible T-cell costimulator
IFN	Interferon
IGF2BP2	Insulin Like Growth Factor 2 MRNA Binding Protein 2
IGH	Immunoglobulin heavy chain
IGHA	Immunoglobulin heavy chain alpha
IGHD	Immunoglobulin heavy chain delta
IGHG	Immunoglobulin heavy chain gamma
IGHM	Immunoglobulin heavy chain mu
IGK	Immunoglobulin kappa chain
IGKC	Immunoglobulin heavy chain kappa constant region
IGL	Immunoglobulin light chain
IL	Interleukin
IM	Intestinal metaplasia
irAE	Immune related adverse event
irRECIST	Immune related response evaluation criteria in solid tumors
KCTD21	Potassium Channel Tetramerization Domain Containing 21
KIR	Killer-cell immunoglobulin-like receptor
KLRG1	Killer Cell Lectin Like Receptor G1
LAC	Lung adenocarcinoma
LAG3	Lymphocyte-activation gene 3
LAP	Latency associated peptide
LC	Lung cancer
LPS	Lipopolysaccharide
LTR	Long term responder
MACS	Magnetic activated cell sorting
MAD	Median absolute deviation
MFI	Mean fluorescence intensity
MHC	Major histocompatibility complex
MMP14	Matrix metalloprotease 14
MMR	DNA mismatch repair
MNC	Mononuclear cell
MPO	Myeloperoxidase
MS	Multiple sclerosis
MSI	Microsatellite instability
MZ	Marginal zone
NCB	No clinical benefit

NGS	Next generation sequencing
NK	Natural killer
NKG2	Natural killer group 2
NKT	Natural killer T
NSCLC	Non-small cell lung cancer
NT	Nucleotide
OAC	Oesophageal adenocarcinoma
OC	Oesophageal cancer
OS	Overall survival
OSCC	Oesophageal squamous cell carcinoma
OX40L	OX40 ligand
P1	Panel1
P2	Panel2
P3	Panel3
PAMP	Pattern associated molecular pattern
PBMC	Peripheral blood mononuclar cell
PBS	Phosphate buffered saline
PC	Principle component
PCA	Principle component analysis
PCR	Polymerase chain reaction
PD1	Programmed cell death protein 1
PDL1	Programmed death ligand 1
PDL2	Programmed death ligand 2
PE	Phycoerythrin
PEI	Polyethylenimine
PFA	Paraformaldehyde
PFS	Progression free survival
PLZF	Promyelocytic leukemia zinc-finger
PR	Partial response
PRR	Pattern recognition receptor
PRTN3	Proteinase 3
Q1	Quadrant
Q2W	Twice weekly administration
QC	Quality control
RA	Rheumatoid arthritis
RAG	Recombination-activating gene
RB	RayBiotech
RCC	Renal cell carcinoma
REC	Research ethics comittee
RECIST	Response evaluation criteria in solid tumors
RNA	Ribonucleic acid
RNP70	70 kDa protein of the U1-snRNP complex
RPMI	Roswell Park Memorial Institute Medium
RT	Room temperature
RTX	Radiotherapy
SARS	Severe acute respiratory syndrome

SD	Standard deviation
SHM	Somatic hypermutation
SIGLEC7	Sialic Acid Binding Ig Like Lectin 7
SLE	Systemic lupus erythematosus
SLL	SciLifeLab
SLO	Secondary lymphoid organ
SOC	Standard of care
SSC	Side scatter
SSc	Systemic sclerosis
TAA	Tumour associated antigens
TAMM41	TAM41 Mitochondrial Translocator Assembly And Maintenance Homolog
TBE	Tris/Borate/EDTA
TBET	T-box expressed in T cells
TBI	Tumour burden index
TCM	T central memory
TCR	T cell receptor
TEM	T effector memory
TEMRA	T effector memory re-expressing CD45RA
TG	Thyroglobulin
TGF-β	Tumour growth factor beta
TGM2	Transglutaminase 2
TIGIT	T Cell Immunoreceptor With Ig And ITIM Domains
TIL	Tumour infiltrating lymphocyte
TLR	Toll like receptor
TLS	Tertiary lymphoid structure
TMB	Tumour mutation burden
TMC	Tumour monocyte content
TME	Tumour microenvironment
TNF	Tumour necrosis factor
TOP1	Topoisomerase 1
TORC1	CREB-regulated transcription coactivator 1
TP53	Tumour protein p53
TPM	Transcripts per million
TRIM33	Tripartite Motif Containing 33
UBE2L5	Ubiquitin Conjugating Enzyme E2 L5
UBE3A	Ubiquitin-protein ligase E3A
UMAP	Uniform Manifold Approximation and Projection
UMI	Unique Molecular Identifier
UV	Ultra violet
V	Variable
VDJ	Variable diversity joining
VH	Variable heavy
WGS	Whole genomic sequencing
ZNF	Zinc finger

Table of contents

Abstract i

Declaration iii

Acknowledgments iv

Abbreviations vi

Table of contents xi

1 Introduction 1

1.1 Oesophageal cancer 2

1.1.1 Oesophageal cancer subtypes and Barrett's Oesophagus 2

1.1.2 Detection of oesophageal cancer and current treatments 4

1.2 Immune tolerance and Immune checkpoint inhibition 5

1.2.1 The immune system, immune tolerance, and immune checkpoint inhibition 6

1.2.1.2 Immune checkpoint inhibition 9

1.2.2 Prognostic factors for clinical outcome to immunotherapy 10

1.2.2.1 Factors intrinsic to the tumour or tumour microenvironment 11

1.2.2.2 Factors extrinsic to the tumour and tumour microenvironment 13

1.3 B cells and antibody-mediated immunity 16

1.3.1 B cell development and maturation 17

1.3.1.1 B cell receptor V(D)J recombination 18

1.3.1.2 B cell maturation in secondary lymphoid organs 20

1.3.1.3 The germinal centre reaction 23

1.3.2 Immunological tolerance in the B cell compartment 24

1.3.3 Antibody isotypes and function 27

1.3.3.1 Antibody isotypes 27

1.3.3.2 Antibody functions 30

1.3.4 B cells and antibodies in human cancer 32

1.3.4.1 Anti-tumour roles of B cells and antibodies 33

1.3.4.2 Pro-tumour roles of B cells and antibodies 33

1.3.4.3 Applications of autoantibody detection in cancer 35

1.4 LUD2015-005 clinical trial of ICI treated oesophageal cancer 36

1.5 Thesis aims and hypotheses 42

2 Materials and Methods 43

2.1 LUD2015-005 clinical trial inoperable cohort 44

2.1.1 Trial design and patient selection criteria 44
2.1.2 Assessment of patient clinical outcomes 45
2.1.3 Sample collection 46
2.2 Additional clinical sample collection 47
2.2.1 Endoscopic biopsies 47
2.3 Autoantibody based antigen reactivity 48
2.3.1 Custom Luminex bead-based antigen arrays 48
2.3.1.1 Antigen screening 48
2.3.1.2 Quality control and initial data processing 49
2.3.2 RayBio autoimmune array 50
2.3.3 Determining cytomegalovirus (CMV) serostatus 50
2.4 DNA and RNA extraction 51
2.4.1 Bulk NGS library preparation and sequencing 51
2.4.1.1 Whole-transcriptome RNA sequencing 51
2.4.1.2 Whole genome sequencing 52
2.5 Single-cell RNA sequencing 52
2.5.1 Tissue dissociation 52
2.5.2 Cell encapsulation, library preparation and sequencing 53
2.6 Processing of whole blood 54
2.7 Cloning, expression, and purification of recombinant antibodies 54
2.7.1 Cloning into expression vectors 54
2.7.2 HEK 293T/F transfection 55
2.7.3 Antibody purification 56
2.8 Cell culture 57
2.9 Microscopy 57
2.10 Flow cytometry 58
2.10.1 Ex vivo immunophenotyping 58
2.10.2 Recombinant antibody binding to untreated and cisplatin treated cell lines 64
2.10.3 Antibody dependent cellular phagocytosis assay 64
2.11 Bioinformatics 65
2.11.1 Autoantibody profiling 65
2.11.1.1 Custom Luminex bead-based antigen arrays 65
2.11.1.2 Raybiotech AAb array 66
2.11.2 Ex vivo immunophenotyping 66
2.11.3 Whole-transcriptome RNA sequencing 68
2.11.3.1 Processing of raw data 68

2.11.3.2 Deconvolution of bulk RNA-seq data 68
2.11.4 Whole genome sequencing 69
2.11.5 Single-cell RNA sequencing 69
2.11.6 Single-cell B cell receptor analysis 70
2.11.7 Statistical analysis 70
2.12 Figure generation 71
3 Autoantibody profiling to predict clinical outcome to immunotherapy 72
3.1 Rationale 73
3.2 Autoantibody profiling of blood plasma from oesophageal cancer patients using two custom antigen panels 75
3.2.1 Antigen panel 1: Reactivity to DNA- and RNA- binding proteins in oesophageal cancer 76
3.2.1.1 Rationale 76
3.2.1.2 Antigen panel design 77
3.2.1.3 Sample selection for the discovery cohort screened against antigen panel 1 78
3.2.1.4 Pre-treatment number of reactive antigens are positively associated with clinical outcome 78
3.2.2 Antigen panel 2: Reactivity to immune associated and common autoantigens in oesophageal cancer, Barrett's oesophagus, and lung cancer 80
3.2.2.1 Antigen panel design 81
3.2.2.2 Sample selection 81
3.2.2.3 Analysis of data from antigen panel 2 does not reproduce the findings from antigen panel 1 82
3.3. Development of a novel analysis workflow which improves reproducibility of binary data 86
3.3.1 Rationale 86
3.3.2 A Gaussian finite mixture modelling workflow to assign binary serostatus 87
3.3.3 Binary serostatus data generated by the FMM workflow is reproducible across panels 91
3.3.4 Reproducibility of binary data is improved when modelling on just the LUD2015-005 cohort 93
3.4. Autoantibodies and clinical outcome in the LUD2015-005 cohort 99
3.4.1 Rationale 99
3.4.2 Associations between the total number of reactive antigens and clinical outcome 99
3.4.2.1 The pre-treatment number of reactive antigens shows a non-significant negative association with change in tumour size following four weeks of ICI treatment 99
3.4.2.2 Analysis of samples from the ICI-4W timepoint shows the same trends as the PreTx samples 102
3.4.2.3 Autoantibody profiles remain relatively stable following four weeks of ICI treatment 105
3.4.3 Analysing reactivity to individual antigens as biomarkers 106

3.4.3.1 Pre-treatment no antigens show differential reactivity between CB and NCB patients 106
3.4.3.2 After four weeks of ICI treatment no antigens show differential reactivity between CB and NCB patients 107
3.5. Exploratory cohort: autoantibodies in Lung cancer, Barrett's Oesophagus, and Oesophageal cancer 109
3.5.1 Rationale 109
3.5.2 Biomarkers of response to immunotherapy in lung cancer 109
3.5.2.1 Clinical summary 109
3.5.2.2 irRECIST responders show an increased number of reactive antigens PreTx 111
3.5.2.3 Pre-treatment no antigens show differential reactivity between irRECIST responders and non-responders 112
3.5.3 Anti-p53 autoantibodies are a promising biomarker to non-invasively distinguish Oesophageal cancer patients from Barrett's oesophagus patients 114
3.6. Autoimmune associated autoantibodies and clinical outcome in the LUD2015-005 cohort . 116
3.6.1 Rationale 116
3.6.2 Data overview and workflow 116
3.6.3 No association of number of autoantibodies PreTx with clinical benefit or irAEs 119
3.7 Exploring the aetiology of autoantigen reactivity 121
3.7.1 Rationale 121
3.7.2 Mutations 121
3.8.2 Cross-reactivity with infectious agents 122
3.9 Discussion 124
3.9.1 Development of a novel analysis workflow to evaluating AAb based antigen reactivity to an array of autoantigens 124
3.9.2 AAb profiling and clinical outcome to ICI treatment 126
3.9.3 Anti-p53 autoantibodies are a highly specific biomarker for distinguishing OC patients from BO patients 127
3.9.4 Aetiology of autoantigen reactivity 128
4 Peripheral immune cell profiling to assess signatures of immune activation 130
4.1 Rationale 131
4.2 Samples, experimental approach, and data processing 132
4.2.1. Samples 132
4.2.2. Experimental approach and data processing 133
4.3 Identifying cell subsets using unsupervised clustering 141
4.3.1 Clustering of mononuclear cells identifies multiple cell subsets 142
4.3.1.1 Seven subsets of MNCs identified by clustering with the markers from panel 1 (T cell centric) 142

4.3.1.3 Six subsets of MNCs identified by clustering with the markers from panel 2 (NK cell centric) 143
4.3.1.3 Seven subsets of MNCs identified by clustering with the markers from panel 3 (B cell and myeloid centric) 145
4.3.2. Nine subsets of B cells identified by clustering with the markers from panel 3 (B cell and myeloid centric) 146
4.3.3. Seven subsets of CD4 T cells identified by clustering with the markers from panel 1 (T cell centric) 148
4.3.4. Clustering of CD8 T cells identifies multiple cell subsets 150
4.3.4.1 Nine subsets of CD8 T cells identified by clustering with the markers from panel 1 (T cell centric) 150
4.3.4.2 Seven subsets of CD8 T cells identified by clustering with the markers from panel 2 (NK cell centric) 152
4.3.5. 11 subsets of NK cells identified by clustering with the markers from panel 2 (NK cell centric) 154
4.3.6. Three subsets of monocytes identified by clustering with the markers from panel 3 (B cell and myeloid centric) 157
4.3.7. Four subsets of myeloid dendritic cells (mDCs) identified by clustering with the markers from panel 3 (B cell and myeloid centric) 159
4.4 Associations of abundance and expression profiles of cell subsets with clinical factors 161
4.4.1. Rationale 161
4.4.2. Clinical benefit of patients from the LUD2015-005 inoperable cohort 162
4.4.2.1. Differences in CD8 T cells and B cells are detectable pre-treatment 162
4.4.2.2. Increased abundance of NK cells after four weeks of ICI is associated with clinical benefit 167
4.4.3. Development of irAEs in patients from the LUD2015-005 inoperable cohort 171
4.4.3.1. At the PreTx timepoint, irAE developers show signatures of reduced LAP expression and a skew towards a more immature and activated B cell compartment 171
4.4.3.2. After four weeks of ICI treatment, irAE developers no longer show signatures of reduced LAP expression and increased B cell activation 175
4.4.4. Autoantibody seropositivity is associated with decreased CD22 expression by DN2/activated naïve-like B cells, and an increased IgA staining of myeloid cells 177
4.4.5. Barrett's Oesophagus patients show significantly increased IgA expression by B cells than Oesophageal cancer patients 184
4.5 Discussion 188
4.5.1 Comprehensive immunophenotyping of peripheral immune cells from patients with OC and BO 188
4.5.2 Signatures predictive of clinical outcome of OC patients to ICI treatment 188
4.5.3.1. Clinical benefit patients show signatures of increased activation of NKG2C+ CD8 T cells 188

4.5.3.2. Development of immune related adverse events is associated with signatures of reduced LAP expression and a skew towards a more immature and activated B cell compartment 190
4.5.3 Autoantibody seropositive patients show signatures of a dysregulated antibody response 192
4.5.4 IgA expression by B cells is a promising biomarker to distinguish BO and OC patients 194
4.6 Conclusion 195
5 Intra-tumour B cells and antibodies in immunotherapy treated oesophageal cancer 196
5.1 Rationale 197
5.2 Intratumour transcriptomic signatures of tertiary lymphoid structures and B cells in patient clinical outcome 199
5.3 Single cell RNA sequencing identifies expanded BCR clones within patient tumours 208
5.3.1. Transcriptomic signals of germinal centre responses 209
5.3.2. Selection of expanded BCR clones for in vitro screening 216
5.4 Recombinant antibodies bind to human oesophageal adenocarcinoma tissue and cell lines 217
5.4.1 Recombinant antibodies bind to cancer cell lines 218
5.4.2 Recombinant antibodies preferentially bind to annexin V-positive cancer cell lines and promote their phagocytosis by myeloid cells 220
5.5 Antigen characterisation of recombinant antibodies 223
5.6 Discussion 227
5.6.1 Transcriptomic measurements of intra-tumour tertiary lymphoid structures and B cells 227
5.6.2 Evaluating clonal expansion and subtype and isotype distribution of tissue infiltrating B cells 228
5.6.3 Tumour binding capability, function, and antigen reactivity of antibodies expressed by expanded BCR clones identified from OC tumours 230
5.7 Conclusion 234
6 Conclusion and future directions 235
6.1 Overall conclusions 236
6.1.1 Signatures of peripheral blood CD8 T cells and B cells, but not autoantibodies, reflect likelihood of clinical benefit and irAE development 236
6.1.2 Tumour infiltrating B cells undergo germinal centre responses and, in some cases, express tumour-reactive antibodies 238
6.1.3 Circulating anti-p53 autoantibodies and IgA-expressing B cells distinguish individuals with oesophageal cancer from those with Barrett's Oesophagus 239
6.2 Limitations 240
6.3 Future directions 240
7 References 242

1 Introduction

1.1 Oesophageal cancer

There were over 19 million new cancer cases and almost 10 million deaths from cancer globally in 2020[1]. Amongst these, oesophageal cancer (OC) was the 10th most prevalent cancer and the cancer type with the 6th leading cause of mortality in the same timeframe, with over 600,000 new cases and over 540,000 deaths in 2020[1]. Concerningly, the number of new cases of OC of the oesophageal adenocarcinoma (OAC) subtype are predicted to rise dramatically over the coming years, particularly in high-income countries, due to increasing risk associated with lifestyle choices[2]. Given the high mortality rate of OC, which is in part due to the frequent late-stage diagnosis that requires an invasive endoscopy, the key challenges in the field are to identify patients at high risk of OC development and identify alternative treatments to the standard of care (SOC) chemotherapy (CTX), radiotherapy (RTX), and surgery[3,4].

In this section I will describe the differences between the two major subtypes of OC and the associated pre-malignant inflammatory condition Barrett's Oesophagus (BO), the challenges involved in early detection of OC and current treatments, including recent developments in harnessing immunotherapy to treat advanced OC.

1.1.1 Oesophageal cancer subtypes and Barrett's Oesophagus

Although the majority of OC cases occur in Eastern Asia and parts of Africa, there is substantial geographical variation in incidence between the two main histological subtypes of OC, which are OAC and oesophageal squamous cell carcinoma (OSCC)[1]. Despite OSCC accounting for the vast majority of OC cases globally[1,5,6], over the past few decades OAC incidence in North American and Western European countries, including the United Kingdom, has increased substantially to the point that it is now the predominant subtype of OC in these countries[1,6–8]. This geographical variation in incidence is thought to be reflective of the differing risk factors which have been identified for each OC subtype. For OSCC, risk factors include nutrient deficiencies, alcohol consumption, tobacco use and

consumption of very hot liquids and food[1,5–9]. In contrast, risk factors for OAC include being of male sex, obesity, or suffering from gastro-oesophageal reflux disease (GORD) and/or BO[1,6–9].

In BO, the stratified layers of squamous epithelial cells that usually make up the normal oesophagus are replaced by columnar epithelium, in a process termed intestinal metaplasia (IM)[10,11]. This process is thought to be initiated by damage to the epithelium of the distal oesophagus, most likely by reflux of acid and bile[10,12]. BO is thought to progress to OAC sequentially via an intermediate step known as dysplasia, which is classified as low or high grade, and in which the BO cells appear more abnormal than they do following IM[13].

As typical symptoms of BO are similar to those of GORD, but can also include unintentional weight loss and dysphagia, endoscopic examination of the oesophagus and histological examination of a tissue biopsy is required to diagnose BO[7,10–12]. However, a large proportion of patients with BO are thought to be asymptomatic, increasing the likelihood that many cases of BO are undetected[14–16]. The patients that are diagnosed with BO have a 30-fold increased risk of OAC development compared to the general population[9], although this reportedly only equates to approximately 0.3% of BO patients developing OAC within 1 year of diagnosis with BO[17]. However, when only BO patients with high grade dysplasia are considered, the percentage of patients progressing to OAC is estimated to be much higher[18,19]. Given the increased risk of OAC development for patients with BO, once diagnosed with BO, patients are often routinely monitored via endoscopy and there are procedures to eradicate the BO, such as endoscopic mucosal resection and radiofrequency ablation, available for patients with high grade dysplasia[11,13,20].

1.1.2 Detection of oesophageal cancer and current treatments

Symptoms associated with OAC are similar to those of GORD and BO and so, like BO, OC diagnosis requires endoscopy and often pathological analysis of a biopsy[6]. This means that OC is often detected at a late stage. Cancer staging, by the TNM (tumour, node, metastasis) staging system often based on a computed tomographic (CT) scan, is reflective of how advanced the growth of the cancer is and whether it has spread from the organ of origin to another part of the body, known as metastasis[6,21]. The T state refers to how large the primary tumour is and the depth at which it has invaded local tissue[21]. Briefly, stage T1 cancers have not grown further than the submucosa of the oesophagus, stage T2 cancers have grown into the muscle layer of the oesophagus but not further, stage T3 cancers have grown into the adventitia which is the external membrane of the oesophagus, and stage T4 cancers have grown into structures external to but adjacent to the oesophagus[21,22]. The N status refers to the number of lymph node that tumour cells have been detected in and the M status refers to whether the tumour has metastasised to other organs of the body[21]. Together the T, N and M statuses of a tumour are considered to determine the treatment options available for OC[21].

Patients with stage T1 can often be treated without any CTX and can undergo endoscopic resection or ablation procedures similar to those used for BO[4,23]. For patients with stage T2, T3, or T4 cancers that are deemed to be resectable, surgical removal of part of the oesophagus by oesophagectomy, usually after a course of neoadjuvant CTX with or without concurrent RTX, is the common treatment pathway[4,21,23]. Although this can be effective at removing the cancer, oesophagectomy is a major operation which is associated with substantial morbidity and can still leave the patient susceptible to recurrence of the cancer[21]. According to the NIH National Cancer Institute Surveillance, Epidemiology, and End Results Program (SEER) registry of cancer cases, more than 30% of diagnosed OC show distant metastasis and so curative surgery is not a viable treatment option[24]. In these cases, patients are often given palliative therapy which is usually CTX and medication to manage symptoms[4,25]. For patients with Human Epidermal Growth Factor Receptor 2 (HER2, also known as ERBB2) positive metastatic

OC, the anti-HER2 monoclonal antibody (mAb) trastuzumab (Herceptin) is also offered in addition to CTX[4]. However, a recent phase III clinical trial showed that the addition of trastuzumab to neoadjuvant CTX+RTX for OC overexpressing HER2 did not show improved clinical outcomes compared to CTX+RTX alone[26]. Consequently, more effective treatments are needed for these patients with advanced OC.

Immune checkpoint inhibitor (ICI) therapy, alone or in combination with concurrent CTX, was recently approved by the FDA for treatment of advanced OC[27–29]. However, despite some patients showing an improvement in clinical outcome compared to SOC CTX, there were still many patients who did not benefit clinically from the addition of ICI treatment to CTX[30–36]. Hence, understanding the heterogeneity of patient clinical outcome and discovering biomarkers that can robustly predict which OC patients are likely to respond to ICI+CTX are key aims for the field[37,38].

1.2 Immune tolerance and Immune checkpoint inhibition

ICI is the most common type of immunotherapy treatment and is used to treat multiple types of cancer[39–41]. Unlike CTX and other common cancer therapies which aim to directly kill the cancer, ICI treatment aims to activate the host immune system to eradicate the cancer[42]. However, there is considerable variability in patient response to ICI and so there has been much research into prognostic factors which predict patient clinical outcome[43].

In this section I will describe the basic structure of the immune system, the concepts of immune homeostasis and tolerance and how they are disrupted in autoimmunity and cancer, the basic mechanism by which ICI is believed to act, and reported prognostic factors for clinical outcome to ICI treatment.

1.2.1 The immune system, immune tolerance, and immune checkpoint inhibition

The primary function of the mammalian immune system is to maintain health and prevent disease. It is classically split into two main arms, innate immunity and adaptive immunity, which have equally crucial but largely distinct roles.

Innate immunity is the first line of defence against infectious agents and acts rapidly in a largely antigen agnostic manner through the use of germline encoded molecules such as Pattern Recognition Receptors (PRRs)[44]. PRRs are bound by pathogen associated molecular patterns (PAMPs) which are molecules conserved across many pathogens, such as components of cell walls and viral DNA or RNA[44,45]. PRRs also bind to damage associated molecular patterns (DAMPs) which are endogenous intracellular molecules that are released by dying or damaged cells[44,46]. The main cell types expressing PRRs and contributing to innate immunity are myeloid cells, natural killer (NK) cells and dendritic cells (DCs), although other immune and non-immune cells also play a role[44,47]. Binding of PAMPs or DAMPs to PRRs activates immune cells and induces inflammation to eliminate pathogens or repair damaged tissue[48]. In addition to providing an immediate inflammatory response, innate immunity is also important for initiation of adaptive immunity[49].

In contrast to the rapid and largely non-specific innate immune response, adaptive immunity is characterised by an antigen-specific response, which requires gene rearrangement after antigen encounter and is therefore considerably slower[47,49]. The main effector cells of adaptive immunity are lymphocytes which consist primarily of B cells and T cells. Activated innate immune cells express signalling molecules called cytokines and costimulatory molecules such as Cluster of Differentiation 80 (CD80) to stimulate lymphocytes[45,47,49]. In particular, DCs play a vital role in initiating an effective and appropriate adaptive immune response[50,51]. Different subsets of DCs express different PRRs which determines the type of pathogen they can detect, and which set of cytokines they produce when stimulated[45,52]. These sets of cytokines then appropriately direct differentiation, migration, and

response of lymphocytes[45,53]. In addition to producing cytokines, DCs are the primary antigen presenting cells (APCs) and so are responsible for processing and displaying peptide antigen on their surface via the Major Histocompatibility Complex (MHC) proteins to T cells so that they can recognise antigen[50,51]. There are two classes of MHC which can be primarily distinguished by the cell types they are expressed by and which subset of T cell they are recognised by[54]. MHC class I is expressed on all nucleated cells and presents antigen to CD8+ T cells which predominantly function as cytotoxic cells that are important for killing of tumour cells and virus-infected cells[54–56]. In contrast, MHC class II is primarily expressed by APCs and presents antigen to CD4+ T cells which mainly function to coordinate the immune response via cytokine secretion and interaction with other immune cells[54,57]. Antigen presentation ensures that only T cells which receive co-stimulation and have a T cell receptor (TCR) that binds to the presented antigen are activated[55].

Unlike T cells, B cells can recognise soluble antigen without it being presented by MHC[58]. However, B cells still require additional stimulatory and survival signals from innate and adaptive immune cells to become activated[59]. Once activated, the primary role of B cells is to produce antigen specific antibodies[59]. Upon antigen binding, antibodies function to neutralise the infectivity of pathogens and target them for destruction by effector cells or molecules[60]. In addition to antibody production, B cells can also modulate the immune responses through cytokine production and antigen presentation[59,61].

Although primarily intended to recognise foreign antigens, the immune repertoire of an individual's lymphocyte pool is also capable of recognising antigens expressed by host cells. This can lead to immune cells erroneously attack healthy host cells, which is known as autoimmunity. To limit autoimmunity, immune tolerance mechanisms are present throughout lymphocyte development and maturation to ensure that any lymphocytes that recognise host antigen, known as autoreactive, are prevented from maturing or are rendered functionally inactive[62,63].

Immune homeostasis refers to the balance between limiting autoimmunity through immune tolerance and eradicating infectious agents through immune reactivity[64] (**Figure 1.1**). In health, immune homeostasis is successfully maintained but in disease may be disrupted, and inappropriate immune reactivity or tolerance persists. However, autoreactive lymphocytes can evade tolerance mechanisms and lead to pathogenic autoimmune responses if they receive sufficient stimulation.

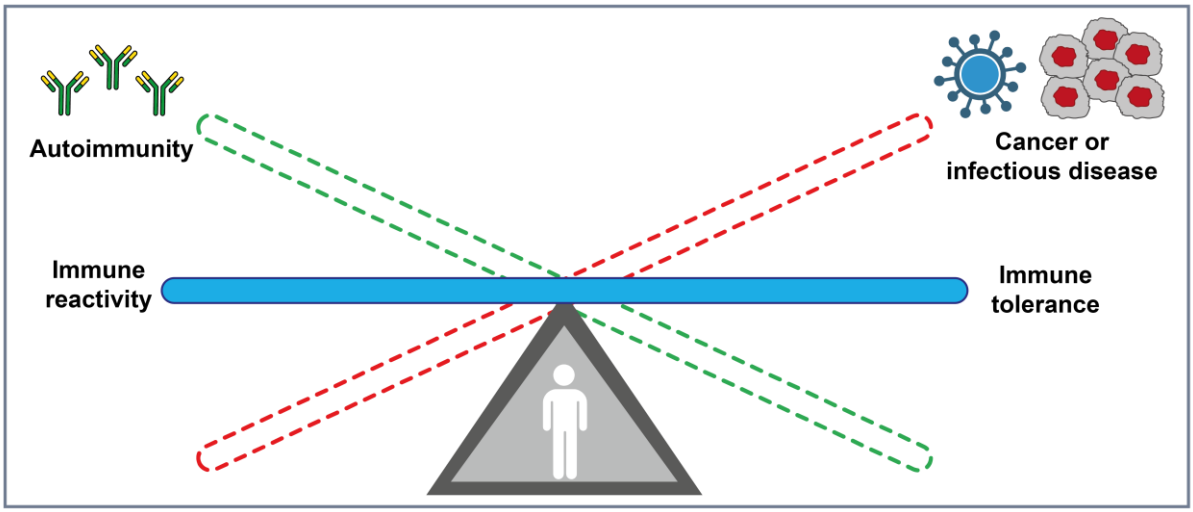


Figure 1.1. Immune homeostasis maintains the health of an organism. Immune homeostasis is the balance between immune tolerance, which is the process of limiting immune responses to self-antigens, and immune reactivity, which is the process in which infectious agents and cellular debris are eliminated from the host. If immune tolerance is not maintained, then disease can occur. Inappropriate immune reactivity to self-antigens can lead to autoimmunity, whilst insufficient immune reactivity to malignant cells or infectious agents can lead to cancer or infectious disease. This figure was created using some icons from Biorender.com.

Tumours are often sites of chronic inflammation so there can be a plethora of immune cells in the tumour microenvironment (TME) which can eliminate the tumour cells[65]. However, one of the hallmarks of cancer is immune evasion, which reflects the multiple mechanisms by which tumour cells can limit the anti-tumour immune response[66]. Mechanisms of immune evasion are known but the most well studied example is overexpression of immune checkpoint molecules such as Programmed Death Ligand 1 (PD-L1)[67]. In health, the main function of PD-L1 is to suppress autoreactive T cell responses by inducing inhibitory signalling after binding to its receptor Programmed Cell Death Protein 1 (PD-1)[68–70]. However, many tumours have evolved to exploit this by expressing PD-L1 themselves or increasing PD-L1 expression by other cell types in the TME[67]. There are many other immune

checkpoint molecules expressed by the human immune system in addition to PD-1 and PD-L1, such as cytotoxic T-lymphocyte associated protein 4 (CTLA-4), and many of these are dysregulated in cancer [71,72].

1.2.1.2 Immune checkpoint inhibition

ICI treatment aims to overcome immune tolerance and initiate robust immune reactivity against the tumour of the recipient[73]. To achieve this, ICI utilises monoclonal antibodies (mAbs) to block interactions between immune checkpoint molecules and their ligands[74]. PD-1 and CTLA-4 are the two most targeted immune checkpoint molecules[74]. Many immune cell types express PD-1, including T cells, and interaction of PD-1 with its ligands PD-L1 and PD-L2 leads to suppression of the PD-1 expressing cell[75]. CTLA-4 is expressed by both activated CD4+ and CD8+ T cells and has a similar suppressive function as PD-1. However, CTLA-4 is mechanistically slightly different to PD-1 in that the ligands for CTLA-4, CD80 and CD86 (also known as B7-1 and B7-2 respectfully), also bind to the costimulatory molecule CD28 which is vital for effective T cell activation[76]. CTLA-4 therefore functions to prevent excessive T cell activation by competing with CD28 for CD80/CD86 binding[76,77]. By blocking these interactions ICI aims to reduce immune checkpoint mediated inhibition of tumour infiltrating T cells by tumour cells, therefore enhancing anti-tumour immunity[71,74,76,77].

ICI has been used to treat many different types of cancer and has shown improved clinical outcomes compared to SOC treatments[74]. However, although some patients experience desirable and durable clinical outcomes to ICI, many patients treated with ICI do not fare any better clinically than they would be expected to if treated with SOC CTX[42,43]. It is unclear why there is such variability in clinical outcome to ICI treatment, but studies of prognostic factors associated with clinical outcome have provided clues into the biology underlying this heterogeneity.

1.2.2 Prognostic factors for clinical outcome to immunotherapy

There are two main interests in identifying prognostic factors associated with clinical outcome to ICI. Firstly, these factors could provide insight into potential mechanisms underlying response and resistance to ICI. Understanding these mechanisms could help to guide development of novel therapeutics which could be used in combination with ICI. Secondly, prognostic biomarkers would allow for stratification of patients to a treatment that they are most likely to benefit from. This is particularly important as ICI treatment is expensive and has potentially serious side effects associated with it which can detrimentally affect patient quality of life, particularly when given in combination with CTX[42,73,78].

As ICI was only recently approved for treatment of OC, there are limited studies that have reported prognostic markers of clinical outcome[37,38,79]. For OSCC, despite suggestions from previous studies[80,81], predictive biomarker analysis of pre-treatment (PreTx) tumour samples from more recent studies did not identify histological expression of PD-L1 as a biomarker associated with improved clinical outcomes[34,82,83]. Instead, the ORIENT-2 study of patients treated with the ICI anti-PD1 mAb sintilimab alone reported that increased prevalence of T follicular helper cells (Tfh) and activated B cells within tumours PreTx were associated with improved progression free survival (PFS)[82]. Analysis of PreTx blood samples from this study also identified low neutrophil-to-lymphocyte ratio (NLR), increased TCR clonality and lower molecular tumour burden index (mTBI) from circulating tumour DNA (ctDNA) as predictive biomarkers of improved clinical outcome[82]. Another study showed CTX refractory OSCC patients treated with the anti-PD-1 mAb toripalimab showed better clinical outcome if their tumour did not show amplifications of chromosome region 11q13[83].

For OAC, a previous study by our lab used whole genome sequencing (WGS), bulk- and single-cell ribonucleic acid (RNA) sequencing of OAC tumour biopsies to identify increased PreTx tumour monocyte content (TMC) and tumour mutational burden (TMB) as predictive biomarkers of good

clinical outcome to ICI+CTX in the inoperable arm of the LUD2015-005 clinical trial of OC (unpublished). This research also reported a novel transcriptional signature of 70 immune related genes, termed INCITE, which were upregulated in the tumour biopsies of all OAC patients following four weeks of treatment with ICI alone. Further, the magnitude of increase in INCITE genes over the four weeks of ICI treatment positively correlated with tumour shrinkage over the same timeframe, perhaps suggesting that early tumour shrinkage following treatment with ICI alone is mediated by influx of immune cells.

Although our current knowledge of prognostic biomarkers of clinical outcome to ICI in OC is limited, much has been reported in other cancer types in which ICI treatment has been more thoroughly studied, such as melanoma[41,84–87]. These prognostic factors can be broadly divided into two groups based on whether they are intrinsic or extrinsic to the tumour or TME[87].

1.2.2.1 Factors intrinsic to the tumour or tumour microenvironment

Most previous studies have focused on prognostic factors intrinsic to the tumour or TME as these are thought to be the most informative for understanding mechanisms of response and resistance to ICI[87]. The most notable of these are histological staining of PD-L1, tumour DNA alterations such as TMB, and intra-tumour tertiary lymphoid structures (TLSs).

As briefly discussed in the context of OC, greater expression of PD-L1 in histological tumour sections has been reported by some studies to have shown predictive power in determining which patients are likely to respond to ICI treatment in multiple cancer types[80,81,88,89]. Histological PD-L1 expression is commonly assessed by the combined positive score (CPS) method which is calculated based on the number of PD-L1 positive cells, including both tumour and immune cells, relative to the total number of viable tumour cells in a histological section[90,91]. Mechanistically, high histological PD-L1 expression in the tumour is thought to indicate high immune checkpoint mediated immune suppression within

the tumour which can be relieved by ICI treatment[92]. However, other studies have also shown that patients with low histological PD-L1 expression by CPS assessment can still go on to show good clinical outcome to ICI treatment, suggesting that other factors influence patient outcome[93,94].

This is demonstrated by another well studied tumour intrinsic factor, TMB[95,96]. TMB is the number of mutations, base substitutions, insertions and deletions that a tumour has relative to normal somatic tissue from the same patient[40,97]. Several clinical studies have shown that patients that have tumours with a high TMB are more likely to have good clinical outcome to ICI treatment[96,98–100]. Mechanistically, a high TMB is believed to predispose to improved clinical outcome to ICI by increasing the number of neoantigens that a tumour expresses[99]. Neoantigens are absent from the normal human genome and are created by changes in the tumour DNA which result in translation of novel protein sequences which can be targeted by activated immune cells, particularly after they are activated by ICI[101]. Other genomic aberrations of tumours such as microsatellite instability (MSI) and mismatch repair deficiency (MMR-D) have also been shown to predict good clinical outcome to ICI in multiple cancer types[102–104]. Like TMB, MSI and MMR are thought to increase the number of neoantigens that a tumour expresses[104]. However, like histological PD-L1 expression, high TMB, MSI and MRR have shown inconsistent associations with prognosis, highlighting the complexity of response to ICI and subsequent need for research into alternative factors[95,105].

More recently, the presence of tertiary lymphoid structures (TLSs) within or close to the tumour prior to ICI treatment has been found to be associated with good clinical outcome in multiple cancer types[106–109]. TLSs are cell aggregates which are similar in structure to lymph nodes and usually contain lymphocytes, stromal cells, high endothelial cells and DCs[110]. Unlike lymph nodes, TLSs are not present under normal conditions but transiently form in non-lymphoid tissue, usually in the response to infection or chronic inflammation[110,111]. Interestingly, some studies have shown that intra-tumour TLSs have been positively associated with TMB and predicted neoantigen counts, perhaps suggesting TLSs

preferentially form in the presence of a highly immunogenic tumour[112,113]. Mechanistically, intra-tumour TLSs are thought to provide privileged sites for interaction and signalling between immune cells which culminates in lymphocyte activation, proliferation, and differentiation into anti-tumour effectors[114,115]. Usually, intra-tumour TLSs are identified histologically but transcriptional signatures for TLSs have also been identified and were shown to correlate with histological features of TLSs and in some cases improved clinical outcome of cancer patients to ICI treatment[113,114,116–119].

1.2.2.2 Factors extrinsic to the tumour and tumour microenvironment

Despite focus on tumour and TME intrinsic factors, there is a growing body of evidence, from studies of multiple cancer types, that factors extrinsic to the tumour and the TME may play a crucial role in influencing clinical outcome to ICI treatment[120]. Considering that ICI is not targeted to the tumour, and so affects the host immune system globally, it is reasonable to assume that the baseline immune state of a cancer patient would influence their clinical outcome to ICI treatment. Accordingly, several immune related prognostic factors have been reported, some of which are considered to potentially be reflective of the baseline immune state of an individual[42]. The most notable of these are human leukocyte antigen (HLA) type, signatures from peripheral blood immune cells, and development of immune related adverse events (irAEs).

The HLA system is the human MHC and so is crucial for antigen presentation to T cells[121]. Some studies have shown that the HLA type of a cancer patient, which is tumour independent, can influence their outcome to ICI[86]. The HLA types of an individual determine the affinity of binding to and breadth of peptide antigens that their cells can present to T cells[121]. Subsequently, it is thought that certain HLA types may have higher affinity for tumour neoantigen binding, thus enabling improved neoantigen presentation to T cells[86]. Everyone has two alleles for each HLA type which can increase the breadth of antigen binding if the alleles are different[121]. Two studies reported that patients with HLA homozygosity at one or more HLA types had poorer prognosis to ICI treatment, likely because of the

reduced breadth of neoantigens their cells can present[86,122]. An advantage of using HLA type as a biomarker is that it can be determined from the DNA of any cell type and so can routinely be determined from a blood sample[121]. However, like previously discussed prognostic factors, associations of HLA type with patient clinical outcome have not been recapitulated consistently and so further research using multiple large cohorts with carefully selected patient eligibility criteria, for example exclusion of MSI-high patients, is needed[123,124].

As ICI is primarily thought to affect T cells, the most well studied prognostic signatures from peripheral blood immune cells of ICI treated cancer patients are from T cells, either PreTx or early during treatment[120,125–132]. Markers of increased activation and lower exhaustion on CD8+ T cells PreTx, and high TCR clonality during ICI treatment appear to be the most consistently reported biomarkers of clinical outcome across these studies[120,125–128,130–132]. Despite the focus on T cells, studies using in depth characterisations of peripheral immune cells have reported associations of non-T cells with good clinical outcome[133]. These studies of melanoma and non-small cell lung cancer (NSCLC) found PreTx differences in the myeloid and NK cell compartments of patients who clinically benefited from ICI treatment[133–136]. Although further studies of other cancer types are needed, these findings suggest that immune cell types other than T cells are important and characterisation of peripheral immune cells may be useful in assessing the baseline immune activation of a patient prior to ICI treatment.

irAEs are a group of side effects that many cancer patients develop during ICI treatment[42,137,138]. irAEs present as symptoms like those seen in autoimmune diseases and are thought to occur due to breaking of immune tolerance to self-antigen[73,137,138]. Development of one or more irAEs during ICI treatment, particularly early on, is associated with improved clinical outcome in multiple cancer types, including OC[139,140]. Although, it should be noted that two studies in NSCLC reported significantly increased risk of death in patients who developed immune related pneumonitis during ICI treatment[141,142]. Despite promising findings, a key limitation of using irAEs as a prognostic biomarker

of good clinical outcome to ICI is that, by definition, they are only assessable during ICI treatment and not before. There are also cases where the irAEs are sufficiently serious to cease treatment, particularly during combination ICI treatment[143,144]. Hence, there is considerable interest to discover PreTx biomarkers of high risk of irAE development during ICI treatment[145].

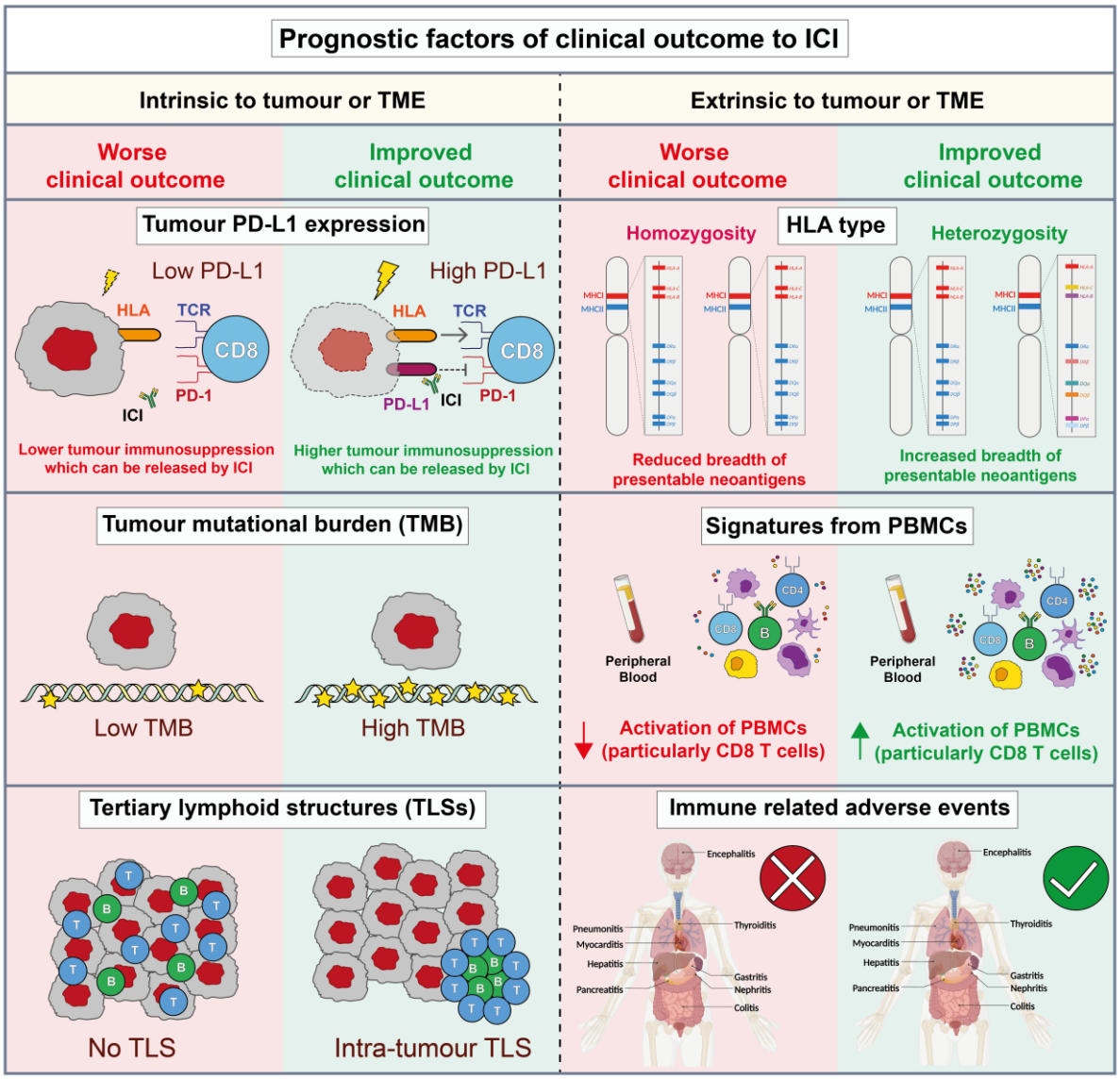


Figure 1.2. A summary of the most notable prognostic factors for good clinical outcome to ICI treatment. (Left) The most notable prognostic factors intrinsic to the tumour or tumour microenvironment (TME) are tumour PD-L1 expression, tumour mutational burden (TMB), and tertiary lymphoid structures (TLSs). Briefly, high tumour PD-L1 expression is thought to indicate high immune checkpoint mediated immune suppression within the tumour which can be relieved by immune checkpoint inhibitor (ICI) treatment. A high TMB is believed to predispose to improved clinical outcome to ICI by increasing the number of neoantigens that a tumour expresses, which can then be targeted by activated immune cells. Intra-tumour TLSs are thought to provide privileged sites for interaction and signalling between immune cells which leads to lymphocyte activation, proliferation, and differentiation into anti-tumour effectors. (Right) The most notable prognostic factors extrinsic to the tumour or TME are human leukocyte antigen (HLA) type, signatures from peripheral blood mononuclear cells (PBMCs), and immune related adverse events (irAEs). Briefly, homozygosity of an individual's HLA type is thought to associate with poor prognosis likely due to the reduced breadth of tumour derived

neoantigens that their cells can present to T cells. Signals of activation from PBMCs, particularly T cells, are thought to associate with improved prognosis by indicating a greater baseline immune activation and capacity of more effective anti-tumour immunity after ICI treatment. Finally, development of irAEs are thought to occur due to breaking of immune tolerance to self-antigens, and so may be an indicator of greater baseline immune activation and potential to overcome immune tolerance to tumour antigens following ICI treatment. This figure was created using some icons from Biorender.com.

Given the similarities in presentation of irAEs and symptoms of autoimmune disease, autoantibodies (AAbs), which are a hallmark of Systemic Lupus Erythamotosous (SLE) and many other autoimmune diseases, and the B cells which produce them have become a recent focus of research in ICI treated cancer[146,147].

1.3 B cells and antibody-mediated immunity

A fundamental aspect of adaptive immunity is production of antibodies by B cells[59]. There are two sub-compartments of B cells, known as B-1 and B-2 cells, which differ in developmental origin and function[148]. In humans, B-1 cells have been found in the umbilical cord blood and adult peripheral blood[149]. These B-1 cells constitutively secrete antibodies which are minimally mutated from germline sequences of antibody genes and are known as natural antibodies[148]. Although these natural antibodies usually have low affinity for antigen, they can often react to commonly conserved motifs such as those found in carbohydrates which enables binding to a wide range of pathogen associated antigens and cellular debris[150–152]. This low affinity binding of an antibody to multiple structurally unrelated antigens is known as polyreactivity[151]. As a result of their polyreactivity, natural antibodies play a vital role in protecting the host from disease[152–156]. B-2 cells are the conventional B cells which originate from common lymphoid progenitor (CLP) cells in the bone marrow (BM)[157]. Unlike B-1 cells, B-2 cells do not constitutively secrete antibodies and instead require help from T cell to produce antibodies with high affinity binding directed towards specific antigens[148]. This antigen driven generation of high affinity antibodies is vital for effective anti-pathogen immune responses that natural antibodies cannot sufficiently provide.

However, in this process of generating high affinity antibodies, the B cell receptor (BCR,) which specifies antigen binding, can acquire affinity for binding to self-antigens[158,159]. B cells expressing BCRs which bind self-antigen are known as autoreactive B cells[158]. To limit autoimmunity, most autoreactive B cells are deleted during the B cell developmental processes[62]. However, some autoreactive B cells evade deletion and can have detrimental effects on the host by mediating autoimmunity[160]. One mechanism by which autoreactive B cells are known to mediate autoimmunity is through production of AAbs[161]. Unsurprisingly given their polyreactivity, natural antibodies are known to bind to self-antigens and so by definition act as AAbs[155]. However, as natural antibodies usually have low affinity, it is instead the self-reactive high affinity antibodies produced by B-2 cells which are most frequently implicated in pathogenic autoimmune disease[161]. B cells and the antibodies they produce have also been shown to influence the human anti-tumour immune response[147].

In this section I will summarise the key steps that are involved in development, maturation and antibody production of B-2 cells; and how immune tolerance mechanisms limit autoreactivity during these processes. I will also briefly describe the known functions of AAbs and B cells, with a particular focus on their roles in influencing anti-tumour immunity in humans.

1.3.1 B cell development and maturation

Development of B-2 cells begins in the BM when CLPs differentiate into pro-B cells which are progenitor cells that are committed to the B cell lineage[162]. From this point onwards there are three key stages in B cell development and maturation. These are BCR V(D)J recombination, maturation in secondary lymphoid organs (SLOs), and generation of B cells with high affinity to antigen in the germinal centre (GC) reaction[162]. Each of these key stages have tolerance checkpoints to limit autoreactivity whilst maximising diversity of the repertoire of antigens that can be recognised. BCR V(D)J recombination, the first of these key stages, takes place in the BM[162].

1.3.1.1 B cell receptor V(D)J recombination

BCRs are membrane bound antibodies which consist of two identical immunoglobulin heavy chains (IgH) and two identical immunoglobulin light chains (IgL) all joined together by disulphide bonds[163,164]. IgH chains are encoded by genes of the *IGH* locus, whereas IgL chains are encoded by either the genes of the *IGK* locus, for Igκ chains, or genes of the *IGL* locus, for Igλ chains. For functional BCR formation, gene rearrangement of the *IGH* gene locus, which occurs first, and either the *IGL* or *IGK* loci is required[163] (**Figure 1.3**).

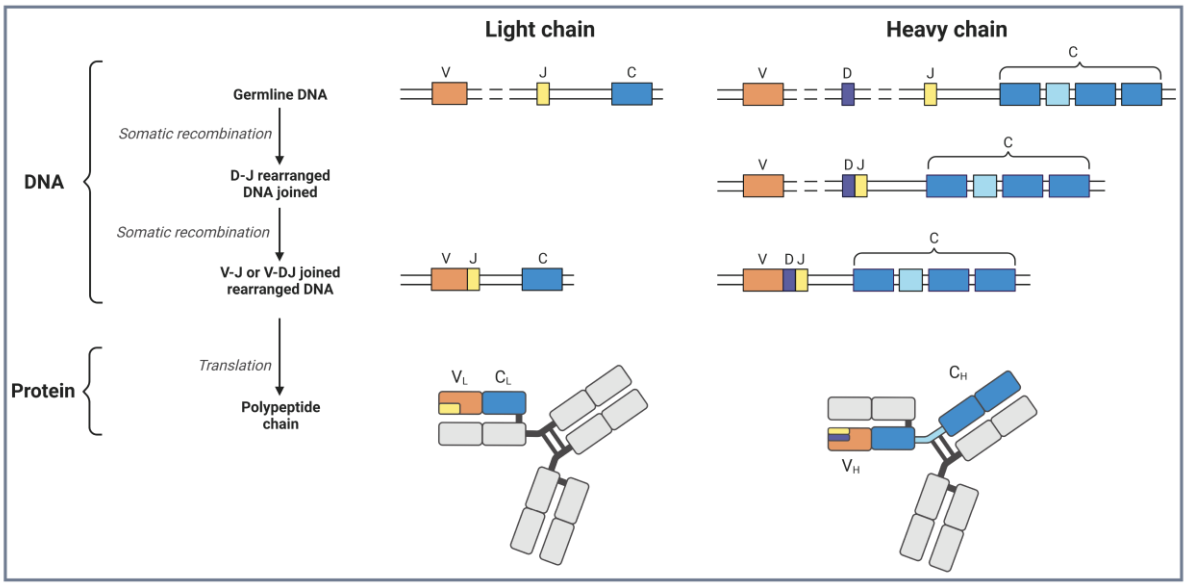


Figure 1.3. Summary of B cell receptor V(D)J recombination. During development, B cells randomly rearrange the gene segments of their *IGL* and *IGH* loci to form a functional B cell receptor (BCR). Briefly, pro-B cells randomly rearrange their variable (V), diversity (D), and joining (J) gene segments of the *IGH* gene locus to form the variable region of their IgH chain, in a process known as somatic recombination. This is two-step process which begins with D-J gene recombination and is followed by V-DJ recombination to produce a functional heavy chain protein. Following successful IgH chain production. Rearrangement of the IgL chain occurs. Unlike the *IGH* gene locus, the *IGL* gene loci do not have any D gene segments so assembly of V and J elements, from either one of the *IGK* or *IGL* loci, is the only recombination step required for formation of an IgL chain. The recombined IgH and IgL chains then come together to form a functional BCR. Created using a Biorender.com template.

At the pro-B cell stage, a combination of *IGH* gene locus accessibility and expression of the Recombination Activating Genes (RAG-1 and RAG-2) initiates pro-B cells to randomly rearrange their variable (V), diversity (D), and joining (J) gene segments of the *IGH* gene locus to form the variable region of their IgH chain[158,165]. This is a two-step process which begins with D-J gene recombination and is followed by V-DJ recombination in pre-B cells to produce a functional heavy chain protein Igμ[163].

Recombined Igμ then associates with surrogate light chains (SLCs) and the Igα and Igβ subunits, which are the signalling portion of the BCR and mediate antigen internalisation, to form the pre-BCR[166–168]. Formation of the pre-BCR initiates allelic exclusion which is the process of halting *IGH* rearrangement to prevent expression of two different IgH chains by the same cell[169]. Signalling through the pre-BCR drives strong B cell proliferation and differentiation which allows for rearrangement of the IgL chain[170–175]. Unlike the *IGH* gene locus, the *IGL* gene loci do not have any D gene segments so assembly of V and J elements, from either one of the *IGK* or *IGL* loci, is the only recombination step required for formation of an IgL chain[163]. *IGK* rearrangement is thought to precede *IGL* rearrangement, with any pre-B cell unable to produce a functional Igκ chain proceeding to attempt to produce a Igλ chain instead[168]. Following downregulation of SLCs, the subsequent rearranged IgL chain comes together with the IgH chain to form a functional BCR, of the IgM isotype, which is expressed on the plasma membrane of immature B cells[158]. It is at this point that these pre-B cells become capable of binding antigen for the first time[158].

Antigen recognition by a BCR is primarily determined by the fragment antigen-binding (Fab) region which is composed of one constant and one V domain from each of the IgH and IgL[163]. In particular, it is subregions of the Fab region, known as complementarity determining regions (CDRs) which mediate antigen contact and are highly variable[163]. CDR1 and CDR2 are encoded by the V domains, whist the CDR3 encompasses the region at which the V, (D) and J domains are joined together, thus making it the most variable CDR[163]. The remaining Fab region is comprised of framework regions (FRs) which are less variable than the CDRs[163]. As the *IGH* gene locus is more complex than the *IGL* or *IGK* gene loci[164], there is much more potential for diversity between different IgH chains. In fact, it has been estimated that greater than 10^{16} different BCRs can be potentially formed through a combination of somatic variation in CDR3 regions, variation in VDJ rearrangement, and association of different combinations of IgH and IgL chains[164].

At this stage, autoreactive immature B cells are eliminated through central tolerance mechanisms, whilst other immature B cells undergo positive selection, in which they cease V(D)J recombination and migrate to SLOs to undergo further development[158,162].

1.3.1.2 B cell maturation in secondary lymphoid organs

SLOs are sites at which adaptive immune cells interact with each other to generate response or tolerance to encountered antigen[176]. SLOs include the spleen, lymph nodes, tonsils, and gut-associated lymphoid tissue (GALT)[176]. It is in SLOs that B-2 cells mature and can eventually receive T cell help to generate T cell dependent antibody responses[176,177].

Once in the SLOs, human immature B-2 cells are thought to differentiate through the transitional stages T1, T2 and T3 into either mature naïve or marginal zone (MZ) B cells[178–180]. However, T1 and T2 B cells have been found in the human BM suggesting that maturation from immature to transitional B cells can take place in the BM before migration to SLOs[181]. As human T2 B cells have also been identified at low frequencies in the peripheral blood, it is possible that T1 and T2 B cells can mature outside of SLOs[182,183].

Progression of B-2 cells from an immature state through T1, T2, and T3 to mature naïve state is characterised by the loss of expression of CD10, CD5 and IgM coupled with increase in expression of CD22, CD40, CD21, CD23 and IgD[178,179,183,184]. Although it is important to note that antigen activated B cells in germinal centres also express CD10, so expression of CD10 is not a definitive marker of immature B cells[178,183,185]. Maturation of T1 B cells to T2 B cells requires Bruton tyrosine kinase (BTK) which is vital for BCR signalling, suggesting that BCR engagement is involved in maturation[184]. The survival of T1 B cells was also shown to be improved in the presence of the cytokine IL-4 or co-culture with BM stromal cells, but not the presence of cell culture supernatant from these stromal cells[182]. This suggests that IL-4 signalling or cell to cell contact is important in survival of T1 B cells and their

maturation into T2 B cells. Due to the dramatic reduction in autoreactivity between T1 and T2 B cells it is thought that a peripheral tolerance mechanism potentially related to BCR signalling exists during this maturation step[184]. Transcriptomic and metabolic analysis of human T1, T2, T3 and mature naïve B cells found that reduced mTORC1 signalling and the subsequent switch in metabolism were crucial for the maturation of T1 and T2 B cells into T3 and mature naïve B cells in both humans and mice[186]. Signalling from the cytokine B cell activating factor (BAFF, also known as BLyS) is also believed to play a key role in development of transitional B cells into mature naïve and MZ B cells[187].

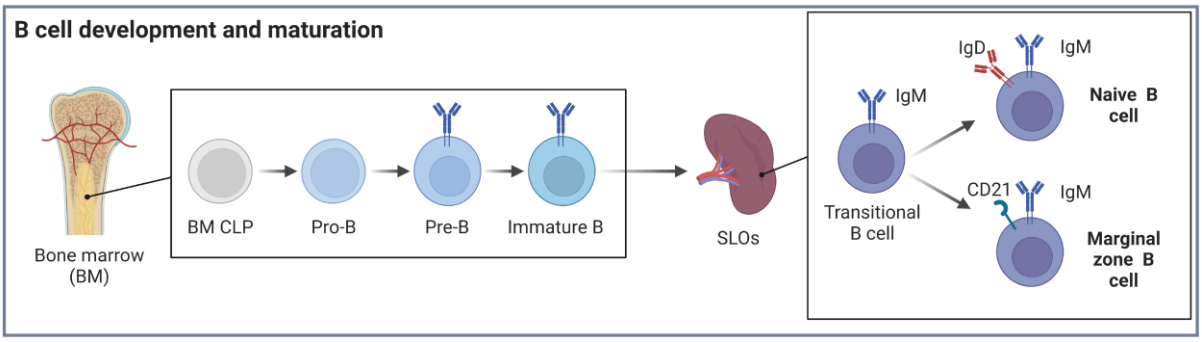


Figure 1.4. Development and maturation of naïve and marginal zone B cells from common lymphoid progenitors. B cells originate from a common lymphoid progenitor (CLP) which is found in the bone marrow (BM) and differentiates into pro-B cells which are progenitor cells that are committed to the B cell lineage**.** Pro-B cells undergo V(D)J recombination of their *IGH* to express a functional heavy chain protein Igμ which then associates with surrogate light chains SLCs and the Igα and Igβ subunits, to form the pre-BCR, completing their differentiation into pre-B cells. Signalling through the pre-BCR in pre-B cells drives proliferation and V(D)J recombination of the *IGL*, resulting in formation of a functional IgL chain which comes together with the IgH chain to form a functional BCR. Expression of a functional BCR completes the differentiation of pre-B cells into immature B cells. Migration of immature B cells from the BM to secondary lymphoid organs (SLOs), such as the spleen, signals their differentiation into transitional B cells. Once in SLOs, transitional B cells can differentiate into one of two distinct B cell subsets, mature naïve or marginal zone (MZ) B cells. However, it is unclear how differentiation of transitional B cells into mature naïve and MZ B cells is controlled. Naïve B cells then interact with T follicular helper cells and differentiate into either extrafollicular plasma cells, germinal centre B cells, or memory B cells. In contrast, MZ B cells are vital for T cell independent antibody responses, particularly those against encapsulated bacteria. Created using a Biorender.com template.

In humans, it is unclear exactly how differentiation of transitional B cells into mature naïve and MZ B cells is controlled[188]. One hypothesis is that the T3 state is further divided into two different precursor populations, previously referred to as T3 and T3', based on their surface expression of CD45RB^{MEM55}, IgM and CD27[178,189,190]. T3 B cells expressing CD45RB^{MEM55} and IgM but lacking CD27 expression were proposed to be precursors to MZ B cells which can mature into MZ B cells following NOTCH

signalling[190,191]. A more recent study proposed that the bifurcation in the B cell development pathway occurs earlier at the T1 to T2 B cell stage with IgM^{hi} T2 B cells sharing many transcriptional and phenotypical features with MZ B cells, unlike IgM^{lo} T2 B cells which may be the precursors of mature naïve B cells[192]. Further work from the same research group identified two subsets of MZ B cells, called MZB-1 and MZB-2, primarily based on greater expression of the chemokine receptor CCR7 in the MZB-1 subset[193]. MZB-1 and MZB-2 were suggested to not be clonally related and so were proposed to have unrelated developmental origins[193]. Key differences between the two subsets were that MZB-1 cells were found to more abundant and more somatically mutated than MZB-2 cells, with MZB-2 showing little evidence of SHM[193]. Although MZ B cells are B-2 cells and can show features of involvement in T cell dependent antibody such as mutations in IgH V genes, like B-1 cells, they are vital for T cell independent antibody responses, particularly those against encapsulated bacteria[156,191,193–196]. Like transitional B cells, human MZ B cells are not restricted to the spleen and have also been detected in the peripheral blood and other SLOs suggesting that they circulate the body looking for antigen[194–196].

For mature naïve B cells, the current model suggests that their activation begins when they encounter their cognate antigen which is bound by APCs in B cell follicles of SLOs[58,59,197,198]. However, BCR signalling alone is not sufficient stimulation for the B cell to survive, proliferate and differentiate[199]. A second stimulatory signal is required and can be provided by TLR activation or T cell mediated co-stimulation[199,200]. To receive T cell mediated co-stimulation, BCR bound antigen is internalised, processed, and presented on MHC class II to Tfh cells which are at the border of the T cell zone[59]. The subsequent interaction with Tfh cells leads to one of three main developmental possibilities for the B cell[59,201–203]. One possibility is differentiation into short lived extrafollicular plasma cells[203]. Like MZ B cells, these are antibody secreting cells (ASCs) which have minimally somatically mutated Ig genes but are important for rapid initial response to antigen[204]. Another developmental possibility is differentiation into GC B cells and subsequent entrance into a GC[203]. Once in a GC, GC B cells can differentiation into memory B cells and plasma cells that express BCRs with high affinity for the

cognate antigen[203]. Both of these cell types are crucial for effective and long-lasting antibody responses[203]. Plasma cells are the main ASCs, whereas memory B cells provide immunological memory and can clonally expand and then differentiate into plasma cells following re-encounter and stimulation with the cognate antigen[205]. The final possibility is differentiation into memory B cells independent of GC formation, which circulate the body but can later enter a GC reaction[201,203,205].

1.3.1.3 The germinal centre reaction

GCs can form in the B cell zone of SLOs in response to antigen[177]. GC B cells undergo rapid proliferation and antibody diversification[177]. Then can differentiate into ASCs such as memory B cells and plasma cells. GCs are split into a dark zone (DZ), where B cells proliferate and undergo somatic hypermutation (SHM), and a light zone (LZ), where B cells are selected for antigen binding affinity[177]. This process of SHM and selection for antigen binding affinity is known as affinity maturation[177].

GC B cells enter the DZ of the GC and become DZ B cells, also known as centroblasts, which proliferate rapidly and undergo SHM of their IgV region genes to revise their BCR and alter binding affinity for cognate antigen[185,206]. During SHM, the enzyme Activation Induced Cytosine Deaminase (AICDA) generates random mutations in the variable domains of the Ig genes[177,207]. Upregulated expression of the transcription factor Bcl-6 in GC B cells is vital at this stage as it allows for toleration of DNA damage induced by SHM[177]. Following proliferation and SHM, DZ B cells then enter the GC LZ and become LZ B cells, also known as centrocytes[177,206]. LZ B cells compete with each other for binding to antigen, which is sequestered on follicular DCs (FDCs), and help from Tfh cells which are present in the LZ[208]. LZ B cells that have the BCRs with the highest affinity binding to antigen survive and differentiate into memory B cells and plasma cells. To generate BCRs with high antigen binding affinity, GC B cells can cycle between the DZ and the LZ to undergo repeated rounds of SHM and affinity binding selection[208,209]. During this process, LZ B cells which do not bind antigen with sufficiently high affinity and do not receive sufficient help from Tfh cells die via apoptosis[208,210].

It had been proposed that the LZ of GCs is also where class switch recombination (CSR) of activated LZ B cells occurs[207]. However, it has recently been shown that CSR of B cells primarily occurs prior to differentiation into GC B cells[211]. During CSR, activated IgD or IgM expressing B cells excise either the μ or δ constant region from the IgH locus so that they can be replaced by any one of the four γ, two α, or one ε constant region genes[212]. CSR of a B cells to a particular antibody class is initiated by the specific combinations of co-stimulation and cytokine signalling that it receives[213–220]. This signalling directs CSR into different isotypes by activating sterile germline transcripts that open up switch regions and make targets for AICDA[218]. CSR is a crucial step in B cell maturation as it allows B cells to alter their effector function but retain their antigen specificity[218].The different isotype classes of Ig are described in more detail in section 1.3.1.

Due to their association with GCs, V gene mutation, memory B cell development, and long-lived antibody responses are considered signs of T cell dependent antibody responses.

1.3.2 Immunological tolerance in the B cell compartment

Immunological tolerance of B cells is essential to limit autoreactivity whilst also preserving the capacity of reacting to foreign insults such as pathogens[62,221]. Tolerance is enforced by multiple mechanisms, known as checkpoints, which endeavour to reduce either the affinity for self-antigen, the functionality, or the frequency of autoreactive B cells within the B cell repertoire[158,222]. These checkpoints are classified as either central or peripheral, depending on the stage of B cell development and location in which they occur[62].

Central B cell tolerance encompasses the first checkpoints which negatively select newly generated IgM expressing immature B cells in the BM, primarily through mechanisms such as clonal deletion, and receptor editing[158,223–227]. Clonal deletion is apoptosis induced in immature B cells following high affinity ligation of the IgM BCR by self-antigen in the BM[158,225]. It has been suggested that clonal

deletion is the primary mechanism of central tolerance and results in the death of over 85% of immature B cells in the BM[228]. Interestingly, clonal deletion does not occur if the self-antigen is not membrane bound[229]. Rather than apoptosis, recognition of soluble self-antigen instead permits survival of the B cell, although peripheral tolerance mechanisms will likely render it anergic[229]. The other primary mechanism of central tolerance is receptor editing. Receptor editing occurs following crosslinking of IgM by membrane bound autoantigen[223]. Following this, RAG-1 and RAG-2 mediated secondary recombination of the IgL V genes occurs to replace the original IgL chain and thus alter the antigen specificity of a BCR[223,224]. Unlike clonal deletion, B cells that undergo receptor editing are not deleted and remain in the B cell pool, providing they succeed in generating a new non-autoreactive BCR through IgL recombination[158,223,224]. One study demonstrated that receptor editing is a common process that contributes approximately 25% of antibodies in the antibody repertoire[230].

Despite central tolerance mechanisms, autoreactive B cells are still present in the transitional and naïve compartments, with up to 30% of the mature naïve B cell pool estimated to be autoreactive[62]. To limit the potential detrimental effects of autoreactive B cells in these mature B cell compartments, further tolerance mechanisms exist outside of the BM[62]. These are collectively known as peripheral tolerance.

Peripheral tolerance of B cells aims to prevent autoreactive B cells from undergoing affinity maturation and at early stages is primarily enforced through anergy, follicular exclusion and receptor dilution[62]. Anergy refers to B cells which have been rendered functionally inactivated despite expression of BCR and presence of cognate antigen[226,227]. Anergic B cells do not respond efficiently to Tfh interaction, do not proliferate after BCR binding, and rely on outcompeting other B cells for BAFF signalling to survive[231,232]. However, anergy is thought to be largely reversible due to the potential need for tolerised self-reactive B cells to react against pathogen antigens with similar epitopes[158,233]. Follicular exclusion is the prevention of autoreactive B cells from entering GCs[234,235]. Follicular

exclusion is mediated through downregulation of the chemokine receptor CXCR5, which usually functions to home mature B cells to B cell follicles which are rich in its ligand CXCL13, and maintained expression of CCR7, which homes B cells to the T cell zone[236]. This results in sequestering of autoreactive B cells at the border between the B cell and T cell zones, at which they undergo apoptosis[234,236]. Receptor dilution is the expression of both a Igλ and Igκ chain by a single B cell[62,237–239]. This functions to reduce autoreactivity by either diluting the autoreactive BCR pairing of the original light chain with the heavy chain, or by inducing sequestration of the original light chain which is then replaced completely in the BCR pairing by the new light chain[62].

Despite early peripheral tolerance mechanisms, some autoreactive B cells escape and enter GC reactions[240]. This, in combination with non-autoreactive B cells potentially acquiring autoreactivity via SHM in the GC DZ, poses a risk for development of high affinity autoreactive ASCs[159]. However, there are GC based peripheral tolerance mechanisms which exist to limit their development[62,240]. These include receptor revision, FAS-mediated apoptosis, FDC and FcγRIIB induced inhibition, and suppression by regulatory T cell subsets[62]. Receptor revision is the process in which autoreactive GC B cells undergo SHM to reduce binding to self-antigens whilst potentially increasing binding to foreign antigens[241]. If successful, these B cells with revised BCRs reduce their autoreactivity and can continue their development into ASCs[241]. Fas mediated apoptosis functions like clonal deletion, with signalling through the Fas death receptor resulting in death of the autoreactive B cell unless strong survival signals induce high expression of anti-apoptotic Bcl-2[242,243]. Autoreactive B cells can also be inhibited and eliminated following interaction with FDCs[244,245]. Binding to self-antigen sequestered on FDCs and signalling through the inhibitory FcR FcγRIIB is thought to regulate autoreactive GC B cells, potentially by inducing some of them into an anergic like state[244–246]. Finally, regulatory T cells (Tregs) and T follicular regulatory (Tfr) cells have been shown to directly supress GC B cells, via mechanisms including CTLA-4 signalling[247]. GC B cells, whether autoreactive or not, which pass through or escape all of these tolerance checkpoints become terminally differentiated ASCs[62].

1.3.3 Antibody isotypes and function

Upon activation, ASCs can produce antibodies which can have diverse effects on the host immune system. Although the combination of the V(D)J domains of the IgH and IgL chains confer antigen specificity, the effector function of an antibody depends largely on the antibody isotype class[164,248–250] (**Figure 1.5**). The isotype class of an antibody is defined by the Fc fragment of the constant (C) domain of its IgH chain, which is encoded by any one of nine genes downstream of the VDJ loci[248]. Alternative splicing of each of these nine IgH C genes allows for production of either membrane bound Ig chains as the BCR, or soluble secreted Ig in the form of antibodies[164,248]. The nine IgH C genes each confers a subclass of Ig which are grouped into five major Ig isotypes: IgM, IgD, IgG, IgA, and IgE[164,248,250].



Figure 1.5. Effector functions of antibodies by isotype class. The effector function of antibodies depends largely on its isotype. There are five major isotypes, IgD, IgE, IgM, IgA, and IgG, with IgG composed of four subclasses, IgG1, IgG2, IgG3, and IgG4, which differ in their potency of inducing effector functions**.** Created using a Biorender.com template.

1.3.3.1 Antibody isotypes

As mentioned earlier, IgM is the first Ig expressed during development of a B cell. Immature and naïve B cells express membrane bound monomeric IgM as their BCR [164,251]. Following further maturation and cognate antigen binding, IgM can be secreted predominantly as a pentamer that is composed of multiple IgM monomers joined by disulphide bonds and the polypeptide J Chain[164,249]. As IgM

antibodies are produced by immature and naïve B cells, they are usually encoded by V regions that have undergone no, or little, SHM[249,251]. Although this means that the repertoire of IgM antibodies is more restricted than IgG and other Ig isotypes, IgM antibodies play a vital role in protection from pathogenic infection[150,249,251]. This is due to IgM antibodies being the first antibody class to be produced and their contribution to most of the natural antibodies which appear in the absence of antigen stimulation or T cell help[150,251–253]. Due to their multimeric structure, natural antibodies of the IgM isotype can achieve high avidity through multimeric interactions with the antigen[164,249]. Pentameric IgM primarily functions by coating antigen, known as opsonisation, which targets the antigen for destruction by other immune cells or complement activation[164,249,254]. However, due to the pentameric nature of IgM, prevents it from permeating into tissues and results in a distribution largely restricted to the blood[249,255].

Circulating IgD antibodies are very uncommon, and until recently their function was unclear[164,256]. Secreted IgD was found to bind to the CD44-binding protein galectin-9 which is expressed on basophils[256]. Following IgD binding, basophils increased their secretion of the cytokines IL-4, IL-5, and IL-13 which leads stimulated Tfh cells to drive CSR mediated IgG and IgE antibody responses by B cells to the initial antigen bound by the secreted IgD[256]. This same study also found that IgD binding reduced basophil degranulation usually initiated by IgE binding[256].

IgG makes up over half of the total circulating Ig in blood and consist of four subclasses, IgG1, IgG2, IgG3, and IgG4, which were named in decreasing order of abundance[257]. Each IgG subclass differs in its combination of typical antigen targets and effector functions[164,255,257]. IgG1 antibody responses are mainly induced by soluble, and membrane bound proteins[257]. Following antigen binding, IgG1 can elicit an array of effector functions through direct neutralisation, activation of the complement cascade, and binding to Fc receptors (FcRs) on immune cells[60,250]. IgG2 is the IgG subclass which is predominantly responsible for binding to polysaccharide antigens expressed by bacteria and

viruses[258,259]. The importance of IgG2 antibodies for protection against pathogenic bacteria has been demonstrated by poor antibody responses and susceptibility to recurrent infections to *Haemophilus influenzae* in IgG2 deficient children[259]. IgG2 binds to FcRs less strongly than IgG1 and so primarily functions through direct neutralisation and complement activation[60,260,261]. Like IgG1, IgG3 responses are mainly induced by protein antigens[164]. IgG3 also has similar effector functions to IgG1, but IgG3 has increased capacity to activate complement following antigen binding due to a greater binding affinity for the complement component C1q[60,164,260]. IgG4 is predominantly associated with responses to allergens and other antigens which the host is repeatedly exposed to in a non-infectious setting[257]. IgG4 is the only subclass of IgG which is unable to activate complement and, like IgG2, binds to FcRs with low affinity, and so is limited in its potential to promote effector functions[60,260,261].

IgA is the most abundant antibody isotype class in secretions and is mostly located at mucosal surfaces, including the oesophagus. IgA is predominantly monomeric in the blood plasma, but in secretions it is usually dimeric with J chain joining the two monomers together[262]. In humans, there are two subclasses of IgA, IgA1 and IgA2, with IgA1 contributing to most of the circulating IgA pool, and IgA2 more abundantly expressed in the intestine and genitourinary tracts[262]. Both T cell dependent and T cell independent mechanisms of class switching to IgA have been described, with T cell independent mechanisms believed to exist to rapidly produce IgA which can provide some level of anti-pathogen immune response whilst high affinity IgA antibodies are generated[262]. Unlike IgG, IgA is not a strong activator of complement or phagocytosis, and so effector functions are primarily mediated via neutralisation[249]. This is believed to be because IgA is secreted mainly in mucosal surfaces in which complement and phagocytes are not as prevalent compared to the sites in which IgG is most commonly present[249].

IgE antibodies are very lowly abundant in blood and are associated with immune responses to parasites, allergy, and hypersensitivity reactions[164,263]. IgE functions by binding to the FcR FcεRI which

is expressed on mast cells and other innate immune cells[263]. Crosslinking of antigen to the FcεRI-bound IgE initiates degranulation of the bound cell and release of immune signalling molecules such as histamine, which subsequently promotes a T helper 2 (Th2) response[263]. Despite being canonically associated with allergy, other research has suggested that IgE may also have additional functions including activation of plasmacytoid DCs (pDCs) and recruitment of monocytes to the TME[263–265].

1.3.3.2 Antibody functions

As briefly touched on , antibodies play a crucial role in protecting the host from infection and disease through their various effector functions which can be largely grouped into neutralisation, activation of complement, and binding to FcRs on immune cells[60,164,250].

Antibody mediated neutralisation is binding of antibody to an antigen, usually a toxin or pathogen, and inhibiting its function without requiring the presence of effector immune cells[250]. Examples include reducing the adherence capability of *Streptococcus pneumoniae* and preventing entry of viruses into cells[266,267]. In addition to toxins and pathogens, neutralisation of cytokines by anti-cytokine autoantibodies (ACAAs) has also been shown to occur in humans and possibly functions to modulate the host immune response[268–270].

The other primary mechanism by which antibodies exhibit an effector function that is independent of immune cells is activation of complement[60,164,249,250]. Binding of C1q, the first component of the complement pathway, to the Fc regions of antigen-antibody complexes leads to activation of the complement cascade[164,249,250,271]. This ultimate results in opsonisation and destruction of the target cell via formation of the membrane attack complex[249,250,271]. Although it can destroy a target cell independently, complement activation can also recruit immune cells to the site of activation, which can phagocytose the opsonised target cell[164,249,250,271].

The other main antibody effector functions are all dependent on interaction with FcRs on immune cells[250]. The two primary examples are antibody dependent cell cytotoxicity (ADCC) and antibody dependent cellular phagocytosis (ADCP)[250]. ADCC is a process where antibody binding to membrane bound antigen targets the expressing cell for destruction by effector cells, such as NK cells[250,272]. Crosslinking of FcRs on the surface of the effector cells induces either release of cytotoxic granules containing perforin and granzymes, Fas signalling, or production of reactive oxygen species[272]. ADCP is a similar process to ADCC. Antibody binding to membrane bound antigen targets the expressing cell for engulfment by phagocytic myeloid cells[250,273]. Following phagocytosis of the antibody-bound cell, some myeloid cells can elicit further immune activation either by presenting peptides from the engulfed cell as antigen to T cells via MHC class II, or by producing pro-inflammatory mediators[250,274]. Both ADCC and ADCP are important effector functions for killing of virally infected host cells and tumour cells[250,272–274].

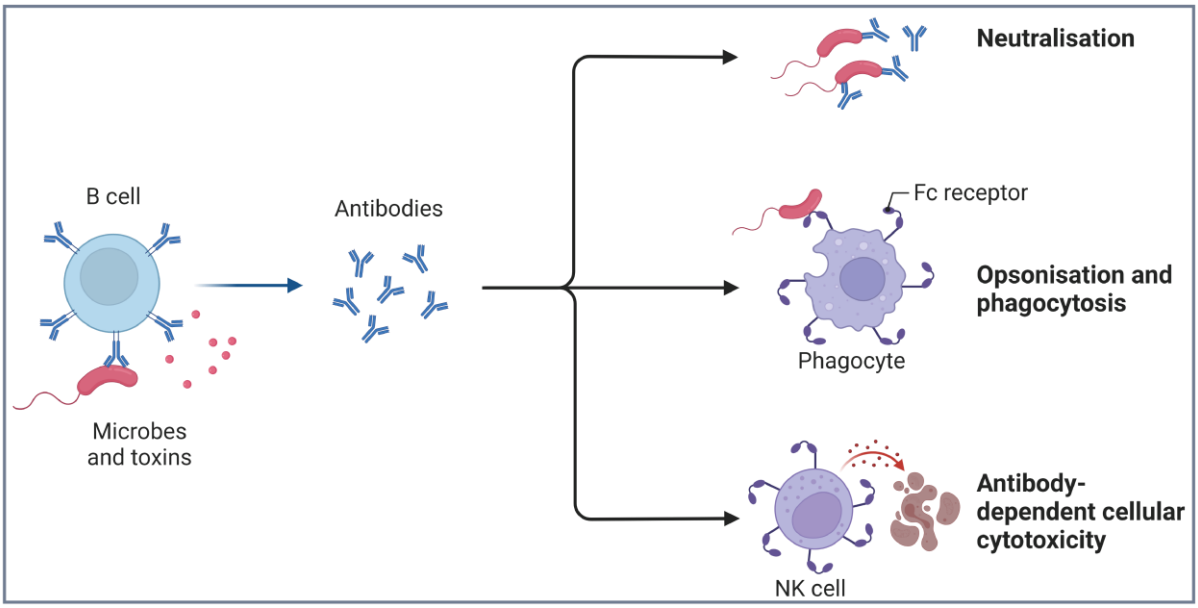


Figure 1.6. Neutralisation, opsonisation, phagocytosis, and antibody-dependent cellular cytotoxicity are some of the major effector functions of antibodies. Antibodies play a crucial role in protecting the host from infection and disease through their various effector functions. One example is neutralisation, which is binding of antibody to an antigen, usually a toxin or pathogen, and inhibiting its function without requiring the presence of effector immune cells. The two major effector functions which require the presence of effector immune cells are phagocytosis of opsonised target cells, and antibody dependent cellular cytotoxicity (ADCC). Opsonisation is the binding of antibodies to a target cell to mark if for engulfment by phagocytic myeloid cells. In contrast, ADCC is a process where antibody binding to membrane bound antigen targets the expressing cell, such as a virally infected or malignant cell, for destruction by natural killer (NK) cells. Created using a Biorender.com template.

1.3.4 B cells and antibodies in human cancer

There is a growing body of research which shows that tumour infiltrating B cells are associated with the clinical outcome of cancer patients[146,275–280]. Although many studies identify a positive association of tumour infiltrating B cells with prognosis, other studies show no association or even a negative association[275,278]. This has been proposed to be due to the difference in factors including the subset of tumour infiltrating B cell, their localisation regarding TLSs and other tumour infiltrating immune cell types, and the isotype of tumour-binding antibodies they produce[275–278]. Although many studies have been published on this subject, and anti- and pro-tumour mechanisms have been proposed, our understanding of the processes by which tumour infiltrating B cells and antibodies modulate the human anti-tumour immune response remains incomplete[147,275,276,281,282] (**Figure 1.7**).

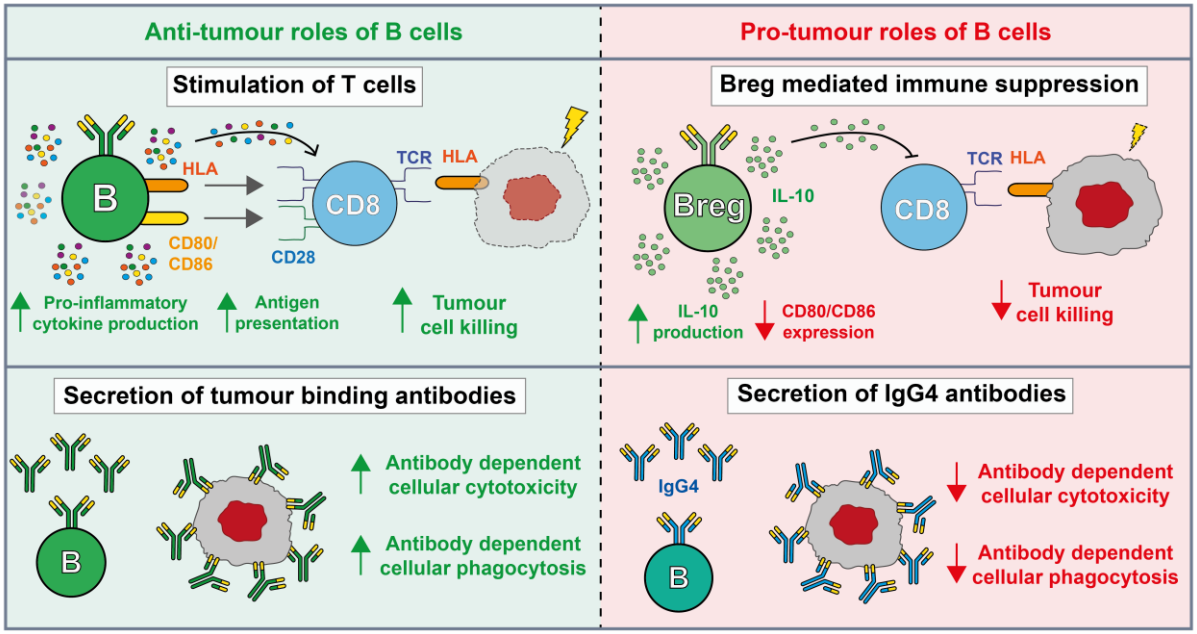


Figure 1.7. B cells can have anti- and pro-tumour functions. (Left) The most notable anti-tumour roles of B cells are stimulation of T cells, and secretion of tumour-binding antibodies. Briefly, B cells can stimulate T cells through proinflammatory cytokine production and presentation of tumour associated antigens or neoantigens, which can promote T cell mediated tumour killing. B cells can also promote anti-tumour immunity through secretion of tumour-binding antibodies. Binding of these antibodies to tumour cells can enhance tumour cell killing via effector functions like antibody dependent cellular cytotoxicity (ADCC) and antibody dependent phagocytosis (ADCP). (Right) The most notable pro-tumour roles of B cells are regulatory B cell (Breg) mediated immune suppression, and secretion of IgG4 antibodies. Briefly, Bregs can elicit a pro-tumour role by functioning to suppress T cell activation and proliferation, primarily through production of the anti-inflammatory cytokine IL-10, and low expression of the co-stimulatory molecules CD80 and CD86 which are required for T cell co-stimulation. Finally, secretion of tumour-binding IgG4 antibodies by B cells is thought to elicit a pro-tumour function due to the poor affinity of IgG4 antibodies for FcRs. Consequently, tumour-binding IgG4 antibodies are

thought to block binding of tumour-binding antibodies of other Ig isotypes, thus reducing tumour killing by ADCC and ADCP. This figure was created using some icons from Biorender.com.

1.3.4.1 Anti-tumour roles of B cells and antibodies

Proposed anti-tumour mechanisms of tumour infiltrating B cells include production of proinflammatory cytokines, presentation of antigen to T cells, and secretion of tumour-binding antibodies[147,275,276,278,283]. Proinflammatory cytokine production and presentation of tumour associated antigens or neoantigens by tumour infiltrating B cells has been suggested to promote T cell mediated tumour killing[283–285]. However, the main mechanism by which tumour infiltrating B cells are thought to promote anti-tumour immunity is through GC driven production of tumour-binding antibodies[109,286–288]. Binding of these antibodies to tumour cells can enhance tumour cell killing via FcR mediated effector functions like ADCC and ADCP[147,276,289]. Although it should be noted that this requires the presence of NK cells or other effector immune cells which may be excluded from the TME[147]. Immune complexes of tumour-binding antibodies have also been proposed to indirectly enhance antigen presentation following their internalisation by DCs[147]. A recent study of ovarian cancer proposed a novel mechanism in which tumour infiltrating B cell derived polyclonal IgA inhibits growth of tumour cells and promotes their T cell mediated killing via transcytosis[290].

1.3.4.2 Pro-tumour roles of B cells and antibodies

The primary proposed pro-tumour mechanism of tumour infiltrating B cells is regulatory B cell (Breg) mediated immune suppression[147,275,281,291]. Bregs, which are phenotypically defined by expression of the anti-inflammatory cytokine IL-10 and their ability to suppress T cell activation and proliferation, have been proposed to infiltrate the TME and inhibit the anti-tumour immune response[61,281,291,292]. However, studies of Bregs in human cancer are challenging because, unlike some other cell subsets such as Tregs, Bregs are not defined by the expression of a specific surface marker or transcription factor and are present in many B cell compartments[281]. Therefore, although tumour infiltrating Bregs can be detected in human tumours, it is difficult to specifically isolate them and establish their

mechanism of immune suppression[275,281,291,293–296]. One mechanism was proposed by a study of hepatocellular carcinoma (HCC) which identified a tumour infiltrating Breg population that expressed high levels of PD-1[292]. Subsequent isolation and *ex vivo* culture of tumour infiltrating B cells revealed that they had capacity to produce IL-10 and suppress T cell cytotoxicity[292]. They also showed that B cells with the same PD-1 high phenotype could be induced from primary peripheral blood B cells by culture with supernatant from primary HCC cells[292]. A similar mechanism was suggested by a study of ovarian cancer[297]. In this study, culture of B cells from the ascites of ovarian cancer patients were found to suppress production of the pro-inflammatory cytokine interferon-γ by autologous CD8+ T cells[297]. The mechanism of this suppression was revealed to be in part due to via production of IL-10 and low CD80/CD86 expression by the B cells[297].

Tumour infiltrating B cells expressing IgG4 are another B cell subset which have been suggested to promote tumour growth, primarily by production of pro-angiogenic cytokines and tumour-binding IgG4 antibodies[298,299]. Mechanistically, tumour-binding IgG4 antibodies are thought to compete with tumour-binding antibodies of other Ig isotypes, such as IgG1 and IgG3, for binding to antigen expressed on tumour cells. Due to the poor affinity of IgG4 antibodies for FcRs, and their much weaker capacity to elicit ADCC and ADCP, their binding to tumour cells results in reduced FcR mediated tumour killing[299].

Aside from Bregs and IgG4 expressing B cells, additional studies have shown that tumour infiltrating B cells can promote tumour growth by inducing a suppressive phenotype in CD4+ T cells *ex vivo* via antigen presentation, or production of the growth factor IGF-1[284,300].

1.3.4.3 Applications of autoantibody detection in cancer

Given their proposed role in anti-tumour immunity, antibodies against tumour associated antigens (TAAs) have been well studied in multiple cancer types, primarily with a focus on cancer diagnosis and prognoisis[301–304].

Numerous studies have shown the potential of using AAbs, either alone or in a small panel, to non-invasively distinguish patients with cancer from those without cancer, although often with limited sensitivity[303–309]. Particularly promising results were seen from a recent study which screened 150 lung cancer samples against 154 TAAs to yield a model based on a panel of 7 AAbs which showed a sensitivity of close to 90%[307]. Amongst AAbs to TAAs, AAbs that bind to the tumour protein p53 (anti-p53 AAbs) are one of the most well studied[305,308]. Despite promise as a diagnostic biomarker of cancer, there have been conflicting reports of the association of anti-p53 AAbs with clinical and prognostic factors of cancer patients, namely cancer stage and patient survival[310–312].

AAbs against non-TAAs have also shown interesting associations with cancer. For example, development of cancer in dermatomyositis (DM) patients was associated with seroprevalence of AAbs against transcriptional intermediary factor-1-γ (TIF-1γ, also known as TRIM33)[313]. However, TIF-1γ AAb seropositive DM patients who were also seropositive for AAbs against cell division cycle and apoptosis regulator protein 1 (CCAR1) had reduced risk of cancer development[313]. Similarly, although patients with systemic sclerosis (SSc) have an increased risk of cancer development, SSc patients with anti-centromere AAbs had reduced risk of developing cancer compared to anti-centromere seronegative SSc patients[314].

Other studies have shown that AAbs may also be associated with risk of developing irAEs during ICI treatment[315–318]. For example, a recent study of multiple cancer types evaluated the association of irAE development with the presence of anti-nuclear autoantibodies (ANAs) in cancer patients both

PreTx and after ICI treatment[318]. Interestingly, this study showed that the presence of ANAs PreTx was not linked with development of irAEs, but patients who showed new ANA patterns at any time after beginning ICI treatment showed higher rates of development of severe irAEs[318]. Two recent studies of panels of AAbs in melanoma patients showed similar results[316,317]. The first of these studies showed that development of an AAb response against one of twenty-three common autoantigens during ICI treatment was associated with increased risk of irAE development, whilst the second study used a microarray of 120 autoantigens to show that patients with a greater fold change in AAb concentration over six weeks of ICI treatment had an increased risk of developing an organ-specific irAE[316,317]. Together these studies suggest that development or changes in AAb responses during ICI treatment may predict the risk of developing an irAE, and its severity.

1.4 LUD2015-005 clinical trial of ICI treated oesophageal cancer

In this thesis I will utilise valuable peripheral blood and tissue samples collected from ICI treated OC patients as part of the LUD2015-005 clinical trial study (ClinicalTrials.gov Identifier: NCT02735239).The LUD2015-005 study is a phase I/II study of ICI in combination with SOC treatment in previously untreated OC[319]. The study is divided into two arms. An inoperable arm (cohorts A1, A2, and B, 38 patients in total) in which patients received treatment of ICI+CTX, and an operable arm (Cohorts C, C-FLOT, D, and D2, 34 patients in total) in which patients received treatment of ICI+CTX +/- radiotherapy followed by resective surgery[319]. A notably unique aspect of the LUD2015-005 study is the initial four-week window of ICI-only treatment prior to the combined ICI+CTX treatment. This provides a rare opportunity for early assessment of ICI response without the confounding influence of concurrent CTX treatment.

This thesis will focus on the inoperable arm of the LUD2015-005 clinical trial as the clinical data for these patients are more matured which allows for more robust assignment of patients into those who showed good clinical outcome and those who did not.

In the inoperable arm, all patients received ICI alone for four weeks in the form of anti-PD-L1 (Durvalumab) only (n=12), anti-PD-L1 plus a single 37.5mg dose of anti-CTLA-4 (Tremelimumab) (n=5), or anti-PD-L1 plus a single dose of 75mg Tremelimumab (n=21). Following this, patients continued anti-PD-L1 treatment and also received SOC CTX (Capecitabine and Oxaliplatin) for up to 18 weeks[319]. $2mm^3$ biopsies from tumour and non-malignant control tissues were collected via endoscopy at up to three timepoints throughout treatment: prior to treatment (Pre-treatment, PreTx), after the four-week ICI-only treatment window (ICI-4W), and at the end of the ICI+CTX treatment regime (Post-treatment) (**Figure 1.8**). Molecular profiling of these tissue biopsies was performed using bulk RNA-sequencing (RNA-seq), single-cell RNA-seq (scRNA-seq), and WGS (**Figure 1.8B**). Whole blood samples were also sampled at these same three timepoints throughout treatment, as well as at an additional timepoint during the third week of ICI+CTX treatment (C1W3). These whole blood samples were separated into blood plasma, for antibody profiling, and PBMCs, for immunophenotyping by spectral flow cytometry (**Figure 1.8C**). Full details regarding the treatment that participants received, and the clinical samples that were collected are described in chapter 2 section 2.1.

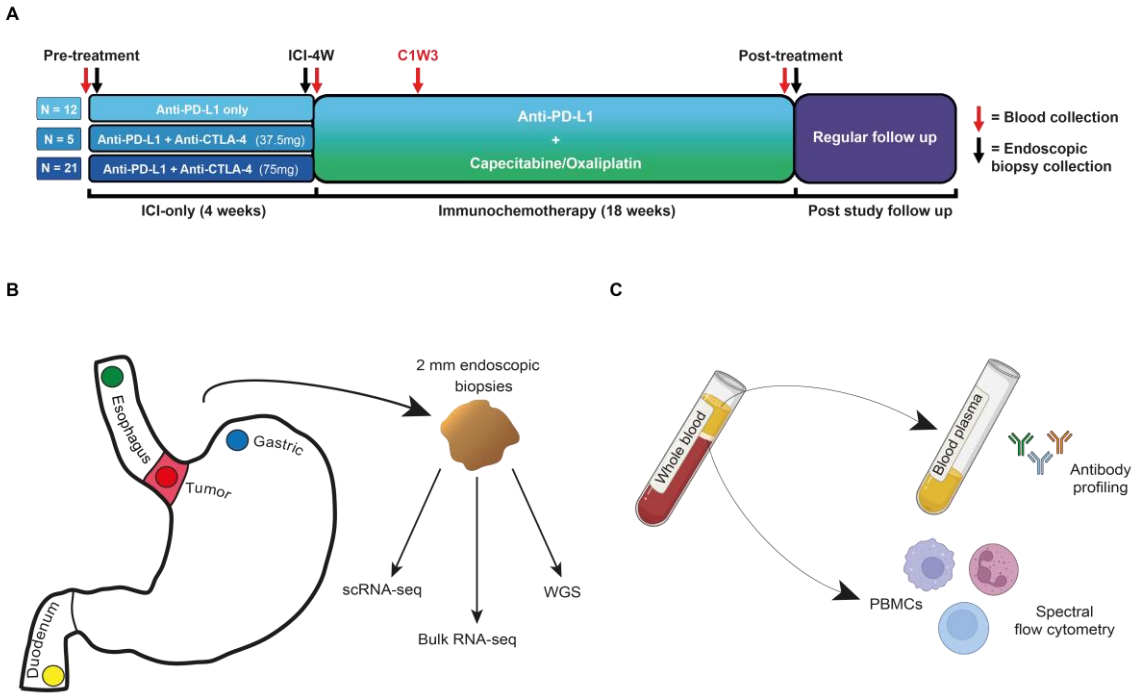


Figure 1.8. LUD2015-005 treatment regime and sampling for inoperable arm. (**A**) A timeline showing the treatment regime that the patients received. All patients received ICI alone for four weeks in the form of anti-PD-L1 (Durvalumab) only (n=12), anti-PD-L1 plus a single 37.5mg dose of anti-CTLA4 (Tremelimumab) (n=5), or anti-PD-L1 plus a single dose of 75mg Tremelimumab (n=21). Following this, patients continued anti-PD-L1 treatment and also received standard of care (SOC) CTX (Capecitabine and Oxaliplatin) for up to 18 weeks. Tissue biopsies were collected via endoscopy prior to treatment (Pre-treatment), after the four weeks of ICI-only (ICI-4W), and after the 18 weeks of ICI+CTX (Post-treatment). Whole blood was collected at the same three treatment timepoints and additionally during the third week of ICI+CTX treatment (C1W3). (**B**) The endoscopic tissue biopsies from tumour and non-malignant control tissue sites that were collected were molecularly profiled using scRNA-seq, bulk RNA-seq, and whole genomic sequencing (WGS). (**C**) The whole bloods that were collected were separated into blood plasma, which was used for autoantibody profiling, and PBMCs, for spectral flow cytometry analysis.

Overall survival (OS) outcomes for all 38 OC patients (**Figure 1.9A**) and the 35 OAC patients (**Figure 1.9B**) enrolled in the inoperable arm of this trial were similar to those reported by the recent phase III clinical trials[30–32,36,81,320]. An additional metric to capture patient clinical outcome, termed clinical benefit (CB), was conceived retrospectively to denote the patients who experienced long-term disease control. A patient was deemed to have experienced CB if they showed no clinical progression for 12 months or longer following treatment initiation. Patient demographics and the rate of CB are summarised below for both the full cohort of 38 patients and for the subset of 35 OAC patients (**Table 1.1**). Histological expression of PD-L1 expression by CPS was not assessed in patients enrolled in the

inoperable arm of the LUD2015-005 study due to sample limitations and so is not reported in this thesis.

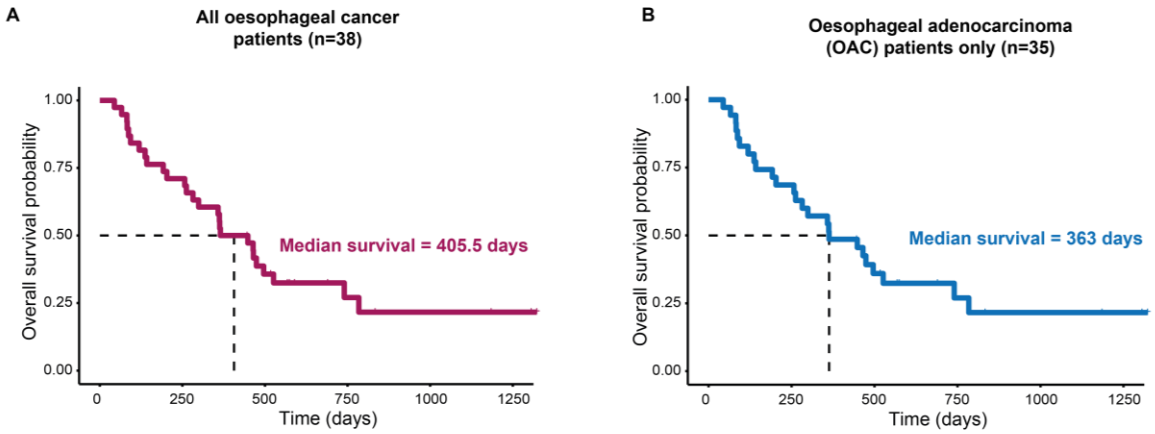


Fig 1.9. Overall survival outcomes in the inoperable arm of the LUD2015-005 clinical trial. Kaplan Meier plots of overall survival for **(A)** all 38 oesophageal cancer patients or **(B)** the 35 oesophageal adenocarcinoma patients enrolled in the inoperable arm of the LUD2015-005 clinical trial. The median survival for each cohort is marked on each plot with dashed lines.

Cohort	*Full cohort (OAC + OSCC, n=38)*	*OAC subset (n=35)*
Age in years at study enrolment		
Median (range)	59 (23 - 78)	59 (23 - 78)
***Sex**: Number (%)*		
Male	32 (84%)	29 (83%)
Female	6 (16%)	6 (17%)
***Stage at study enrolment**: Number (%)*		
III	3 (8%)	3 (9%)
IV	35 (92%)	32 (91%)
***Clinical benefit**: Number (%)*		
Yes	15 (39%)	14 (40%)
No	22 (58%)	20 (57%)
Not assessable	1 (3%)	1 (3%)

Table 1.1. Summary of patient demographics and clinical benefit for the inoperable arm of the LUD2015-005 clinical trial. Summary metrics are shown for the full cohort of 38 OAC and OSCC patients, and for the subset of 35 OAC patients. Clinical benefit was defined as no clinical progression for 12 months or longer following treatment initiation.

Many of these 38 patients developed at least one irAE during treatment. The rate of irAE development, the number of irAEs developed, and the grade of irAEs developed are summarised below for both the anti-PD-L1 only and anti-PD-L1 + anti-CTLA-4 treated cohorts (**Table 1.2**).

Immune checkpoint inhibitor treatment	*Anti-PD-L1 only (n=12)*	*Anti-PD-L1 + Anti-CTLA-4 (n=26)*
***Development of an irAE of CTCAE grade 1 or greater:** Number (%)*		
No	5 (42%)	7 (27%)
Yes	7 (58%)	19 (73%)
***Number of irAEs developed**: Number (%)*		
0	5 (42%)	7 (27%)
1-2	3 (25%)	9 (35%)
3-5	3 (25%)	6 (23%)
>5	1 (8%)	4 (15%)
***Grade of irAEs developed**: Number (%)*		
Total	23	63
Grade 1	16 (70%)	38 (60%)
Grade 2	3 (13%)	19 (30%)
Grade 3	4 (17%)	5 (8%)
Grade 4	0 (0%)	1 (2%)
Grade 5	0 (0%)	0 (0%)

Table 1.2. Summary of immune related adverse event (irAE) development for the inoperable arm of the LUD2015-005 clinical trial. Summary metrics are shown for the subset of 12 patients who were treated with anti-PD-L1 only as their ICI treatment, and for the subset of 26 patients who were treated with anti-PD-L1 + anti-CTLA-4 as their ICI treatment. The grades of irAEs listed are according to Common Terminology Criteria for Adverse Events (CTCAE; v4.03) grading[321].

In both the full cohort of 38 patients and the subset of 35 OAC patients, development of at least one irAE during treatment was significantly associated with improved OS (**Figure 1.10**). For this analysis, and all subsequent analyses related to development of irAEs that is described in this thesis, patients were stratified into two groups based on whether at least one irAE of grade 1 or greater, according to Common Terminology Criteria for Adverse Events (CTCAE; v4.03)[321] grading, was reported during their treatment.

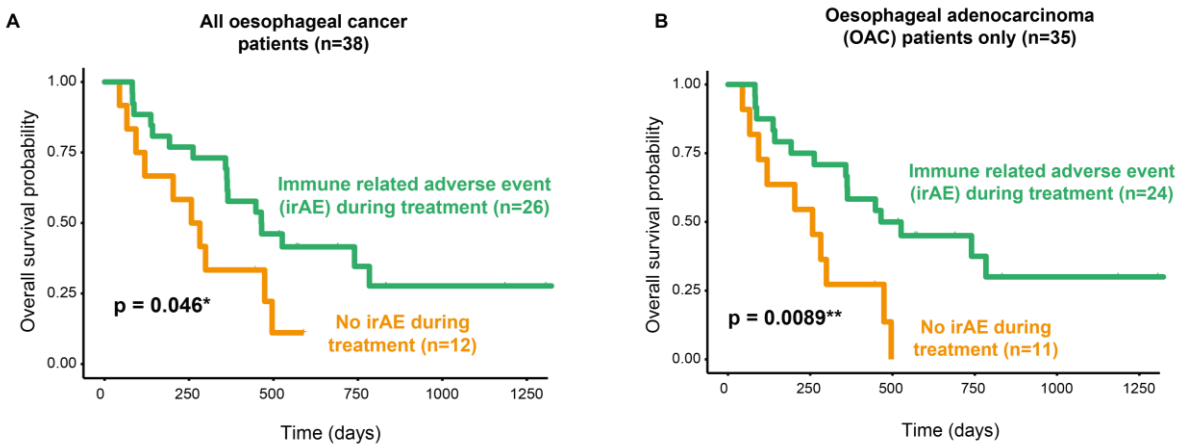


Fig 1.10. Development of at least one immune related adverse event in the LUD2015-005 inoperable arm is associated with improved patient survival. Kaplan Meier plots of overall survival (OS) for (**A**) all 38 oesophageal cancer patients or (**B**) the 35 oesophageal adenocarcinoma patients enrolled in the inoperable arm of the LUD2015-005 clinical trial, with patients split into those that developed at least one immune related adverse event (irAE) (green) and those who did not (yellow). Development of an irAE during treatment was significantly associated with longer OS in both patient groups. p-values were calculated by the log-rank method.

Due to the previously reported molecular differences in OAC and OSCC[322], analysis of inoperable LUD2015-005 samples in this thesis will be performed on the subset of 35 OAC patients.

1.5 Thesis aims and hypotheses

Despite treatment of OC with ICI+CTX having shown encouraging clinical outcomes overall, a number of key questions relating to the high variability in patient clinical outcome remain unanswered. Studies of other cancer types have identified promising prognostic factors which have potential to predict clinical outcome and further understanding of the mechanisms underlying patient response. Of these, the most promising seem to be the baseline immune activation of a patient, which could reflect the likelihood for ICI to sufficiently overcome immune tolerance, and features of intra-tumour TLSs and tumour infiltrating B cells, which could reflect the efficiency of anti-tumour immunity. However, similar studies evaluating both of these factors in ICI treated OC are lacking.

Subsequently, the aim of this thesis is to investigate the role of baseline immune activation and tumour infiltrating B cells in the clinical outcome of OC patients to ICI+CTX treatment. I hypothesised that clinical outcome in this setting is influenced by: 1) the patient's level of baseline immune activation, and 2) the abundance and transcriptional profiles of intra-tumour B cells.

This thesis has the following aims:

- **a)** Establish whether peripheral AAb profiling, as a surrogate for baseline immune activation, can predict clinical outcome to ICI+CTX treatment in OC patients (Chapter 3)
- **b)** Analyse whether peripheral signatures of increased immune activation are associated with improved clinical outcome and greater AAb numbers (Chapter 4)
- **c)** Determine whether transcriptional profiles of intra-tumour B cells show evidence of GC responses and whether the antibodies expressed by expanded BCR clones bind to identifiable autoantigens (Chapter 5)

2 Materials and Methods

2.1 LUD2015-005 clinical trial inoperable cohort

2.1.1 Trial design and patient selection criteria

The LUD2015-005 clinical trial (NCT02735239; REC reference: 16/SC/0015) is a phase I/II study of durvalumab as first-line therapy for patients with oesophageal or gastro-oesophageal cancer. Design and implementation of this trial involved key contributions from many individuals, including Professor Mark Middleton, Professor Xin Lu, Dr Richard Owen, Dr Michael White, Dr Simon Lord, Dr Ioannis Karydis, and many clinical and research collaborators (see Acknowledgements). The Oncology Clinical Trials Office at the University of Oxford (OCTO) and Ludwig Cancer Research provided key oversight of study procedures and data management and processing. This research was conducted with AstraZeneca UK Limited, who provided the investigational drugs for this trial.

Adult patients with a histologically confirmed diagnosis of locally advanced or metastatic oesophageal or gastro-oesophageal cancer were eligible for enrolment into the inoperable arms of the LUD2015-005 trial. Patients with previous systemic anti-cancer therapy for this same advanced disease were excluded from this study, as were patients who had been treated with ICI in any previous setting. Inclusion criteria included an Eastern Cooperative Oncology Group (ECOG) performance status of 0 or 1, and an anticipated lifespan of greater than 4 months.

Eligible patients were enrolled into one of three treatment arms for inoperable patients based on a dose escalation strategy, which proceeded to the next stage following establishment of acceptable safety profile at each step. Patients in the first stage received 750 mg of intravenous durvalumab (AstraZeneca) administered biweekly (Q2W) as the ICI agent, those in the second received the same regimen of durvalumab plus a single 37.5 mg priming dose of tremelimumab (AstraZeneca), and those in the third received durvalumab plus a single 75 mg priming dose of tremelimumab. After 4 weeks of treatment with these ICI agents alone, patients in all cohorts received a maximum of six cycles of CTX of oxaliplatin and capecitabine (Ox/Cap), in addition to continued Q2W durvalumab. In the absence

of a reason to discontinue treatment earlier, durvalumab infusions continued until the end of the final CTX cycle.

2.1.2 Assessment of patient clinical outcomes

Patients from the LUD2015-005 trial were consented for research by research nurses during registration on the trial. Research using their samples has ethical approval (16/SC/0015).

The primary outcomes of this phase I/II study involved safety: the number of subjects reporting adverse events, number of subjects experiencing a dose-limiting toxicity, and any changes from treatment onset in laboratory evaluations. Secondary outcome measures were overall survival (OS), progression free survival (PFS), and tumour response. Response was assessed using Immune-related Response Evaluation Criteria in Solid Tumours (irRECIST) criteria[323] from CT scans conducted during the screening period and every 6 weeks thereafter during treatment. irRECIST outcomes were classified as complete responses (irCR), partial responses (irPR), progressive disease (irPD), or stable disease (irSD). One patient with an overall response of irNN, which occurs when there is no measurable target lesion at baseline and an absence of on-treatment irPD or irCR for nontarget lesions, was considered to be irSD for the purposes of this work. Three patients who did not have an on-treatment CT scan for assessment of irRECIST criteria but were deemed to have clear on-treatment clinical progression were classified as irPD. irCR and irPR were classified as responders, and all other response categories were classified as non-responders. Tumour shrinkage during the ICI-4W window was calculated as the difference in sum of diameters of target lesions between pre-treatment and C1D8 CT scans. Overall survival (OS) and progression free survival (PFS) were calculated including data from the post-study follow-up period (data cut-off 16 October 2020). A post-hoc outcome metric for long-term disease control was also defined, which was termed “clinical benefit”. A patient was deemed to have attained clinical benefit (CB) if a centralised review process determined they had 12 months of survival from the initiation of treatment without confirmed disease progression. One

patient who received an alternative therapy within 12 months of treatment onset was excluded from CB analyses.

2.1.3 Sample collection

Endoscopic biopsies were collected for research analysis prior to the onset of treatment, during the fourth week of treatment with ICI alone, and at the end of the study treatment protocol (approximately 5-6 months after treatment onset). At each endoscopy, up to five 2 mm biopsy pairs were taken from the site of the tumour, while 1-2 pairs of normal control tissue were also taken from the descending duodenum (D2), gastric cardia, and normal oesophagus. Each biopsy pair was split into two aliquots, one of which was immediately snap-frozen using dry ice or liquid nitrogen, while the other was slow frozen with 1 mL of foetal bovine serum (FBS) or heat-inactivated human serum (Sigma-Aldrich, Cat.# H4522) with 10% DMSO and placed at -80°C in a controlled-rate freezing container. Snap and slow-frozen biopsies collected at local enrolment sites were stored at -80°C, and were shipped on dry ice to the central site for long-term storage at -80°C and subsequent analysis.

Whole blood samples were collected at the same treatment timepoints at which endoscopic biopsies were collected, and at an additional timepoint during the third week of the first cycle of CTX. These whole blood samples were processed into peripheral blood mononuclear cells (PBMCs) and blood plasma as previously described. Blood plasma was aliquoted and stored at -80°C. PBMCs were cryopreserved with 1 mL of FBS with 10% DMSO and placed at -80°C in a controlled-rate freezing container, before being transferred to liquid nitrogen for long-term storage.

2.2 Additional clinical sample collection

2.2.1 Endoscopic biopsies

For scRNA-sequencing, additional biopsies were also taken from three patients from the operable arms of the LUD2015-005 trial, which enrolled patients deemed suitable for surgery with curative intent. These operable OACs were treated in the neoadjuvant setting with a similar regimen of durvalumab, to which CTX or chemoradiotherapy was added following an ICI-only window. Additional biopsies were also collected from patients undergoing routine monitoring of known BO by the Translational Gastroenterology Unit (TGU) Biobank (John Radcliffe Hospital; REC reference 11/YH/0020) and released to our team for analysis.

For these supplemental operable OC and BO patients, a similar protocol for collection of endoscopic biopsies was followed as for the LUD2015-005 inoperable cohort. For operable OAC patients, 2 mm endoscopic biopsies were obtained from tumour and normal oesophagus prior to treatment and after 4 weeks of durvalumab alone, and were processed for snap and slow-freezing as above. At the surgery marking the end of the study protocol for these operable OC patients, core biopsies of the tumour site and normal oesophagus were also obtained from resected material. For BO patients, 2mm endoscopic biopsies of BO tissue and paired normal oesophagus were collected at a single timepoint.

2.2.2 Whole blood and blood plasma

Whole blood or blood plasma samples from patients with BO were obtained from the TGU biobank at the John Radcliffe hospital (Oxford). Whole blood samples or blood plasma samples from healthy volunteers (HV) or SOC treated OC patients were obtained from Oxford Radcliffe Biobank (ORB) (project ORB20/A016) at the John Radcliffe hospital (Oxford). These patients were each consented by research nurses prior to providing samples. Research using samples from both BO and HV patients is covered by my supervisor Professor Xin Lu's NHS REC approval 09/H0606/78, and/or ORB research tissue bank ethics, reference 19/SC/0173

2.3 Autoantibody based antigen reactivity

2.3.1 Custom Luminex bead-based antigen arrays

2.3.1.1 Antigen screening

Both antigen arrays were performed according to the manufacturer's instructions. Briefly, for each array, the selected 380 antigens, plus four control antigens, were immobilised on colour coded magnetic beads (MagPlex, Luminex) using COOH-NH2 chemistry. Controls consisted of uncoupled magnetic beads (bare-bead), His6ABP (control of binding to protein fragment tag, HPA resource), rabbit anti human IgG (loading control, Jackson immunoresearch, Cat.# 309-005-082), and Epstein-Barr nuclear antigen 1 (EBNA1, high frequency expected, Abcam, Cat.# ab138345). The coupling efficiency was tested using antibodies towards His6ABP (Human Protein Atlas Project (HPA) resource). Human plasma samples were each diluted 1:250 in assay buffer (3% BSA (Saveen-Werner, Cat.# B2000-500) and 5% milk powder (Sigma-Aldrich, Cat.# 70166-500G) in PBS supplemented with 0.05% Tween-20 and 160 µg/ml His6ABP tag (produced within HPA)). To help assess reproducibility, commercially available human plasma (Sigma-Aldrich, Cat.# P9523-5ML) was included in 3 wells per plate, and 3 buffer-only wells were included per plate as negative controls. Before incubation with the antigen bead array, all diluted samples were incubated for 1 hour (hr) at room temperature (RT) for the His6ABP in the buffer to pre-block any potential antibodies toward the albumin binding protein domain, since this is derived from gram-positive bacteria. Then samples were incubated with the antigen bead array for 2 hrs and then fixed using 0.02% paraformaldehyde (PFA) for 10 minutes. Beads were washed and re-suspended in 0.05% PBS-T and PE labelled anti-human IgG secondary antibody (PE) (Invitrogen, Cat.# H1010) was added for 30 minutes before washing with PBS-T. For each sample, the median fluorescent intensity (MFI) for each antigen was assessed using a FlexMap 3D instrument (Luminex).

2.3.1.2 Quality control and initial data processing

Quality control was and initial data processing was performed by Dr Ronald Sjöberg and his team at SciLifeLab. Briefly, two antigens were filtered out due to not reaching a high enough bead count (beads ≥ 16) in at least one sample, but no samples were filtered out.

To adjust for sample specific backgrounds, raw mean fluorescent intensity (MFI) values per sample were transformed into number of median absolute deviations (MADs) from the sample median. This calculation assumes that at least half of the antigens are unreactive in any given sample, and uses the following algorithm:

$$MADs_{\ sample} = \frac{MFI - median_{\ sample(MFI)}}{MAD_{sample\ (MFI)}}$$

The MADs data was further binned into 16 discrete bins to generate reactivity scores for all antigens within each sample, as shown below:

MADs cutoff (≥)	<0	0	5	10	15	20	25	30	35	40	45	50	55	60	65	70
Reactivity score	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5

Antigen specific cut-offs were then assigned from the distribution of the reactivity scores by determining where the slop is nearing a local minimum besides the highest peak. To calculate the number of reactivities per antigen and per sample, these cut-offs were used to transform the reactivity score for all antigens into a binary serostatus data set for all samples. Following this, a data file was provided by SciLifeLab which included the following information: raw MFI data; MADs data; reactivity score; binary serostatus data; antigen cut-offs; and the number of beads counted per antigen and per sample.

2.3.2 RayBio autoimmune array

Detection of 33 AAbs associated with autoimmune disease were performed using the Human Autoimmune Disease IgG Autoantibody Array G1 (RayBiotech Life, Cat.# PAH-AIDG-G1-16), according to the manufacturer's instructions. Briefly, each array was blocked with the provided blocking buffer, then plasma samples were diluted 1:100 with blocking buffer and incubated with the array. Unbound plasma was removed using the provided wash buffer and then biotinylated anti-human IgG was incubated with each array. Unbound antibody was removed by array washing and PE-labelled streptavidin was incubated with each array. Unbound streptavidin-PE was removed by washing and the air-dried slide was scanned using a FLA5100 scanner (Fujifilm Life Science). Scans of the array were analysed and saved as TIFF images using Multi Gauge v2.2 software (Fujifilm Life Science). Data extraction and quantification from TIFF images was performed using Spotxel software (Sicasys).

2.3.3 Determining cytomegalovirus (CMV) serostatus

CMV serostatus for clinical samples was determined based on Luminex bead-based detection of antibodies in blood plasma which were reactive to the CMV antigens pp150 N-terminus, pp52 and pp28. This assay was performed in collaboration with Dr. Tim Waterboer and Dr. Nicole Brenner (both DKFZ, Germany) and was based on multiplex serology methodology previously published[324]. Briefly, CMV antigens were expressed in *E. coli* as proteins fused to an N-terminal glutathione S-transferase (GST) domain. Each CMV antigen was then loaded onto a distinct Luminex bead set. The subsequent bead sets were mixed and incubated with blood plasma diluted 1:1000 in assay buffer (1 mg/mL casein (Sigma-Aldrich, Cat.# C-5890) in PBS) for 1 hr. Beads were washed and re-suspended in assay buffer and a biotinylated goat-α-human IgG secondary antibody (Dianova), diluted 1:1000 in assay buffer, was added for 1 hr before washing with assay buffer. Beads were washed and re-suspended in assay buffer and streptavidin-PE, diluted 1:750 in assay buffer, was added for 30 minutes before washing with assay buffer. CMV reactive antibodies were detected the following day using a Luminex 100

instrument (Luminex) with at least 100 beads were measured for each bead set and MFI were calculated.

Cut-offs for seropositivity to individual antigens were either based on published assay validation against serological reference assays or were arbitrarily defined by percentile plots. Overall CMV seropositivity was determined by seropositivity against at least two of the three antigens.

2.4 DNA and RNA extraction

2.4.1 Bulk NGS library preparation and sequencing

2.4.1.1 Whole-transcriptome RNA sequencing

Total RNA was extracted from whole endoscopic biopsies using the mirVana miRNA Isolation Kit (ThermoFisher Scientific, Cat.# AM1560). Briefly, snap-frozen biopsies were transferred directly into pre-chilled lysis/binding buffer in a 1.5 mL microcentrifuge tube, and the tissue was homogenised on ice using a disposable RNAse-free pestle (ThermoFisher Scientific, Cat.# 13236679). Acid-Phenol:Chloroform RNA extraction was then performed following manufacturer's instructions. Genomic DNA was digested from the eluted RNA using a 30 minute incubation with TURBO DNase (TURBO DNA-free Kit, ThermoFisher Scientific, Cat.# AM1907). RNA concentration was then calculated using the Qubit RNA BR Assay (ThermoFisher Scientific, Cat.# Q10211). RNA quality was assessed for each sample by on-chip electrophoresis using the RNA Pico Chip (Agilent, Cat.# 5067-1513) and the Total RNA Assay of the Agilent 2100 Bioanalyzer system. RNA was then plated and stored at -80°C. These steps were carried out together with Dr Thomas Carroll.

Prior to library preparation, thawed RNA was purified and concentrated with AMPure XP (BeckmanCoulter, Cat.# A63880) using a 2.8x ratio. Bead-based rRNA depletion and total RNA-sequencing library preparation with dual sample indexing was then conducted using the TruSeq Stranded Total RNA Library Prep Human/Mouse/Rat Kit (Illumina, Cat.# 20020596). Libraries were

sequenced on the Illumina HiSeq (75 cycles read 1, 8 cycles i7 index, 8 cycles i5 index, 75 cycles read 2) to a targeted depth of 50 million reads per library. Library preparation and sequencing was performed by the Oxford Genomics Centre (University of Oxford).

2.4.1.2 Whole genome sequencing

Genomic DNA was extracted from whole endoscopic biopsies using GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific, Cat.# K0721) according to manufacturer's instructions. Extracted DNA samples were then sent out for library preparation using a PCR-free protocol and sequencing using the Illumina NovaSeq6000. DNA extraction was performed by Dr Ayo Omiyale, Jaeho Chang, and Phil Xie. Library preparation and sequencing was performed by the Oxford Genomics Centre (University of Oxford).

2.5 Single-cell RNA sequencing

These steps were carried out in collaboration with Dr Thomas Carroll.

2.5.1 Tissue dissociation

Prior to dissociation, cryo-preserved biopsies were thawed by agitation of the cryovial in a 37°C water bath. For both fresh and cryopreserved material, biopsies were washed in PBS, and then transferred into a tube containing 5 mL of fresh or frozen tissue dissociation medium (Table 2.1) pre-warmed to 37°C. Biopsies were dissociated under constant rotation in an incubator at 37°C and 5% CO_2 until fragments reduced substantially in size, approximately 60-90 minutes for fresh tissue, or 30-60 minutes for frozen tissue. The dissociated solution was then filtered using a 100 µM CellTrics filter (Sysmex, Cat.# 04-004-2328), to remove any undigested tissue fragments or cell clumps. At this stage, for fresh material, lysis of red blood cells was performed using red blood cell lysis solution (Miltenyi, Cat.# 130-094-183), and dead cells were then depleted using Dead Cell Removal Kit (Miltenyi, Cat.# 130-090-101). These steps were not performed for frozen material because red blood cell contamination of dissociated pellets was not visible, and dead cell removal did not improve viability

in optimisation tests. Cells were then washed with pre-chilled cell resuspension buffer (0.04% Bovine Serum Albumin in PBS), and were passed through a 70 µM Flowmi Cell tip strainer (Sigma-Aldrich, Cat.# BAH136800070) to remove cell clumps. Filtered cell suspensions were then counted using a TC20 automated cell counter (Bio-Rad), and kept on ice until encapsulation.

Media	*Composition*
Fresh tissue dissociation medium	DMEM/F12 (Gibco, Cat#. 11330032) with 10% Human serum (Sigma-Aldrich, Cat.# H4522), 0.24 mg/mL Animal Free Collagenase/Dispase Blend II (Sigma-Aldrich, Cat.#, SCR140), and 0.1 mg/mL DNAse I (Sigma-Aldrich, Cat.# D5025)
Frozen tissue dissociation medium	DMEM (Gibco, Cat#. 10566016) with 10% FBS, 2.5 mg/mL Collagenase D (Sigma-Aldrich, Cat#. 11088866001), 0.5 mg/mL Liberase DL (Sigma-Aldrich, Cat#. 5466202001), 0.2 mg/mL DNAse I (Sigma-Aldrich, Cat#. D5025)

Table 2.1. Components of fresh and frozen dissociation media.

2.5.2 Cell encapsulation, library preparation and sequencing

Single-cell RNA-sequencing (scRNA-seq) was performed using the 5' Chromium Single Cell Immune Profiling v1.1 workflow (10x Genomics). Briefly, encapsulation of cells was performed using the Chromium Controller (10x Genomics), and then reverse transcription of RNA, cDNA amplification, and construction of final libraries was conducted according to the manufacturer's instructions. Size profiles of amplified cDNA and final libraries were assessed by on-chip electrophoresis (Agilent 2100 Bioanalyzer system) using the Bioanalyzer High Sensitivity DNA Kit (Agilent, Cat.# 5067-4626). Concentration of final libraries was quantified by the Qubit BR DNA Assay (ThermoFisher Scientific, Cat.# Q32853). Libraries were sequenced on an Illumina NovaSeq 6000 or NextSeq 550 (26 cycles read 1, 8 cycles i7 index, 98 cycles read 2), targeting a minimum of 20,000 reads per cell for gene expression (GEX) libraries, and a minimum of 5,000 reads per cell for BCR libraries.

2.6 Processing of whole blood

Whole blood was collected from BO and HV patients, as described in section 2.2.2, were processed into PBMCs and blood plasma using a similar protocol as that used for the LUD2015-005 trial inoperable cohort. Briefly, undiluted EDTA blood was centrifuged at 800g for 10 minutes and plasma was harvested. Residual cellular debris was removed from plasma by centrifugation at 12,00g for 10 minutes and the subsequent purified plasma was stored at -80^{o}C. PBMCs were isolated from the remaining whole blood using density gradient centrifugation over Lymphoprep (STEMCELL Technologies, Cat.# 07861) as previously described.

2.7 Cloning, expression, and purification of recombinant antibodies

2.7.1 Cloning into expression vectors

Selected B-cell clones from the were produced by DNA synthesis by Genscript and cloned into mammalian expression vectors containing the human IgG1 constant regions of IgH, Igκ, or Igλ which were gifted to us from Professor Gavin Screaton (University of Oxford), have been previously used to express recombinant antibodies[325,326], and are available from Addgene (plasmid# 80795, 80796, 99575). Briefly, the constant region vectors and synthesised inserts containing the V(D)J regions were separately digested using the appropriate restriction enzymes (**Table 2.2**). V(D)J inserts were then ligated with the appropriate linearised constant region vector, by overnight incubation at 16˚C in the presence of T4 DNA ligase (New England Biolabs, Cat.# M0202), and used to transform Max Efficiency DH5α Competent Cells (Invitrogen, Cat.# 18258012), which were then plated out onto ampicillin containing agar plates. Three colonies per construct were selected and DNA was extracted using a QuickLyse Miniprep Kit (Qiagen, Cat.# 27406), according to the manufacturer's instructions.

Vector	*Conditions used for digestion of insert and vector*	*Addgene plasmid name and ID*
IgH	45 µL insert/vector + 0.125 µl AgeI-HF (New England Biolabs, Cat.# R3552) + 2µl SalI-HF (New England Biolabs, Cat.# R3138) + 5µl rCutSmart buffer (New England Biolabs, Cat.# B6004) at 37°C for 6hrs	AbVec2.0-IGHG1 (ID: 80795)
Igκ	45 µL insert/vector + 0.125 µl AgeI-HF (New England Biolabs, Cat.# R3552) + 5µl rCutSmart buffer (New England Biolabs, Cat.# B6004) at 37°C for 3hrs. Then add 0.5µl BsiWI (New England Biolabs, Cat.# R0553) and digest at 37°C for a further 3hrs.	AbVec1.1-IGKC (ID: 80796)
Igλ	45 µL insert/vector + 0.125 µl AgeI-HF (New England Biolabs, Cat.# R3552) + 0.5µl XhoI (New England Biolabs, Cat.# R0146) + 5µl rCutSmart buffer (New England Biolabs, Cat.# B6004) at 37°C for 6hrs	AbVec1.1-IGLC2-XhoI (ID: 99575)

Table 2.2. Summary of mammalian expression vectors containing the human IgG1 constant regions of IgH, Igκ, or Igλ.

The presence of insert and vector was confirmed by digestion of the extracted DNA with the appropriate restriction enzymes, and subsequent agarose gel electrophoresis at 180V for 45 minutes, using UltraPure TBE buffer (Life Technologies, Cat.# 15581028) and 1kb plus DNA ladder (Invitrogen, Cat.# 10787018). Extracted DNA was also sent to Genewiz to confirm sequences. These steps were performed in collaboration with Dr Hantao Lou.

2.7.2 HEK 293T/F transfection

For each antibody successfully cloned, the appropriate constructs containing the IgH and Igκ, or Igλ inserts were transfecting into either HEK-293T (gifted from Professor Gavin Screaton, University of Oxford) or Expi293F cells (ThermoFisher Scientific, Cat. A14635) using polyethylenimine (PEI) (Polysciences, Cat. # 23966-1). HEK-293T or Expi293F cells were then cultured with equal amounts of heavy- and light-chain plasmids and PEI in DMEM Glutamax media (Gibco, Cat.# Cat#. 10566016) for

24hrs. Then supernatant was removed and media was changed to UltraDOMA Protein-free Hybridoma medium (Lonza, Cat.# 12-727F). After five days of culture, the supernatant was collected, and antibody concentration was assessed by ELISA. Briefly, F96 Maxisorp NUNC Immuno plates (VWR, Cat.# 442404) were coated with an anti-Human IgG (γ-chain specific) antibody (Sigma, Cat.# I5885) diluted 1:5000 in carbonate buffer (Merck, Cat #. C3041) at 4°C overnight. Plates were washed twice with PBS-T (0.1% Tween-20 (Sigma-Aldrich, Cat.# P1379) in PBS) and then blocked with blocking buffer (3% BSA in PBS) at RT for 1hr. Plates were washed three times with PBS-T and then incubated with the appropriate antibody-containing supernatant at 37°C for 1hr. Plates were washed three times with PBS-T before being incubated with an alkaline phosphatase conjugated anti-Human IgG (Fc specific) (Merck, Cat.# A9544) secondary antibody, diluted 1:5000 in blocking buffer, at 37°C for 1hr. Plates were washed four times with PBS-T before AP substrate kit (Bio-rad, Cat.# 1721063) was added. Reaction was stopped by adding 0.4M NaOH and absorbance at 405-420 nm was measured using a Sunrise microplate reader (Tecan) and concentration was calculated using a standard curve generated from standards of known concentrations. These steps were performed with Dr Hantao Lou.

2.7.3 Antibody purification

A subset of antibodies were purified into IgG using a column purification method with protein A agarose beads (Merck, Cat.# P9424). Briefly, the column and tubing was treated with 0.5M NaOH for 30 minutes and then washed once with ultrapure water and washed three times with wash buffer (0.02M NaH_2PO_4, 0.15M NaCl, pH8 in ultrapure water). Protein A agarose beads were then added to the column and then washed three times with wash buffer. Antibody containing supernatant was then mixed 1:1 with wash buffer and passed through the column at a dropwise rate, with one drop every 3-5 seconds. After all the solution passed through the column, the column was washed three times with wash buffer. Purified IgG was then eluted from the protein A beads using 0.1M Glycine pH2.7 and collected directly into tubes containing an equal volume of Tris-HCl pH8. Purified antibody was then concentrated using Amicon Ultra-15 Centrifugal Filter Unit (Millipore, Cat.# UFC901096),

according to the manufacturer's instructions. Briefly, the Amicon unit was shaken three times with PBS to remove glycerol from the membrane, then unit was washed through centrifugation with PBS at 3000 rpm for 5 minutes. Purified antibody was then added to the unit and the unit was centrifuged at 3000 rpm for 20 minutes. The unit was then washed with PBS three times, as above. Antibody concentration was then measured by ELISA, as described in section 2.7.2. These steps were performed with Dr Hantao Lou.

2.8 Cell culture

Gastric and oesophageal cancer cell lines were cultured according to recommendations by the European Collection of Authenticated Cell Cultures. FLO-1 cells were cultured in DMEM Glutamax media (Gibco, Cat.# Cat#. 10566016) supplemented with 10% FBS. SKGT4, OE19, and OE33 cells were cultured in RPMI Glutamax media (Gibco, Cat.# 61870010) supplemented with 10% FBS and pen/strep. AGS cells were cultured in Ham's F12 with L-Glutamine (Lonza, Cat.# BE12-615F) supplemented with 10% FBS. Adherent cells were detached from culture flasks for splitting or plating out using either 0.25% Trypsin-EDTA (Gibco, Cat.# 25200056) or Accutase solution (Sigma-Aldrich, Cat.# A6964). For the experiments described in this thesis, aliquots of each of these five cell lines were obtained from historic stocks belonging to the Lu Lab.

2.9 Microscopy

Immunofluorescence was used to assess binding of mAbs to gastric and oesophageal cancer cell lines. Briefly, cells were plated into a flat-bottom 96 well plate at $5x10^5$ cells/well and fixed through incubation with 4% PFA (ThermoFisher Scientific, Cat.# 28906) at RT for 10 minutes. Cells were washed with PBS and then permeabilised through incubation with 0.25% Triton-X-100 (Sigma-Aldrich, Cat.# T9284) in PBS at RT for 10 minutes. Cells were washed again with PBS and then incubated with blocking buffer (5% BSA (Sigma-Aldrich, Cat.# A7906) in PBS) at RT with constant agitation for 1hr. Blocking solution was then removed, and cells were incubated with 50 µL of mAb-containing supernatant at

4°C with constant agitation overnight. Cells were then washed three times with PBS and then incubated with Anti-Human IgG-FITC antibody (Sigma-Aldrich, F3512-1ML) secondary antibody, which was diluted 1 in 1000 in blocking buffer and with 1 µg/mL DAPI (Sigma-Aldrich, Cat.# D9542), at RT with constant agitation for 1hr. Cells were washed three times with PBS, then covered with PBS. Binding of mAbs to cells was assessed by measuring fluorescence with an Operetta CLS (PerkinElmer).

2.10 Flow cytometry

2.10.1 Ex vivo immunophenotyping

The antibody panels and protocols used for these experiments were designed by our collaborator Dr Isabela Pedroza-Pacheco and staining of samples was carried out with her help. Briefly, cryo-preserved PBMCs were thawed by agitation of the cryovial in a 37°C water bath, and then transferred into complete media (**Table 2.3**), supplemented with 250U of Benzonase nuclease (Millipore, Cat# E1014-25KU), pre-warmed to 37°C. Cells were then centrifuged at 500g for 10 minutes and then resuspended in complete media supplemented with 250U of Benzonase, before being rested for 1hr in an incubator at 37°C and 5% CO_2. Cells were counted using a TC20 automated cell counter (Bio-Rad) and for each sample between 0.5 and 3 million cells were plated into three 96 well v-bottom plates, one for each antibody panel, ready for antibody staining.

Reagent	*Manufacturer*	*Catalogue number*	*Volume for 500 mL complete media (mL)*
RPMI minus glutamine	Life Technologies	L31870-025	429.5
Heat-inactivated FBS	Sigma-Aldrich	F9665	50
Glutamax	Gibco	35050-038	5
Peneclillin-Streptomycin	Gibco	15070063	5
Non-essential amino acids	Gibco	11140-035	5
Sodium pyruvate	Sigma-Aldrich	S8636-100ML	5
Beta mercaptoethanol	Gibco	31350-010	0.5

Table 2.3. Components of complete media.

Plated cells were washed with Brilliant Stain Buffer (BD Biosciences Cat.# 563794) and then incubated with Human Fc Block (BD Biosciences, Cat.# 564220) for 10 minutes at RT to limit non-specific antibody binding. The appropriate master mix of antibodies for surface staining (**Table 2.4**) pre-mixed in Brilliant stain Buffer was then added to the cells which were then incubated at 37°C for 15 minutes. Cells were then washed with PBS and incubated with LIVE/DEAD Fixable Blue Dead Cell Stain Kit (Invitrogen, Cat.# L34961) and BUV615 conjugated streptavidin (BD Biosciences, Cat.# 613013) at RT for 10 minutes in the dark. Cells were washed with PBS and then fixed and permeabilised using the eBioscience Foxp3/Transcription Factor Staining Buffer Set (ThermoFisher Scientific, Cat.# 00-5523-00), according to the manufacturer's instructions. Then the appropriate master mix of antibodies for intracellular staining (**Table 2.4**) pre-mixed in eBioscience Permablilization Buffer (ThermoFisher Scientific, Cat.# 00-5523-00) was then added to the cells, and cells were incubated at RT for 45 minutes in the dark.

Antigen	*Fluorophore/tag*	*Manufacturer*	*Catalogue number*	*Surface or intracellular staining*	*Panel(s) used in*
2B4	PERCPCY5.5	Biolegend	329515	Surface	Panel 2
BAFFR	BV750	BD Biosciences	747040	Surface	Panel 3
BTLA	PE/Dazzle 594	Biolegend	344521	Surface	Panel 3
CCR4	BV421	Biolegend	359413	Surface	Panel 1
CCR6	Pe/Dazzle	BD Biosciences	564816	Surface	Panel 1
CCR7	BV650	Biolegend	353234	Surface	Panel 1
CD10	BV711	Biolegend	312225	Surface	Panel 3
CD11C	BUV395	BD Biosciences	563788	Surface	Panel 3
CD123	APC	Miltenyi	130-090-901	Surface	Panel 3
CD127	BUV661	BD Biosciences	749981	Surface	Panel 1
CD138	eFluor450	Invitrogen	48-1388-41	Surface	Panel 3
CD14	Spark NIR 685	Biolegend	367149	Surface	Panel 2
CD14	FITC	BD Biosciences	555397	Surface	Panel 3
CD15S	BV510	BD Biosciences	563529	Surface	Panel 1
CD16	BV605	Biolegend	302039	Surface	Panel 3
CD16	Pacific blue	BD Biosciences	558122	Surface	Panel 2
CD161	BV750	BD Biosciences	746921	Surface	Panel 1 and 2
CD19	AF647	Biolegend	302222	Surface	Panel 1
CD19	SBV515	Bio-Rad	MCA1940GA	Surface	Panel 3
CD20	AF532	Thermo	58-0209-42	Surface	Panel 3

CD21	BUV661	BD Biosciences	741605	Surface	Panel 3
CD22	BUV496	BD Biosciences	741142	Surface	Panel 3
CD24	APC-Fire750	Biolegend	31140	Surface	Panel 3
CD25	PEcy7CD25	BD Biosciences	335824	Surface	Panel 1
CD27	APC-Fire 810	Biolegend	393213	Surface	Panel 1, 2 and 3
CD3	Spark Blue 550	Biolegend	344851	Surface	Panel 1 and 2
CD3	Spark NIR 685	Biolegend	344862	Surface	Panel 3
CD38	BUV496	BD Biosciences	564658	Surface	Panel 1 and 2
CD38	BV650	Biolegend	356619	Surface	Panel 3
CD39	PERCPCY5.5	BD Biosciences	564899	Surface	Panel 1
CD4	Spark NIR 685	Biolegend	344657	Surface	Panel 1 and 2
CD4	SBV710	Bio-Rad	MCA1267GA	Surface	Panel 3
CD40	APC-R700	BD Biosciences	565179	Surface	Panel3
CD43	BUV563	BD Biosciences	748876	Surface	Panel 3
CD45RA	BUV563	BD Biosciences	612926	Surface	Panel 1
CD56	BUV737	BD Biosciences	612766	Surface	Panel 2
CD57	BV605	Biolegend	393303	Surface	Panel 1 and 2
CD7	BV650	BD Biosciences	740565	Surface	Panel 2
CD8	BV570	Biolegend	301038	Surface	Panel 1 and 2
CD85J	PECY5	BD Biosciences	551054	Surface	Panel 2 and 3
CD86	PERCPCY5.5	BD Biosciences	561129	Surface	Panel 3
CRACC	Biotin	Miltenyi	130-099-574	Surface	Panel 2

CTLA-4 (CD152)	PerCP-eFluor710	Invitrogen	46-1529-42	Intracellular	Panel 1
CXCR3	PECY5	BD Biosciences	551128	Surface	Panel 1
CXCR5	BB515	BD Biosciences	564624	Surface	Panel 1 and 3
CXCR5	AF647	BD Biosciences	558113	Surface	Panel 2
DNAM-1	BUV563	BD Biosciences	748429	Surface	Panel 2
FAS	BV785	BIolegend	305645	Surface	Panel3
FCERG	FITC	Millipore	FCABS400F	Intracellular	Panel 2
FOXP3	APC	Invitrogen	17-4776-42	Intracellular	Panel 1
GRAB	AF700	BD Biosciences	560213	Intracellular	Panel 1 and 2
Helios	eFluor450	Invitrogen	48-9883-41	Intracellular	Panel 1
HLA-DR	BUV805	BD Biosciences	748338	Surface	Panel 1, 2 and 3
ICOS (CD278)	Biotin	Invitrogen	13-9948-82	Surface	Panel 1
IgA	AF647	Jackson Immuno Research	109-605-011	Surface	Panel 3
IgD	Super Bright 436	Invitrogen	62-9868-41	Surface	Panel 3*
IgG	BUV737	BD Biosciences	612819	Surface	Panel3
IgM	BV570	Biolegend	314517	Surface	Panel 3
IL21R	BV421	Biolegend	359510	Surface	Panel 3
KI-67	BV480	BD Biosciences	566119	Intracellular	Panel 1, 2 and 3
KIR2DL1	APC	R&D	FAB1844A	Surface	Panel 2
KIR2DL2/L3/S2	APC	Beckman	A22333	Surface	Panel 2

KIR3DL1	APC	Miltenyi	130-092-474	Surface	Panel 2
KIR3DL2	APC	R&D	FAB2878A	Surface	Panel 2
KLRG1	BV785	Biolegend	138429	Surface	Panel 1 and 2
LAG-3	Super Bright 436	Invitrogen	62-2239-41	Surface	Panel 1 and 2
LAP	PE	R&D	FAB2463P	Surface	Panel 1
NKG2A	PEcy7	Beckman	B100246	Surface	Panel 2
NKG2C	BV421	BD Biosciences	748169	Surface	Panel 2
NKG2D	BV510	BD Biosciences	563266	Surface	Panel 2
NKp30	BUV661	BD Biosciences	750510	Surface	Panel 2
NKp46	PercpeFluor710	Invitrogen	46-3359-42	Surface	Panel 2
OX40L	PE	Ancell	400-050	Surface	Panel 3
HLA-DR/DP/DQ	BUV737	BD Biosciences	741845	Surface	Panel 1
PD1	BV711	Biolegend	329928	Surface	Panel 1 and 2
PDL1	PercpeFluor710	Invitrogen	46-5983-42	Surface	Panel 3
PDL2	PercpeFluor710	Invitrogen	46-5888-42	Surface	Panel 3
Perforin	APC-Fire 750	Biolegend	353317	Intracellular	Panel 1 and 2
PLZF	PE-CF594	BD Biosciences	565738	Intracellular	Panel 2
SIGLEC7	PE	Biolegend	339204	Surface	Panel 2
TBET	PEcy7	Biolegend	644824	Intracellular	Panel 3
TIGIT	BUV395	BD Biosciences	747845	Surface	Panel 1 and 2
VH4-34	Biotin	IgM_Biosciences	Gift	Surface	Panel 3

Table 2.4. Antibodies used for *ex vivo* immunophenotyping. N.B. The biotinylated 9G4 VH4-34 antibody was given as a gift to our collaborator Dr Isabela Pedroza-Pacheco by IgM Biosciences. * The anti-IgD antibody was only used to stain three out of five batches of samples due to supply issues.

Cells were washed twice with eBioscience Permablilization Buffer and then resuspended in PBS. Samples were analysed using a Cytek Aurora (Cytek). Raw data was unmixed by Dr Isabela Pedroza-Pacheco using SpectroFlo v3.0 (Cytek), with reference control samples prepared with her help.

2.10.2 Recombinant antibody binding to untreated and cisplatin treated cell lines

Cells were detached from culture flasks using Accutase solution (Sigma-Aldrich, Cat.# A6964) and plated into 96 well flat bottom plates at $2x10^5$ cells/well. Cells were washed once with FACS buffer (2% FBS, 2mM EDTA (Gibco, Cat.# 15575020) in PBS)) and then incubated with the appropriate antibody containing supernatant at 4°C for 1 hr. Cells were washed twice with FACS buffer and then stained with PE anti-human IgG antibody (Biolegend, Cat.# 409304) at RT for 20 minutes. Cells were then washed twice with FACS buffer and then fixed through incubation with 4% PFA at RT for 15 minutes. Cells were then washed once and resuspended in FACS buffer. Antibody binding to cells was assessed using a LSRFortessa X-20 (BD Biosciences). For experiments assessing antibody binding to cisplatin treated cells, cells were cultured with cisplatin (Merck, Cat.# BP809) for 48h prior to harvesting from culture.

2.10.3 Antibody dependent cellular phagocytosis assay

Primary human monocytes were isolated from cryopreserved PBMCs of HVs using a magnetic bead based CD14 positive selection kit (Miltenyi, Cat.# 130-050-201), according to the manufacturer's instructions. Briefly, cryo-preserved PBMCs were thawed by agitation of the cryovial in a 37°C water bath, then washed with MACS buffer (0.5% BSA in PBS) and counted using a TC20 (Bio-rad). Cells were washed again with MACS buffer and then resuspended in 80 µL of MACS buffer per $1x10^7$ cells. 20 µL of CD14 microbeads per $1x10^7$ cells were then added to the cells and this mix was incubated at 4°C for 15 minutes. Cells were then washed and resuspended in MACS buffer. Resuspended cells were applied to an MS column (Miltenyi, Cat.# 130-042-201) and the column was washed through three times with

MACS buffer. CD14-positive cells (monocytes) were harvested by plunging MACS buffer through the column.

The isolated monocytes were then stained with celltrace violet (Invitrogen, Cat.# C34557), according to the manufacturer's instructions. FLO1 cells which had been treated with cisplatin (Merck, Cat.# BP809) for 48h before being washed with PBS, harvested, and stained with celltrace yellow (Invitrogen, Cat.# C34567), according to the manufacturer's instructions. These celltrace yellow stained FLO1 cells were then co-cultured at 37°C with the celltrace violet stained monocytes for either 30, 60, or 120 minutes in the presence or absence of 10 µg/ml of antibody. After co-culturing for the set amount of time, cells were harvested and fixed through incubation with 4% PFA for 15 minutes. Then cells were washed with FACS buffer and stained with PE anti-human IgG antibody (Biolegend, Cat.# 409304) at RT for 20 minutes. After washing again with FACS buffer, samples were then analysed using a LSRFortessa X-20 (BD Biosciences).

2.11 Bioinformatics

2.11.1 Autoantibody profiling

2.11.1.1 Custom Luminex bead-based antigen arrays

Firstly, the raw MFI data for each antigen and each sample were processed to net MFI data on a per sample basis by subtracting the MFI value for the bare bead negative control in the corresponding sample in order to adjust for sample specific background signal. Any negative net MFI values were then set to zero to allow for log transformation. The MFI signal for the beads coupled to only the His6ABP tag, which is bound to each antigen, were then plotted for each sample to identify any samples with a high signal and hence potentially incomplete depletion of anti-His6ABP reactive antibodies which should have been depleted prior to plasma incubation with the beads. As each antigen had this His6ABP tag bound, it was essential to remove any samples with a high MFI signal to the His6ABP coupled bead to avoid false positive reactivities. Next, as the net MFI data showed a

strong positive skew, the net MFI values were natural log (ln) transformed, and, on a per antigen basis, had two distributions fitted using the mclust package[327] (v5.4.9). The proportion of overlap between the two distributions fitted by mclust were calculated in python (v3.8) using a script written by Otto Sumray (University of Oxford), and the difference in net MFI between the means of the peaks of the two distributions fitted was calculated in R (v4.1.0). Serostatus data were calculated as seronegative (0) if mclust classified the sample as most likely belonging in the lower distribution, or seropositive (1) if mclust classified the sample as most likely belonging in the upper distribution.

2.11.1.2 Raybiotech AAb array

Extracted and quantified signals which were deemed to be the result of artefacts, based on manual inspection of the corresponding TIFF image and high standard deviation between replicates, were removed. For each sample, the signal for each AAb was calculated as the average of the signals from replicates. Average signals were then normalised across the arrays based on average signals for the two positive controls, IgG and BSA-biotin. Normalised signals for each AAb across all of the samples were then plotted as boxplots and cutoffs for seropositivity were set on a per antigen basis based on the distribution of signals. Binary serostatus data was calculated as seropositive if the normalised signal was above the cutoff. This analysis, processing, and visualising of data was performed using R[328] (v4.1.0) and the ggplot2 package[329] (v3.3.5).

2.11.2 Ex vivo immunophenotyping

Firstly, unmixed flow cytometry data was imported into FlowJo™ Software[330] (v10.8.1) and the live single mononuclear cells (MNCs) from each file were gated, based on the time, SSC-A, FSC-A, FSC-H, and LIVE/DEAD blue variables (**Figure 2.1**). For each sample, these gated MNCs were exported as a new fcs file.

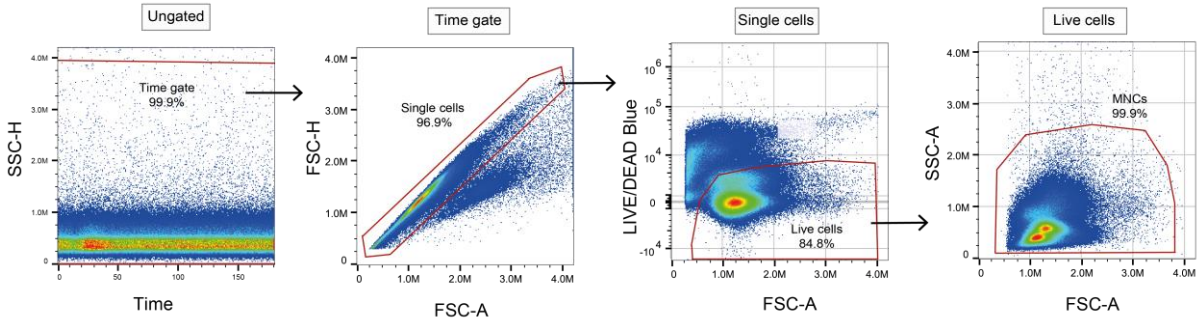


Figure 2.1. Gating strategy used to gate single mononuclear cells. (**A**) Gating of live single mononuclear cells from processed and normalised flow cytometry data was performed using FlowJo™ Software. Firstly, events were gated by time to ensure that a consistent event rate was observed. Next, FSC-H and FSC-A were used to gate single cells, and these single cells were then gated to just live cells based on LIVE/DEAD blue staining and FSC-A profile. Finally, MNCs were gated from the live single cells based on the SSC-A vs FSC-A profile.

For these files of live MNCs, variance was stabilised on a channel-specific basis using the FlowVS package[331] (v1.3.0), then low quality events were removed using the PeacoQC package[332] (v1.8.0), next batch normalisation was performed using the CytoNorm package[333] (v0.0.10), with stimulated and unstimulated control PBMC samples, which were run in each of the five batches of samples, used to train the normalisation model. This quality-controlled batch normalised data was then imported into FlowJo™ Software[330]. For each sample, a set number of cells from each of the 15 gated populations were then randomly down sampled using FlowJo™ Software[330] and the median normalised expression of each marker was calculated using the R package CATALYST[334] (v1.20.1). To further characterise the cell types, clustering and Uniform Manifold Approximation and Projection (UMAP) dimensionality reduction of marker expression data from the major cell types gated into cell subsets was performed using the packages FlowSOM[335] (v2.4.0) and CATALYST[334]. The optimal number of cell subset clusters was determine using the delat_area function and manual inspection of expression profiles of the clusters. Any clusters of cells with an abundance of <0.5% were manually merged with the cluster with the most similar expression profile, using the function mergeClusters. Differential analysis of abundances and states of the cell subsets identified was performed using the package diffcyt[336] (v1.16.0), with the option testDA_edgeR for abundances, and testDS_limma for states. Each analysis was adjusted for the confounding variables age and sex.

2.11.3 Whole-transcriptome RNA sequencing

2.11.3.1 Processing of raw data

Raw RNA-sequencing data were processed following best practices using a reproducible Nextflow[337] pipeline. This workflow was based off of STAR alignment[338] to the same reference used for scRNA-seq alignment, and generation of final counts matrices using featureCounts[339]. This data processing was performed by Dr Thomas Carroll and the reproducible pipeline and associated information can be found at https://github.com/t-carroll/LUD2015-005.

2.11.3.2 Deconvolution of bulk RNA-seq data

Deconvolution of bulk RNA-sequencing data was performed using the MCP-counter package[340] to estimate the proportion of 11 different cell subsets, based on in-built cell type reference matrices.

Single-cell based RNA deconvolution was performed on the same bulk RNA-sequencing data to estimate the abundances of the 45 major cell types identified from scRNA-seq. Briefly, normalised counts in a linear scale were used for deconvolution, which were obtained for scRNA-seq data using the vst function from the sctransform package and for bulk RNA-seq data using the median of ratios method embedded in the estimateSizeFactors function from DESeq2. Potentially confounding gene types (mitochondrial, ribosomal, sex-specific, and TCR/BCR variable region genes) were excluded from the input reference matrix. Deconvolution analysis on bulk RNA-seq samples from LUD2015-005 inoperable cohort was conducted using BayesPrism. with input reference matrix subset to the 3000 most variable genes (as calculated by FindVariableFeatures from Seurat). Potentially confounding genes were removed prior to the run, and OAC cells were labelled as malignant. This data processing was performed by Dr Thomas Carroll.

2.11.4 Whole genome sequencing

Whole genome sequencing data were processed for mutation calling and copy number analysis and then TP53 mutations and copy number alterations were called using a consensus of two of three mutation callers (Strelka2, Mutect2 and Octopus), and Batternberg for copy number alterations. These steps were carried out by Professor Benjamin Schuster-Böckler, Dr Iliana Peneva, and Dr Joseph Kaplinsky.

2.11.5 Single-cell RNA sequencing

Processing of scRNA-seq data was conducted using the Cell Ranger pipeline from 10x Genomics. Briefly, raw BCL files were converted to demultiplexed FASTQ files using cellranger mkfastq with --use-bases-mask=Y26n*,I8,Y98n*. A custom reference sequence was then prepared using the concatenation of the GRCh38 human genome (ENSEMBL100) and supplementary contigs from the Genomic Data Commons alignment reference, namely the decoy contig hs38.d1 and a collection of viral genomes commonly found in human cancer samples. The full GENCODE v34 GTF was used as a gene annotation file, filtering out readthrough transcripts and annotations to PAR regions in chromosome Y as per Cell Ranger recommendations. A custom Cell Ranger reference set was then prepared from these sequence and annotation files using cellranger mkref with default arguments. Finally, the identification of true cells and construction of a UMI counts table for each sample was performed using cellranger count, and these per-sample counts were combined into a single matrix for downstream analysis using cellranger aggr without normalisation. Downstream analysis was conducted using Seurat[341]. To begin, low-quality cells with less than 200 genes detected or more than 25% mitochondrial reads were removed. For initial clustering of the full dataset, normalisation using SCTransform[342], UMAP dimensionality reduction, and clusters identification with the default Louvain algorithm was performed. Broad cell type partitions were then identified and subset into separate Seurat objects. To characterise smaller cell subsets, the above workflow was repeated on each of these partitioned objects. To minimise batch effect between fresh and frozen samples for these

smaller objects, SCTransform normalisation was conducted on fresh and frozen samples separately, and then integrated these two datasets using FastMNN[343] a mutual nearest neighbour batch correction method, before proceeding to dimensionality reduction. Marker genes for each identified cluster were calculated using the FindAllMarkers function, and these markers were used to manually assign cell type and cell subtype labels. This data processing and analysis was performed by Dr Thomas Carroll.

2.11.6 Single-cell B cell receptor analysis

For the B cells which had gene expression (GEX) and paired IgH and Igκ/λ BCR data, subsequent analysis of the BCR sequences, was performed using the Immcantation 10x Genomics V(D)J sequence analysis workflow (https://immcantation.readthedocs.io/en/stable/tutorials/10x_tutorial.html). Briefly, V, D, and J were assigned using IgBLAST, then clonal groups were assigned based on IgH and Igκ/λ data, and germline sequences were reconstructed using Change-O[344] (v1.2.0). The frequency of mutations in the variable region of the *IGH* chain and CDR3 region relative to the relevant germline sequence were calculated using the R package SHazaM[344] (v1.1.0).

2.11.7 Statistical analysis

All statistical tests were performed using R[328] (v4.1.0). For analysis of survival data, univariate and multivariate analyses were performed using the cox proportional hazards model in the survival package (v3.4) in R. Details for the other statistical tests employed can be found in the relevant section of this chapter and the relevant figure legends. In the absence of multiple testing, a threshold of 0.05 was used to determine significance. In the case of multiple testing, p-value adjustment was performed using the Benjamini-Hochberg procedure[345] to control for the false discovery rate (FDR), with a threshold of 0.1 used to determine significance.

2.12 Figure generation

Most figures for this thesis were generated in R[328] (v4.1.0) using the package ggplot2[329] (v3.3.5). Other figures were prepared using either GraphPad Prism (v9.3.1), FlowJo[330] (v10.8.1), Adobe illustrator 2022, BioRender (BioRender.com), or a combination of them.

3 Autoantibody profiling to predict clinical outcome to immunotherapy

3.1 Rationale

Experimentally measuring defects in human B cell tolerance is complex[62]. The ideal approach involves studying the frequency and *in vitro* functional capability of autoreactive B cells, by means such as flow cytometry or single B cell culture, as has been done for SLE patients[62,346–349]. However, these approaches are limited by either the lack if autoantigen information and low throughput. They also require further consideration of which B cell subset(s) should be tested for autoreactivity, as autoreactivity in different subsets reflects inefficiency of different tolerance mechanisms.

Given that the ultimate consequence of inefficient B cell tolerance has been suggested to be an increase in circulating AAbs[158,349–351], measuring the number of AAbs, or the number of reactive antigens, from blood plasma is an alternative approach. Further, production of IgG AAbs can be considered a sign of antigen driven GC reactions and cooperation of B and T cells. It could therefore also act as a surrogate measurement for baseline immune activation rather than just B cell tolerance.

Indeed, studies of infectious disease, autoimmunity and cancer have shown how the profile of circulating AAbs can reflect the immune state and immune modulation of an individual[268,270,313,352]. For example, a study of patients infected with severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) found that a greater number of reactivities to exoproteome antigens was associated with increased disease severity[270]. Several of these reactive antigens had immunomodulatory functions, such as cytokines, and in some cases the corresponding reactive AAb was found to influence immune function[270].

However, production of AAbs is not always detrimental to the health of the host. In SLE, the presence of neutralising ACAAs to interferon alpha (IFNα) was linked with improved disease control which was suggested to be mediated through modulation of B cell differentiation[268]. Another example is a study that found that individuals infected with human immunodeficiency virus (HIV) who developed broadly

neutralising antibodies (bnAbs) were more likely to have clinically relevant autoimmune disease associated AAbs, like AAbs reactive against double stranded DNA (dsDNA)[352]. This suggests that greater AAb numbers may be reflective of more effective antibody responses against non-self-antigens. In cancer, a similar observation was made. TIF1-γ AAb seropositive DM patients, who are at a high-risk of developing cancer, were found to have a decreased risk of developing cancer if they had a greater number of reactivities to other antigens[313]. This suggests that a AAb profile which is reactive to a greater breadth of antigens may reduce cancer incidence in high-risk patients.

Aside from cancer risk, as previously discussed in section 1.3.4.3, AAbs have also been associated with risk of irAE development during ICI treatment[315,316,353,354]. However, all these studies focused on studying the association of overall seropositivity to at least one AAb rather than the number of AAbs or breadth of antigen reactivities. As development of irAEs is thought to be linked to baseline immune activation[139], studies of AAbs in ICI treated cancer patients that evaluate associations between the number of reactive antigens, using a similar approach to that of the SARS-CoV2 study mentioned above, with patient clinical outcomes would be valuable. Studies such as these would help to understand whether increased reactivity to self-antigens could be associated with reduced tolerance to tumour antigens. This could explain the link of irAE development with good clinical outcome to ICI treatment.

This chapter addresses my aim to establish whether AAb profiling can predict clinical outcome to ICI+CTX treatment in OC patients. I hypothesise that the number of reactive antigens associates with good clinical outcome, by perhaps acting as a marker of baseline immune activation and thus indicating an increased capacity to overcome immunological tolerance to TAAs and autoantigens that are highly expressed in the TME. The second aim of this chapter is to determine whether AAb profiling can non-invasively distinguish OC patients from BO patients which could have great clinical application for stratifying the endoscopic screening of individuals who are at high-risk of developing OC.

3.2 Autoantibody profiling of blood plasma from oesophageal cancer patients using two custom antigen panels

Multiple different antigen arrays to screen blood plasma for antigen reactivity have been previously described in the literature[355–358]. These vary in many factors including the number and breadth of antigens screened, sample throughput, source of antigen and detection method[355–358].

For this study, a platform capable of screening hundreds of plasma samples for reactivity against hundreds of antigens was required, so it was decided to use custom Luminex bead-based antigen arrays manufactured by the SciLifeLab (SLL) autoimmunity and serology profiling unit (KTH Royal Institute of Technology, Sweden)[359]. These are plate-based assays that can screen hundreds of samples for IgG reactivity against up to 380 different protein fragments. In addition to high throughput capabilities, there were two other major contributing factors for selection of this platform. The first of these was the comparable performance of the same Luminex technology to enzyme linked immunosorbent assay (ELISA) for detection of anti-pathogen antibodies, autoimmune disease associated AAbs and AAbs to TAAs[360–362]. The second was the ability of the data generated to be converted into binary serostatus data and subsequently the number of antigen reactivities per sample[363].

In total, two custom antigen panels were designed. The first antigen panel, which was used for the discovery cohort, was performed in July 2019. The second antigen panel, which was used for the validation cohort, was performed in August 2021. As there were over 20,000 antigens to select from, selection criteria were defined for each panel. These criteria are described in sections 3.3.1 and 3.3.2 respectfully.

3.2.1 Antigen panel 1: Reactivity to DNA- and RNA- binding proteins in oesophageal cancer

3.2.1.1 Rationale

Although theoretically any molecule could act as an autoantigen, some antigens are more likely to be sufficiently immunogenic to initiate an autoimmune response that is mature enough to produce IgG AAbs. Therefore, the selection criteria for the first antigen panel were based on existing literature on what makes a likely autoantigen, particularly in the context of cancer.

In healthy cells, many autoantigens are located intracellularly and so autoreactive B cells which evade immune tolerance mechanisms are not exposed to their cognate autoantigen under normal physiological conditions[364,365]. However, these antigens can be released into the extracellular environment during apoptosis and other forms of cell death[364,365]. In this scenario, the autoreactive B cells are then able to bind to their cognate antigen and become activated. During necrosis and other inflammatory cell death mechanisms, DAMPs such as DNA are released into the extracellular environment[44,46]. These DAMPs can then be taken up by B cells and provide stimulation by activating their TLRs[366]. Concurrent TLR activation by DAMPs and BCR activation by cognate antigen increases the likelihood of the B cell receiving sufficient stimulation to mature into an ASC[366]. The effect of cell death and TLR signalling on AAb production is demonstrated by the findings that patients with SLE have ineffective clearance of apoptotic cells[367]. Further, a single gain of function mutation in TLR7 is sufficient to cause SLE[368].

Mutations and high- or abnormal expression of proteins are also believed to increase their immunogenicity[364]. In cancer, mutations occur frequently, proteins are often over expressed, and there is known to be regular cell death, particularly by inflammatory mechanisms such as necrosis[66,369,370].

Hence, to determine whether circulating AAbs of OAC patients bound to nucleic acid binding proteins, overexpressed proteins, or proteins altered by mutation in OAC, we developed a custom high-throughput Luminex based screen (**Figure 3.1A**).

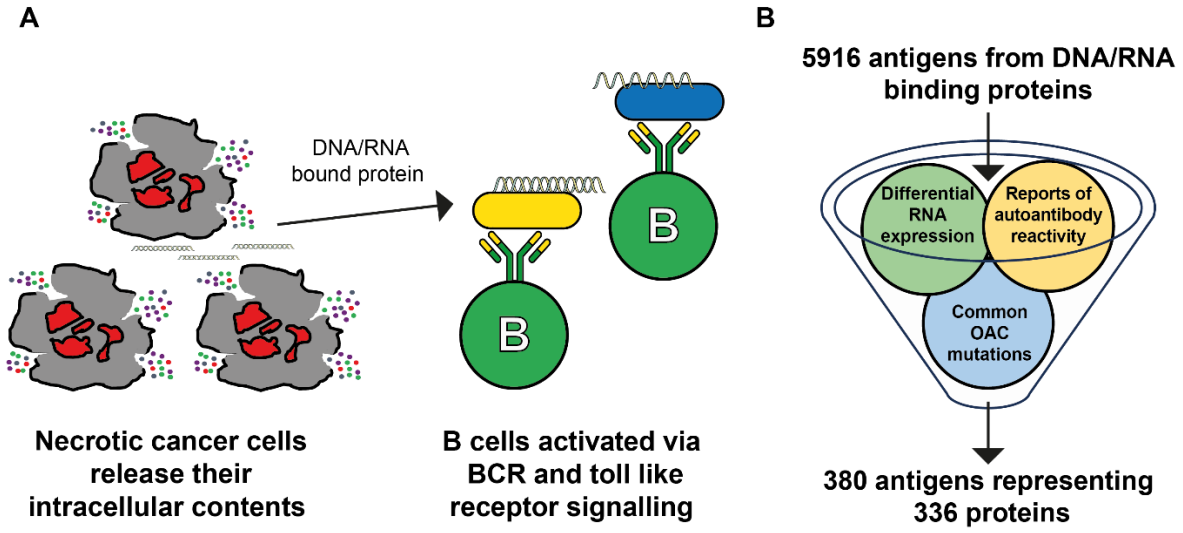


Figure 3.1 Antigen selection criteria for antigen panel 1 were based on DNA/RNA binding proteins being likely autoantigens in cancer. (**A**) Necrosis of cancer cells can release their intracellular contents, including transcription factors and other proteins bound by DNA or RNA. Extracellular DNA and RNA can stimulate B cells via toll like receptors (TLRs) 7 and 9. In combination with stimulation from BCR binding with antigen, TLR stimulation can strongly activate B cells. Based on this, DNA/RNA binding proteins are likely autoantigens in cancer. (**B**) The 5916 antigens derived from DNA/RNA binding proteins were refined to a final list of 380 antigens, derived from 336 different proteins, based on meeting at least one of the following criteria: reports of autoantibody reactivity, differential expression of RNA transcripts in a bulk RNA-seq dataset from a subset of 10 patients from the LUD2015-005 inoperable arm (analysed by Dr. Thomas Carroll), or reports of common mutation in OAC.

3.2.1.2 Antigen panel design

As there were 5916 antigens available which were reported as DNA or RNA binders on UniProt, this antigen list was further refined based on compliance with at least one of the following criteria (**Figure 3.4B**): 1) Reports of AAb reactivity, determined from either a review of previously published AAb studies in cancer or autoimmune patients, or from data from the SciLifeLab Autoimmunity Profiling facility database (provided by Dr. Ronald Sjöberg, unpublished); 2) Differential expression of RNA transcripts, from either tumours versus adjacent normal GI tissues, or tumours of irRECIST responders vs tumours of irRECIST non-responders, in a bulk RNA-seq dataset from a subset of 10 patients from the LUD2015-005 inoperable arm (analysed by Dr. Thomas Carroll, unpublished); and 3) Commonly

mutated genes in OAC[322]. This resulted in a list of 373 antigens which was combined with seven corresponding to one of the following immune checkpoint molecules: CTLA-4, PD-1, and PD-L1, to complete the panel of 380 antigens. Full details are described in chapter 2 section 2.3.1.

3.2.1.3 Sample selection for the discovery cohort screened against antigen panel 1

180 human plasma samples were screened for reactivity against this panel of antigens. 98 samples were from patients from the inoperable arm of the LUD2015-005 clinical trial, with 31 samples from the PreTx timepoint, 29 samples from the ICI-4W timepoint, 25 samples from the C1W3 timepoint, and 13 samples from the post-treatment timepoint. A further 65 samples were from patients from the operable arm of the LUD2015-005 clinical trial, with 21 samples from the PreTx timepoint, 19 samples from the ICI-4W timepoint, 16 samples from the C1W3 timepoint, and 8 samples from the post-treatment timepoint. The remaining 17 samples were comprised of; 12 samples from five BO patients, sampled at 2 to 4 timepoints; 4 samples from healthy volunteers; and 1 sample from a chemotherapy naïve operable OAC patient. Full details regarding these samples are described in chapter 2 sections 2.1 and 2.2.

For each of these 180 samples, raw mean fluorescent intensity (MFI) data corresponding to reactivity to each of the 380 antigens was generated and subsequently processed into binary serostatus data by Dr Ronald Sjöberg and his team at SciLifeLab. Full details regarding this data processing is described in section 2.3.1.

3.2.1.4 Pre-treatment number of reactive antigens are positively associated with clinical outcome

To explore whether AAb profiling could predict clinical outcome to ICI+CTX treatment in OC patients, binary serostatus data from the 29 PreTx OAC samples from the LUD2015-005 inoperable patients was analysed for associations with patient clinical outcomes. In particular, the number of reactive antigens

was explored as a potential surrogate measurement for a higher baseline immune activation and thus a possible increased likelihood to overcome tolerance to the tumour.

Using the binary serostatus data provided by SciLifeLab, the number of reactive antigens, out of a possible 380, were calculated for each of the 29 PreTx OAC samples. The corresponding data was then used to assess whether there was an association of number of reactive antigens with clinical outcome. This analysis showed that, on average, CB patients had a significant greater number of reactive antigens PreTx than NCB patients (**Figure 3.2A**). Similarly, on average, patients who developed an irAE during ICI treatment had a significant greater number of reactive antigens PreTx than patients who did not develop an irAE (**Figure 3.2B**). Moreover, patients who had a high number of reactive antigens (>median) showed significantly improved OS compared to patients who had a low number of reactive antigens (≤median) (**Figure 3.2C**). Subsequent multivariate cox regression analysis of OS data showed that the number of PreTx reactive antigens and development of an irAE during treatment were likely interacting covariates, and irAE development was the factor significantly associated with improved OS (**Figure 3.2D**).

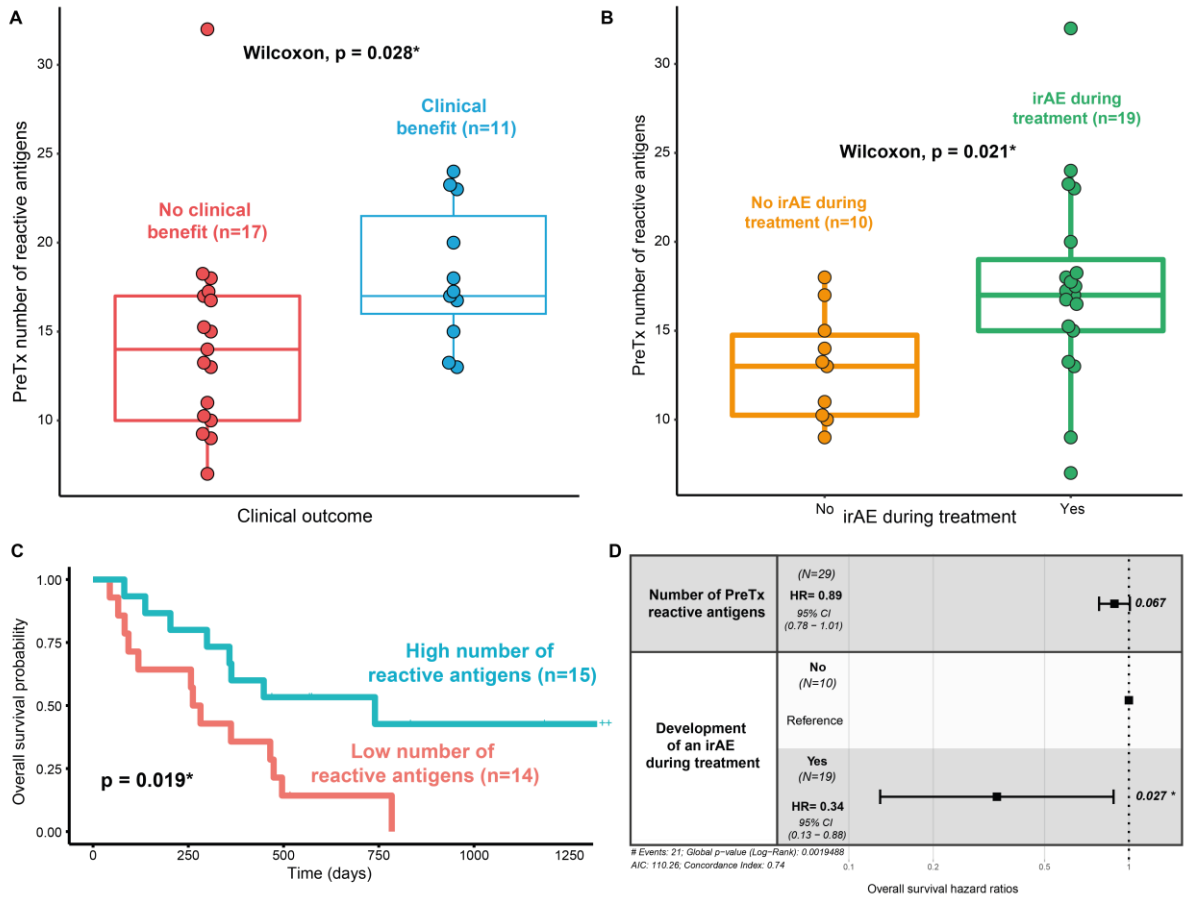


Figure 3.2. PreTx number of reactivities calculated from antigen panel 1 significantly associates with clinical benefit and development of an irAE in OAC patients of the LUD2015-005 cohort. (**A-B**) Boxplots showing the number of reactive antigens for the 28 PreTx plasma samples screened against 180 antigens in antigen panel 1, with samples compared between either clinical benefit (CB) and no clinical benefit patients (NCB) or (**B**) patients who did not develop an irAE during treatment and patients who did develop an irAE. One of the 29 patients could not be assigned a CB status and so this sample was excluded from CB vs NCB comparisons. Statistical comparisons of group medians were carried out using a Wilcoxon test. (**C**). Kaplan Meier plot of overall survival (OS) with patients split into those with a high number of reactive antigens (≥ median) (blue) and those who had a low number of reactive antigens (< median) (red). p value was calculated using the log rank method. (**D**) Forest plot showing the hazard ratios (HRs), 95% confidence interval (CIs) and p values for the association of number of PreTx reactive antigens and development of an irAE during treatment with OS, as calculated by multivariate regression analysis using the cox proportional hazards model in the survival package in R.

3.2.2 Antigen panel 2: Reactivity to immune associated and common autoantigens in oesophageal cancer, Barrett's oesophagus, and lung cancer

To validate the initial findings from analysis of the binary serostatus data from antigen panel 1, a second antigen panel was designed and additional plasma samples, including 95 from ICI treated lung cancer (LC), were sourced. To this end, 443 plasma samples, including 99 that were screened against antigen panel 1, and seven monoclonal antibodies (mAbs), synthesised from single-cell BCR sequencing data (described in section 5.5), were screened for their reactivity to the new antigen panel.

3.2.2.1 Antigen panel design

As most antigens included in antigen panel 1 showed no or minimal reactivity in the plasma samples screened, only the 85 antigens which showed reactivity in more than one patient sample were retained for the second antigen panel. The remaining 295 antigens were selected according to the following criteria: 1) Highly reactive, according to data from the SciLifeLab Autoimmunity Profiling facility database (provided by Dr. Ronald Sjöberg, unpublished); 2) Cytokines and immune related antigens which had shown some reactivity previously in the SciLifeLab Autoimmunity Profiling facility database (unpublished); 3) Related to antigens of interest identified from the DNA/RNA-binding protein focused antigen screen; 4) Surface antigens which had been reported as potential drug targets in OC; 5) Differentially expressed RNA transcripts from the full bulk RNAseq dataset of 33 PreTx tumour tissues of OAC patients from the LUD2015-005 inoperable arm (analysed by Dr. Thomas Carroll, unpublished); and 6) Autoantigens associated with autoimmune diseases. Although this completed a list of 380 antigens, 6 antigens (corresponding to IDH3G, PRTN3, TG, TOP1, TGM2, and MPO) were not available in the quantities required for screening the 450 samples so these antigens were not included in the panel, resulting in a final panel of 374 antigens.

3.2.2.2 Sample selection

443 human plasma samples and seven mAbs were screened for reactivity against antigen panel 2. The full list of plasma samples screened against this panel is as follows: 78 samples were from patients from the inoperable arm of the LUD2015-005 clinical trial, with 37 samples from the PreTx timepoint, 36 samples from the ICI-4W timepoint, and five samples taken at least 12 months after completion of treatment from patients who were classed as long-term responders (LTRs), based on them showing no clinical progression at least 18 months after treatment initiation. A further 67 samples were from patients from the operable arm of the LUD2015-005 clinical trial, with 34 samples from the PreTx timepoint and 33 samples from the ICI-4W timepoint. The remaining 308 samples were comprised of; 95 samples from ICI treated LC patients (provided by our collaborators Professor Christian

Ottensmeier and Dr. Ioannis Karydis, University of Southampton), 102 samples from BO patients, 10 samples from healthy volunteers (HVs) aged 50 years or older; 91 samples from operable OC, gastric cancer, colorectal cancer, or pancreatic cancer patients. Full details regarding these samples are described in chapter 2 sections 2.1 and 2.2.

Once again, binary serostatus data corresponding to antigen reactivity of these samples was calculated by Dr Ronald Sjöberg and his team at SciLifeLab, using the same approach as for antigen panel 1, as described in section 2.3.1.

3.2.2.3 Analysis of data from antigen panel 2 does not reproduce the findings from antigen panel 1

For antigen panel 2, the binary serostatus data provided by SciLifeLab was used to calculate the number of reactive antigens, out of a possible 374, for each of the 34 PreTx OAC samples. The corresponding data was then used to assess whether the significant association of the number of reactive antigens with clinical outcome, which was seen from analysis of antigen panel 1 data, was reproducible. However, this analysis did not show any significant association of number of reactive antigens with clinical outcome (**Figure 3.3**).

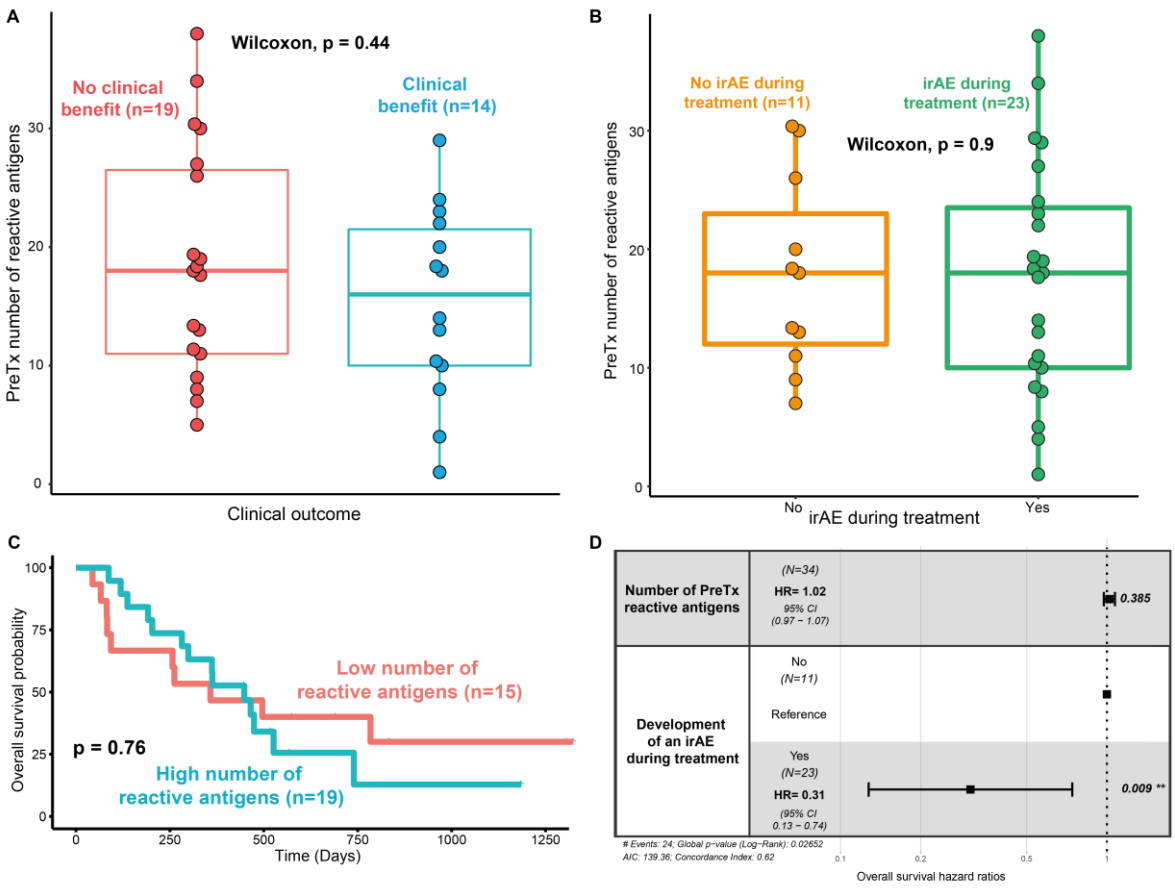


Figure 3.3. PreTx number of reactivities calculated from antigen panel 2 does not with clinical benefit and development of an irAE in OAC patients of the LUD2015-005 cohort. (**A-B**) Boxplots showing the number of reactive antigens for the 34 PreTx plasma samples screened against 180 antigens in antigen panel 1, with samples compared between either clinical benefit (CB) and no clinical benefit patients (NCB) or (**B**) patients who did not develop an irAE during treatment and patients who did develop an irAE. One of the 34 patients could not be assigned a CB status and so this sample was excluded from CB vs NCB comparisons. Statistical comparisons of group medians were carried out using a Wilcoxon test. (**C**). Kaplan Meier plot of overall survival (OS) with patients split into those with a high number of reactive antigens (≥ median) (blue) and those who had a low number of reactive antigens (< median) (red). p value was calculated using the log rank method. (**D**) Forest plot showing the hazard ratios (HRs), 95% confidence interval (CIs) and p values for the association of number of PreTx reactive antigens and development of an irAE during treatment with OS, as calculated by multivariate regression analysis using the cox proportional hazards model in the survival package in R.

Given that antigen panel 2 only shared 85 of its 374 antigens with antigen panel 1, and five extra LUD2015-005 inoperable PreTx OAC plasma samples were screened against these antigens, it is possible that the different antigen panels and samples screened was responsible for the lack of reproducibility. To determine if this was the case, the binary serostatus data from both antigen panels was subset to only the 85 antigens and 29 PreTx OAC samples which were screened against both antigen panels. The number of reactive antigens per sample, out of a possible 85, were then calculated

and the association of this measurement with CB was assessed. Although this analysis showed a significant association of number of reactive antigens with clinical outcome based on the antigen panel 1 data, when using the data from antigen panel 2, no such significant association was seen (**Figure 3.4**).

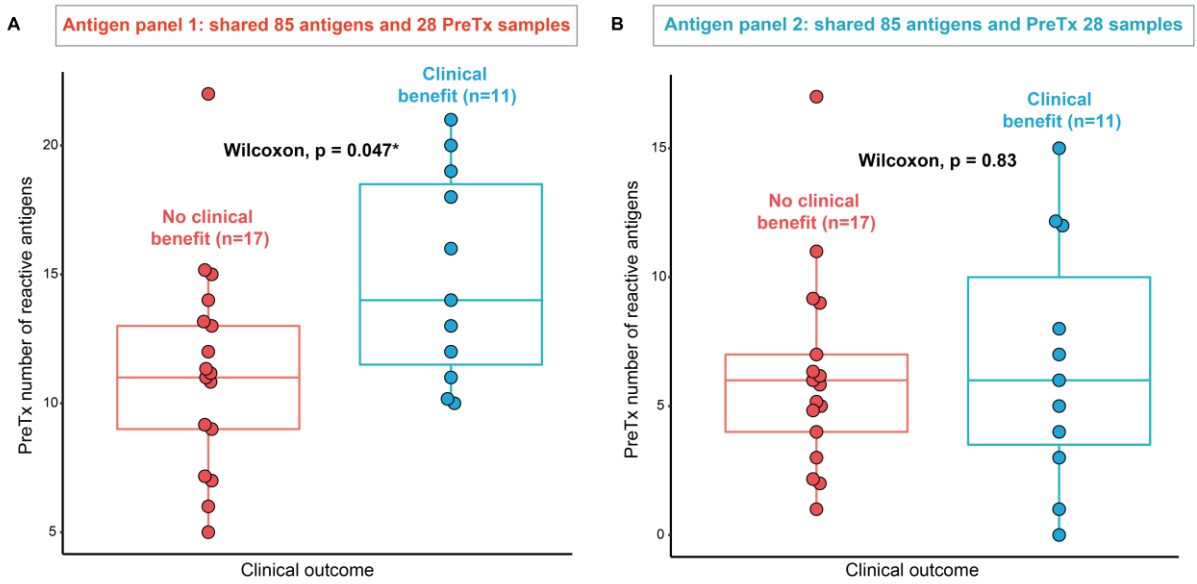


Figure 3.5 The association of PreTx number of reactivities and clinical benefit is not reproducible between the two antigen panels. Boxplots showing the number of reactive antigens between clinical benefit and no clinical benefit patients for the same 28 PreTx plasma samples screened against the same 85 antigens for (**A**) antigen panel 1 and (**B**) antigen panel 2. Statistical comparisons of group medians were carried out using a Wilcoxon test.

To further explore the reproducibility of the assay, the data from both antigen panels was subset to that corresponding to reactivity to the shared 85 antigens for either the commercial plasma samples and buffer-only samples (blanks), which were included as positive and negative controls, or the shared 29 PreTx OAC plasma samples. Clustering was then performed to determine whether clustering of samples was driven by the antigen panel that they originated from.

For the control samples, principal component analysis (PCA) of the binary serostatus data from both panels showed clear clustering of the positive control samples by antigen panel and that PC1 and PC2 were likely driven by a batch effect (**Figure 3.5A**). Similarly, unsupervised hierarchical clustering of the

same data in a heatmap showed clustering of control samples from the same antigen panel together, and that the positive control samples from antigen panel 2 were positive for fewer antigens than the same samples from antigen panel 1 (**Figure 3.5B**). Similar analysis of the 29 PreTx OAC plasma samples showed the same trends, with both PCA and hierarchical clustering of the binary serostatus data from both panels showing clustering of the plasma samples by antigen panel (**Figure 3.5C-D**).

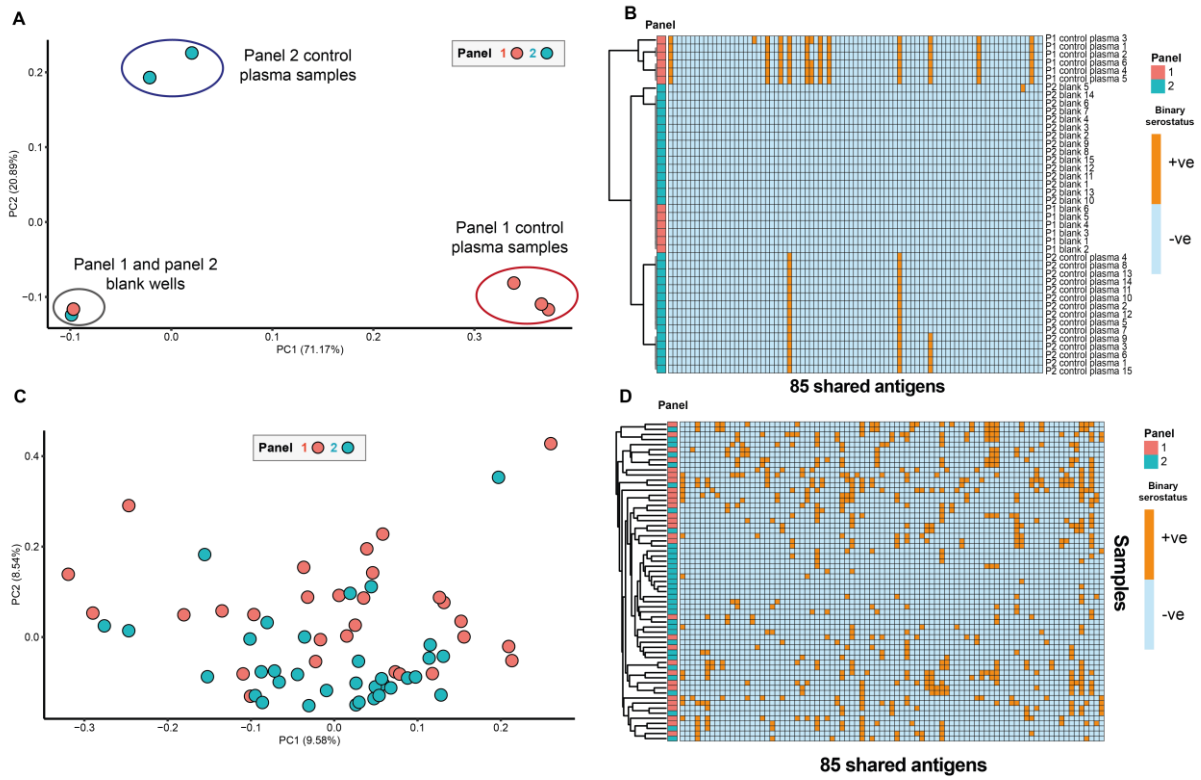


Figure 3.5. Binary serostatus data from antigen panel 1 and antigen panel 2 for the same plasma samples is not reproducible. (**A**) A PCA plot of binary serostatus data of the 85 antigens and control samples that were included on both panel 1 and panel 2, showing PC1 and PC2. One point per sample, per panel. (**B**) Heatmap of the same binary serostatus data with unsupervised hierarchical clustering of both antigens (columns) and control samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were run in panel 1 (red) or panel 2 (blue). (**C-D**) Same as (A and B) except for the 29 PreTx OAC plasma samples rather than control samples.

Given the poor reproducibility of binary serostatus data that was calculated using the approach suggested by SciLifeLab, an alternative analysis workflow which would be capable of accounting for batch variation and calculate highly reproducible binary serostatus data was developed.

3.3. Development of a novel analysis workflow which improves reproducibility of binary data

3.3.1 Rationale

The two most common approaches for analysing antigen reactivity data are the comparison of either normalised continuous data, such as net MFI which is used as a surrogate measurement of antibody titres, and binary serostatus data which stratifies the samples as either seronegative (unreactive to the antigen) or seropositive (reactive to the antigen).

However, as demonstrated in section 3.2.2.3, although antigen arrays allow for high throughput screening of samples against hundreds or thousands of antigens, due to their complexity, variable assay reproducibility has been reported[356,371–374]. An added complication is that most AAbs don't show a bimodal distribution, with clear positive and negative populations, and so lack dichotomy which makes robust assignment of seropositivity complicated[375,376]. To try and counter these issues, several groups studying anti-pathogen antibody responses have taken a finite mixture modelling (FMM) approach which uses the distribution of the data to assign samples into seronegative and seropositive subpopulations[377–379]. A notable advantage of FMM approaches over the conventional assignment of predetermined standardised cut-offs is the reduced bias involved in binary serostatus calls[379].

Subsequently, it was decided to utilise a similar approach for analysis of the data from the two antigen panels described above. Thus, a workflow based on Gaussian FMM using the R package mclust was developed and implemented to identify antigens which had a bimodal-like distribution, and hence would show a more reliable detection of positive reactivity, and then provide a distribution-based cut-off assignment for serostatus. This workflow was applied to the two antigen array datasets described in section 3.2.

3.3.2 A Gaussian finite mixture modelling workflow to assign binary serostatus

Firstly, the raw MFI data for each antigen and each sample were processed to net MFI data on a per sample basis by subtracting the MFI value for the bare bead negative control in the corresponding sample in order to adjust for sample specific background signal. The MFI signal for the beads coupled to only the His6ABP tag, which is bound to each antigen, were then plotted for each sample to identify any samples with a high signal and hence potentially incomplete depletion of anti-His6ABP reactive antibodies which should have been depleted prior to plasma incubation with the beads (details described in section 2.3.1). As each antigen had this His6ABP tag bound, it was essential to remove any samples with a high MFI signal to the His6ABP coupled bead to avoid false positive reactivities. Next, as the net MFI data showed a strong positive skew, the net MFI values were natural log (ln) transformed, and, on a per antigen basis, had two distributions fitted using the mclust package in R[327] (**Figure 3.6D**). Antigens which were deemed to show a bimodal-like distribution, determined from the proportion of overlap between the two distributions and the difference in net MFI between the means of the peaks of the two distributions fitted, were taken forward for further analysis (**Figure 3.6D**). Serostatus data were calculated as seronegative (0) if mclust classified the sample as most likely belonging in the lower distribution, or seropositive (1) if mclust classified the sample as most likely belonging in the upper distribution (**Figure 3.6**).

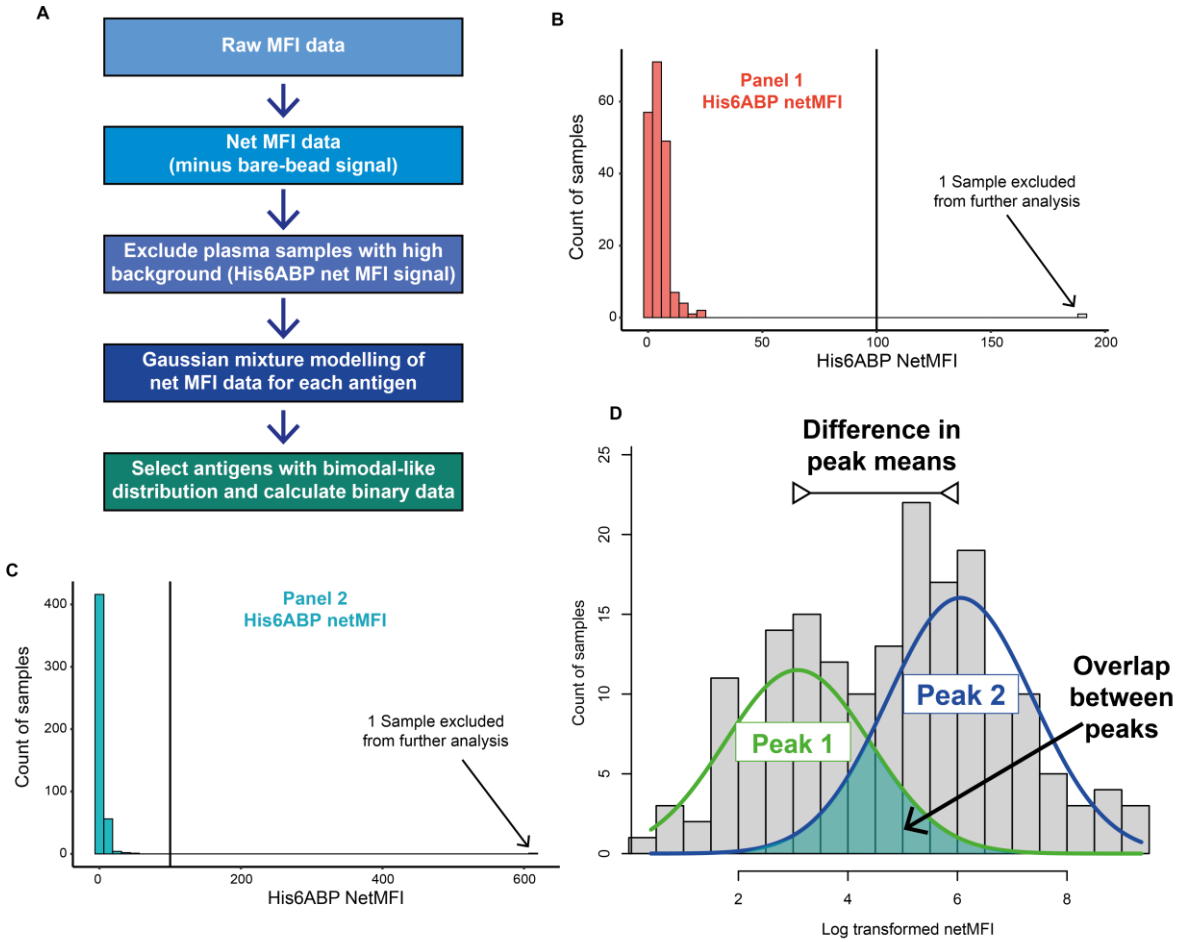


Figure 3.6. A workflow to process raw MFI data into net MFI and serostatus data using a Gaussian finite mixture modelling. (**A**) Workflow for processing the raw MFI data into net MFI and binary serostatus data for each antigen. Raw MFI data was transformed into net MFI data by subtracting the MFI signal to the bare-bead for each sample. Plasma sample(s) that had a high net MFI signal to the His6ABP coupled bead were excluded from downstream analysis. For each of the two panels, the net MFI values for the remaining samples were log transformed and, Gaussian mixture modelling was performed to fit two distributions to the data for each antigen. Binary serostatus data was calculated for the antigens which showed a bimodal-like distribution, and these were taken forward for further downstream analysis. (**B**) Histogram of the net MFI signal for the His6ABP bead for each sample in panel 1. The one sample with a net MFI signal of over 100 was excluded from downstream analysis. (**C**) Histogram of the net MFI signal for the His6ABP bead for each sample in panel 2. The one sample with a net MFI signal of over 100 was excluded from downstream analysis. (**D**) An example of Gaussian mixture modelling of net MFI for an antigen showing the peaks of the two distributions fitted. The overlap between peaks and difference in peak means were calculated and used to determine whether each antigen had a bimodal-like distribution or not.

When applying this workflow to all 180 samples from antigen panel 1, 1 plasma sample was excluded from downstream analysis for having a high net MFI signal to the His6ABP coupled bead (**Figure 3.6B**). To determine suitable cut-offs for both the proportion of overlap between the two peaks and the difference in the net MFI between the two peaks, a histogram was plotted and the cut-offs were set at the point at which there appeared to be a break in the distribution of the data. Using these cut-offs,

78 antigens were deemed to have a bimodal-like distribution with low overlap of peaks and large difference in peak means (**Figure 3.7**).

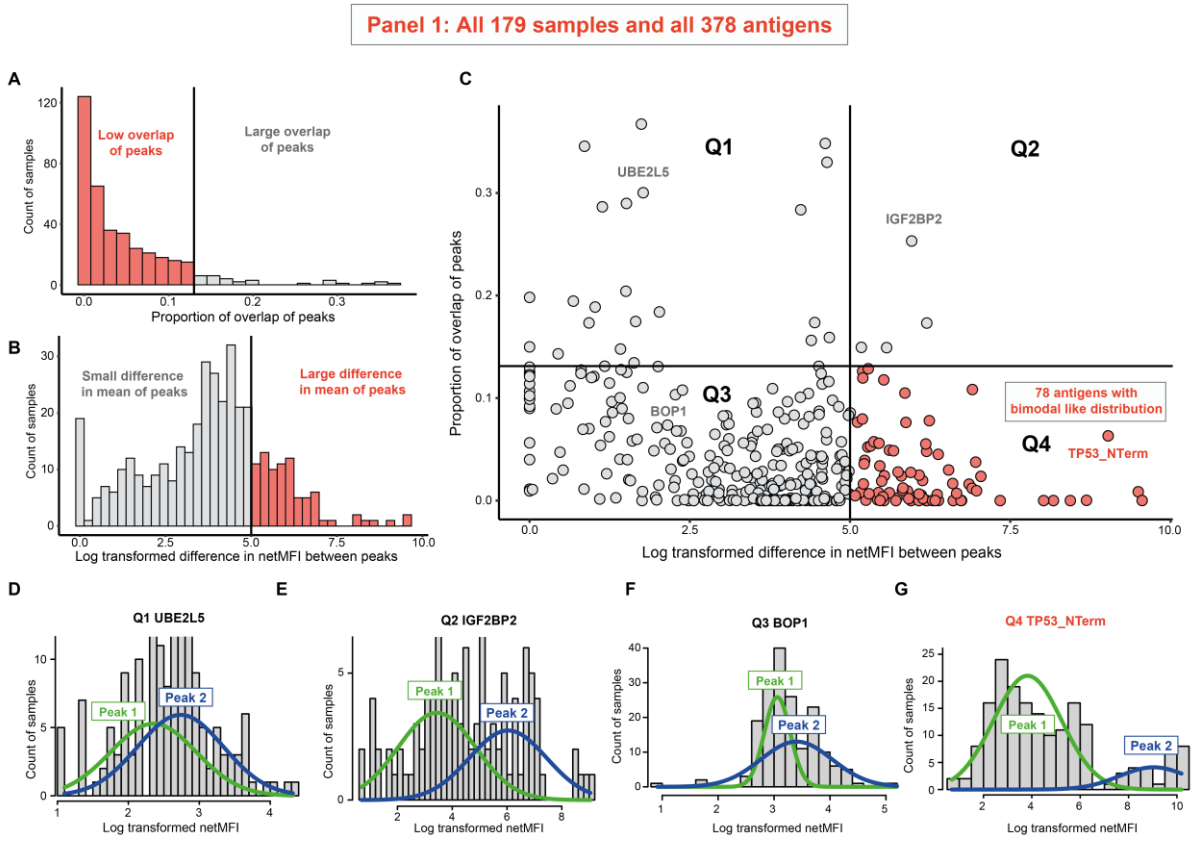


Figure 3.7. Identifying antigens with a bimodal-like distribution in antigen panel 1. (**A-B**) Histograms of (**A**) the proportion of overlap between peaks, and (**B**) the log transformed difference in mean net MFI of the peaks of the two distributions fitted by the Gaussian mixture modelling workflow previously described. On each plot a vertical line shows where a cut-off was set. (**C**) A scatterplot of proportion overlap of peaks and the log transformed difference in mean net MFI of the peaks. One point per antigen. The antigens were split into four quadrants, labelled Q1-4, using the cut-offs shown in (**A**) and (**B**). Q4 contained 78 antigens which had low overlap of peaks and large difference in peak means and so these antigens were deemed to have a bimodal-like distribution. (**D-G**) Representative plots of Gaussian mixture modelling of net MFI for: (**D**) UBE2L5, an antigen from Q1 which was deemed to have large overlap of peaks and a small difference in peak means; (**E**) IGF2BP2, an antigen from Q2 which was deemed to have large overlap of peaks and a large difference in peak means; (**F**) BOP1, an antigen from Q3 which was deemed to have low overlap of peaks and a small difference in peak means; and (**G**) TP53_NTerm (TP53 from amino acid 5 to 64), an antigen from Q4 which was deemed to have low overlap of peaks and a large difference in peak means, and therefore a bimodal-like distribution.

Similarly, when applying this workflow to all 450 samples from antigen panel 2, 1 plasma sample was excluded from downstream analysis for having a high netMFI signal to the His6ABP coupled bead (**Figure 3.6C**). As for antigen panel 1, histograms of the proportion of overlap between the two peaks and the difference in the net MFI between the two peaks were plotted and cut-offs were set at the

point at which there appeared to be a break in the distribution of the data. Using these cut-offs, 172 antigens were deemed to have a bimodal-like distribution with low overlap of peaks and large difference in peak means (**Figure 3.8**).

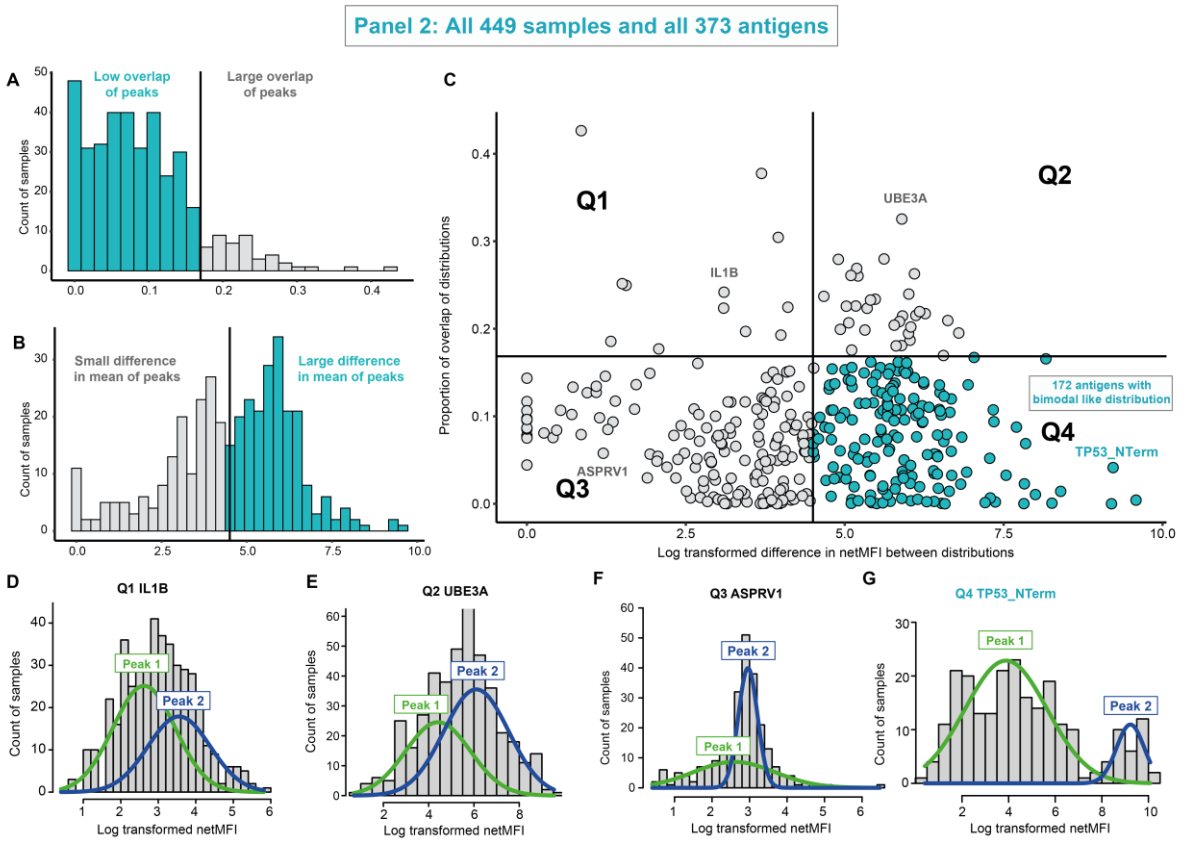


Figure 3.8. Identifying antigens with a bimodal-like distribution in antigen panel 2. (**A-B**) Histograms of (**A**) the proportion of overlap between peaks, and (**B**) the log transformed difference in mean net MFI of the peaks of the two distributions fitted by the Gaussian mixture modelling workflow previously described. On each plot a vertical line shows where a cut-off was set. (**C**) A scatterplot of proportion overlap of peaks and the log transformed difference in mean net MFI of the peaks. One point per antigen. The antigens were split into four quadrants, labelled Q1-4, using the cut-offs shown in (**A**) and (**B**). Q4 contained 172 antigens which had low overlap of peaks and large difference in peak means and so these antigens were deemed to have a bimodal-like distribution. (**D-G**) Representative plots of Gaussian mixture modelling of net MFI for: (**D**) IL1B, an antigen from Q1 which was deemed to have large overlap of peaks and a small difference in peak means; (**E**) UBE3A, an antigen from Q2 which was deemed to have large overlap of peaks and a large difference in peak means; (**F**) ASPRV1, an antigen from Q3 which was deemed to have low overlap of peaks and a small difference in peak means; and (**G**) TP53_NTerm (TP53 from amino acid 5 to 64), an antigen from Q4 which was deemed to have low overlap of peaks and a large difference in peak means, and therefore a bimodal-like distribution.

3.3.3 Binary serostatus data generated by the FMM workflow is reproducible across panels

To assess reproducibility of the data output from the two antigen panels, for the 99 plasma samples that were run on both panels, the net MFI and binary serostatus data were subset to only the antigens which were included in both antigen panels and were deemed to have a bimodal-like distribution, as described in section 3.3.1. This resulted in two sets of net MFI and binary serostatus data of 32 antigens for 99 plasma samples.

PCA of the ln transformed net MFI data for these 32 antigens and 99 plasma samples that were included on both antigen panels showed clear clustering of samples by antigen panel and that PC1 was likely driven by a batch effect (**Figure 3.9A**). Similarly, unsupervised hierarchical clustering of the same data in a heatmap showed clustering of samples from the same antigen panel together (**Figure 3.9B**).

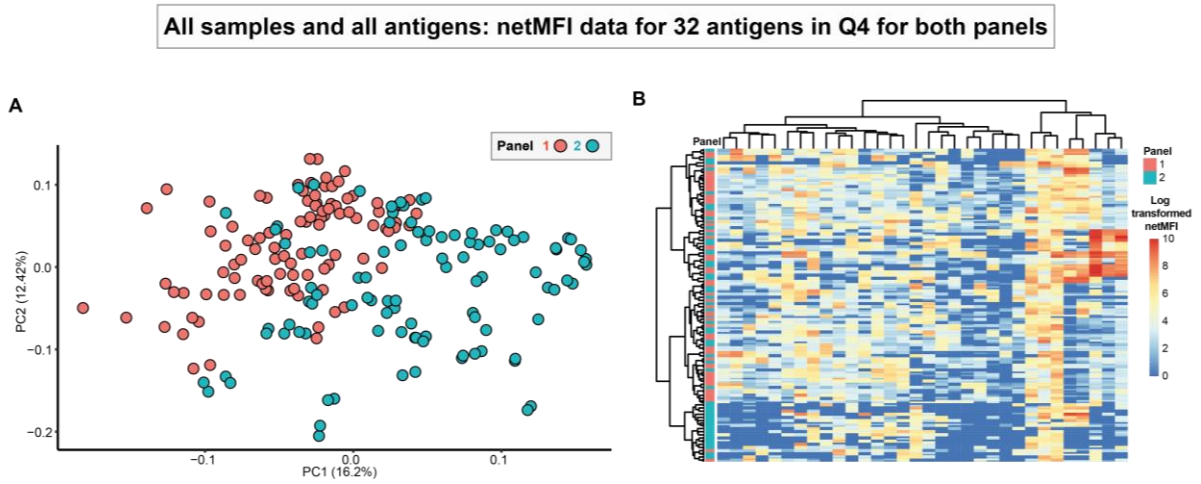


Figure 3.9. Net MFI data for the same plasma samples is highly variable between antigen panel 1 and antigen panel 2. (**A**) A PCA plot of log transformed net MFI data for the 32 antigens and 99 plasma samples that were included on both panel 1 and panel 2, showing PC1 and PC2. One point per plasma sample, per panel (198 points in total). (**B**) Heatmap of the same log transformed net MFI data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were run in panel 1 (red) or panel 2 (blue).

In contrast, the binary serostatus data for the 32 antigens and 99 plasma samples that were included on both antigen panels did not show clear clustering of samples by panel when analysed by PCA or when visualised on a heatmap with unsupervised hierarchical clustering (**Figure 3.10**).

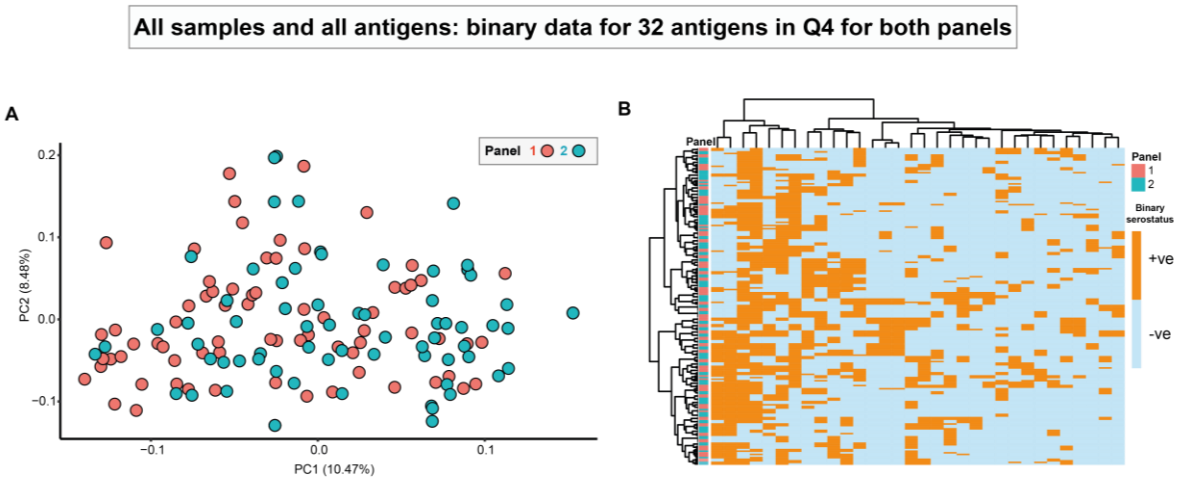


Figure 3.10. Binary serostatus data for the same plasma samples is not highly variable between antigen panel 1 and antigen panel 2. (**A**) A PCA plot of binary serostatus data of the 32 antigens and 99 plasma samples that were included on both panel 1 and panel 2, showing PC1 and PC2. One point per plasma sample, per panel (198 points in total). (**B**) Heatmap of the same binary serostatus data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were run in panel 1 (red) or panel 2 (blue).

To further assess the reproducibility of the binary serostatus data, for the data for each antigen panel, the number of reactive plasma samples for each antigen, and the number of reactive antigens for each plasma sample were both calculated. Most antigens showed a similar number of reactive plasma samples in both antigen panels, and there was a significant correlation between the total number of reactive antigens for the plasma samples screened against both antigen panels (**Figure 3.11**). However, there were some plasma samples which showed a very different number of reactive antigens on each panel.

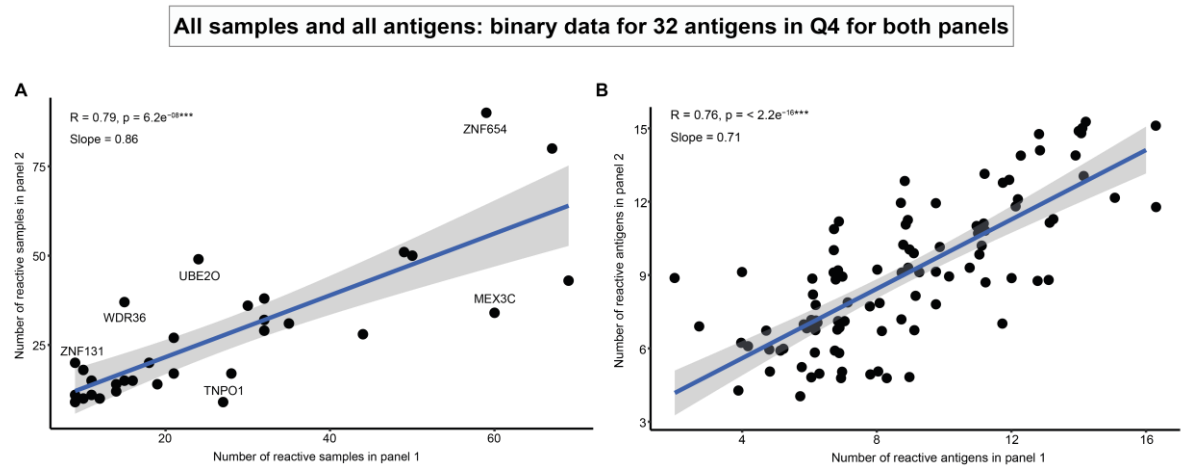


Figure 3.11. Most antigens and plasma samples show a similar number of reactivities between antigen panel 1 and antigen panel 2. (**A**) Scatterplot of number of reactive plasma samples, out of the 99 which were included in both antigen panel 1 and antigen panel 2, in antigen panel 1 and antigen panel 2. One point per antigen. 32 antigens in total. The 5 antigens which showed the largest discrepancy between the two antigen panels are labelled. (**B**) Scatterplot of total number of reactive antigens, out of a possible 32, for each of the 99 plasma samples included in both antigen panels. One point per sample per panel, 198 points in total. The R value, slope and p value for each plot was calculated by fitting a linear model using Pearson correlation method. Shaded area indicates the 95% confidence interval for the straight line fitted.

3.3.4 Reproducibility of binary data is improved when modelling on just the LUD2015-005 cohort

As the Gaussian finite mixture modelling workflow described in section 3.3.1 used a distribution-based cut-off assignment for seropositivity, it is possible that the different number and type of plasma samples included in each antigen panel influenced the robustness of binary serostatus calls. Therefore, to understand whether modelling on the same plasma samples would improve reproducibility of binary serostatus calls, the same workflow was applied to only the 99 samples and 85 antigens that were included on both antigen panels. Going forward, these 99 samples and 85 antigens are referred to as the LUD2015-005 cohort, due to all the plasma samples originating from the LUD2015-005 trial.

For these 99 samples and 85 antigens, cut-offs for the proportion of overlap between peaks and difference in peak means were established for antigen panel 1 (**Figure 3.12A-B**) and antigen panel 2 (**Figure 3.13A-B**), using the same approach as described in section 3.3.1. Using these cut-offs, 47 antigens were deemed to have a bimodal-like distribution with low overlap of peaks and large difference in peak means for antigen panel 1 (**Figure 3.12C**). For antigen panel 2, 33 antigens were

deemed to have a bimodal-like distribution with low overlap of peaks and large difference in peak means (**Figure 3.13C**).

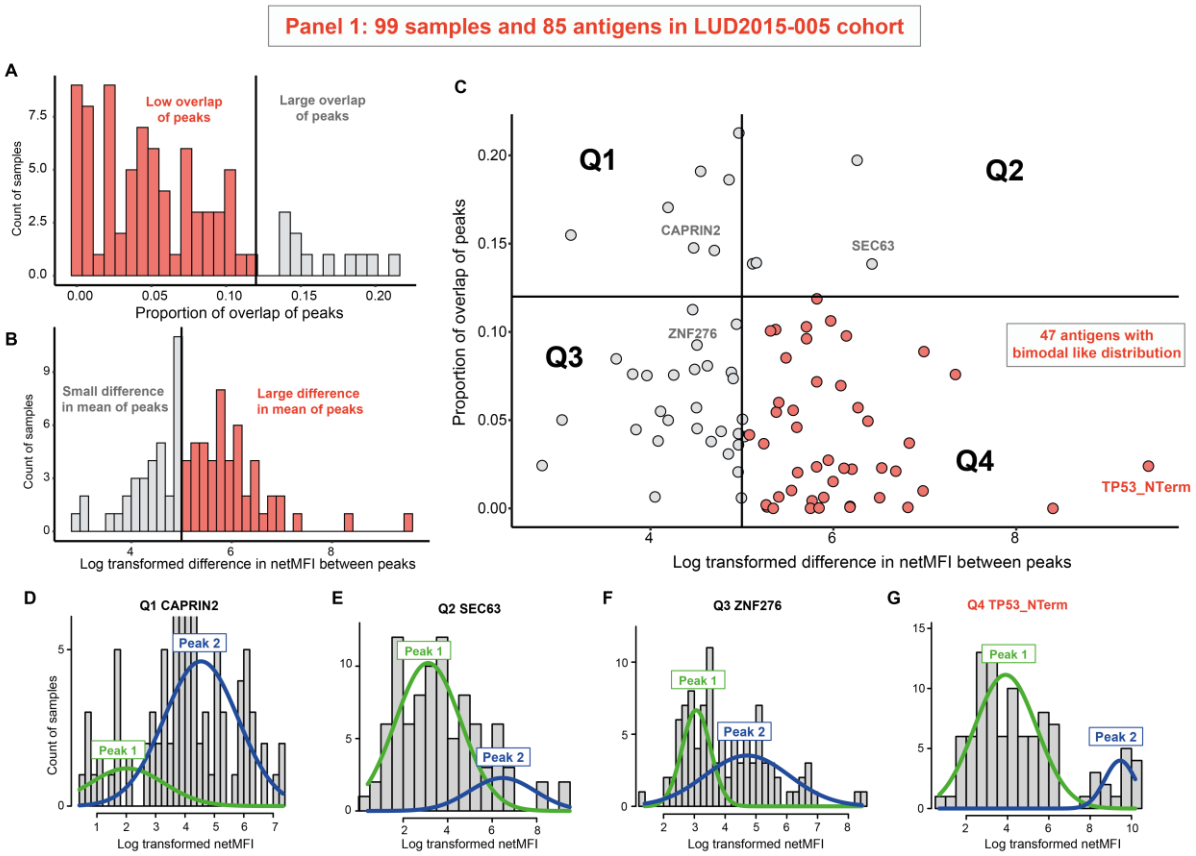


Figure 3.12. Identifying antigens with a bimodal-like distribution in antigen panel 1 when modelling on the LUD2015-005 cohort. (**A-B**) Histograms of (**A**) the proportion of overlap between peaks, and (**B**) the log transformed difference in mean net MFI of the peaks of the two distributions fitted by the Gaussian mixture modelling workflow previously described. On each plot a vertical line shows where a cut-off was set. (**C**) A scatterplot of proportion overlap of peaks and the log transformed difference in mean net MFI of the peaks. One point per antigen. The antigens were split into four quadrants, labelled Q1-4, using the cut-offs shown in (**A**) and (**B**). Q4 contained 47 antigens which had low overlap of peaks and large difference in peak means and so these antigens were deemed to have a bimodal-like distribution. (**D-G**) Representative plots of Gaussian mixture modelling of net MFI for: (**D**) CAPRIN2, an antigen from Q1 which was deemed to have large overlap of peaks and a small difference in peak means; (**E**) SEC63, an antigen from Q2 which was deemed to have large overlap of peaks and a large difference in peak means; (**F**) ZNF276, an antigen from Q3 which was deemed to have low overlap of peaks and a small difference in peak means; and (**G**) TP53_NTerm (TP53 from amino acid 5 to 64), an antigen from Q4 which was deemed to have low overlap of peaks and a large difference in peak means, and therefore a bimodal-like distribution.

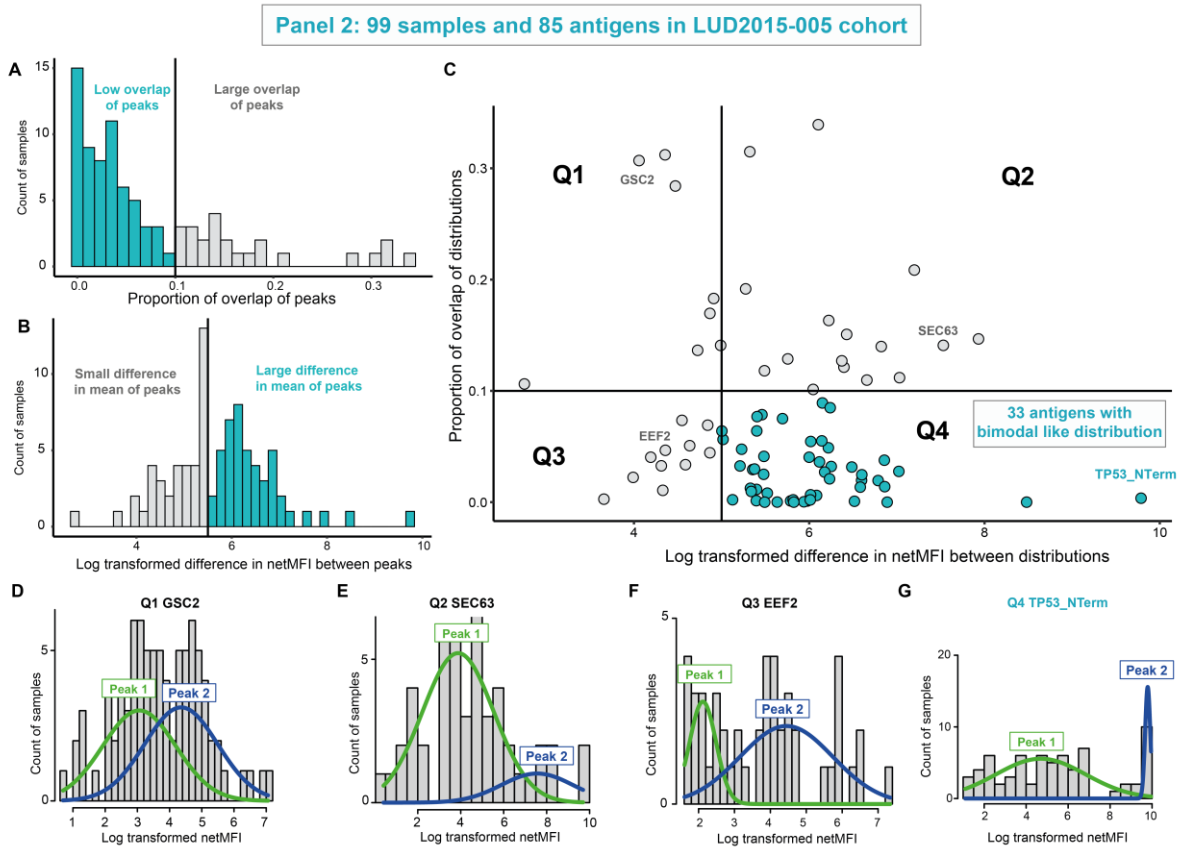


Figure 3.13. Identifying antigens with a bimodal-like distribution in antigen panel 2 when modelling on the LUD2015-005 cohort. (**A-B**) Histograms of (**A**) the proportion of overlap between peaks, and (**B**) the log transformed difference in mean net MFI of the peaks of the two distributions fitted by the Gaussian mixture modelling workflow previously described. On each plot a vertical line shows where a cut-off was set. (**C**) A scatterplot of proportion overlap of peaks and the log transformed difference in mean net MFI of the peaks. One point per antigen. The antigens were split into four quadrants, labelled Q1-4, using the cut-offs shown in (**A**) and (**B**). Q4 contained 33 antigens which had low overlap of peaks and large difference in peak means and so these antigens were deemed to have a bimodal-like distribution. (**D-G**) Representative plots of Gaussian mixture modelling of net MFI for: (**D**) GSC2, an antigen from Q1 which was deemed to have large overlap of peaks and a small difference in peak means; (**E**) SEC63, an antigen from Q2 which was deemed to have large overlap of peaks and a large difference in peak means; (**F**) EEF2, an antigen from Q3 which was deemed to have low overlap of peaks and a small difference in peak means; and (**G**) TP53_NTerm (TP53 from amino acid 5 to 64), an antigen from Q4 which was deemed to have low overlap of peaks and a large difference in peak means, and therefore a bimodal-like distribution.

Overall, LUD2015-005 cohort, 24 antigens were deemed to have a bimodal-like distribution in both antigen panels. Reproducibility of net MFI data and binary serostatus data for these 24 antigens in the LUD2015-005 samples was then assessed. PCA of the ln transformed net MFI data showed clear clustering of samples by antigen panel and that PC1 was likely driven by a batch effect (**Figure 3.14A**). Similarly, unsupervised hierarchical clustering of the same data in a heatmap showed clustering of samples from the same antigen panel close together (**Figure 3.14B**).

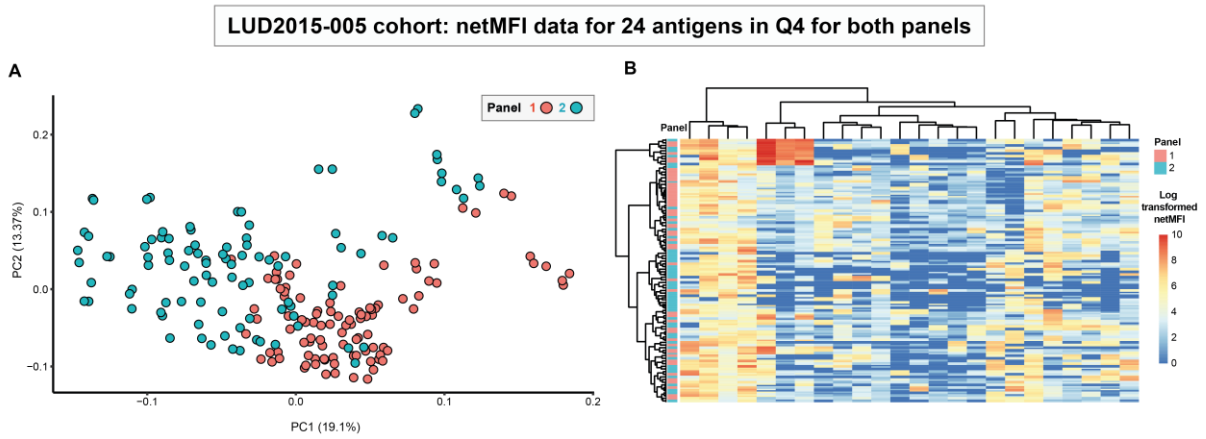


Figure 3.14. Net MFI data for the same plasma samples is highly variable between antigen panel 1 and antigen panel 2 when modelling on the LUD2015-005 cohort. (**A**) A PCA plot of log transformed net MFI data for the 24 antigens and 99 plasma samples that were included on both panel 1 and panel 2, showing PC1 and PC2. One point per plasma sample, per panel (198 points in total). (**B**) Heatmap of the same log transformed net MFI data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were run in panel 1 (red) or panel 2 (blue).

However, once again, the binary serostatus data for the same 24 antigens and 99 plasma samples did not show clear clustering of samples by panel when analysed by PCA or when visualised on a heatmap with unsupervised hierarchical clustering (**Figure 3.15**).

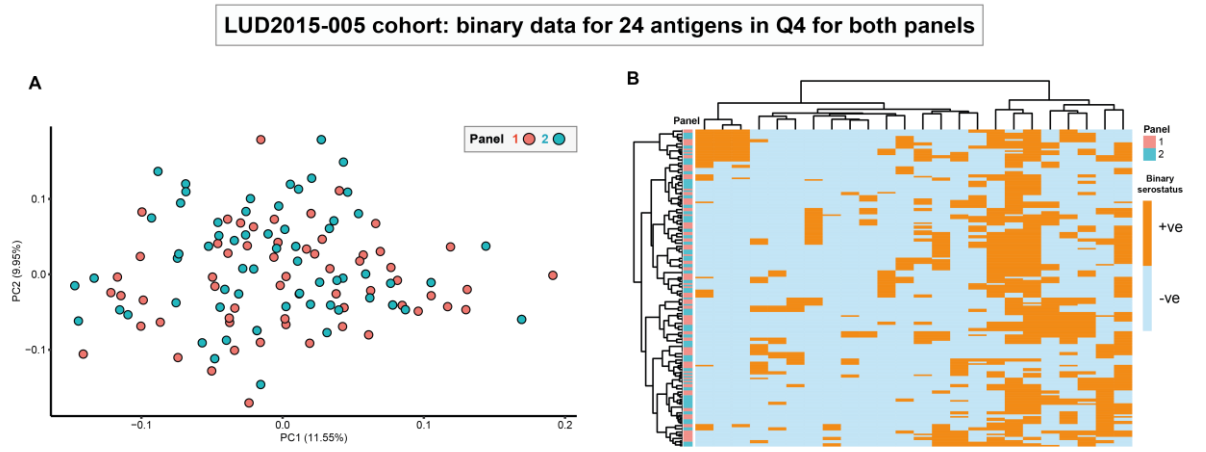


Figure 3.15. Binary serostatus data for the same plasma samples is not highly variable between antigen panel 1 and antigen panel 2 when modelling on the LUD2015-005 cohort. (**A**) A PCA plot of binary serostatus data of the 24 antigens and 99 plasma samples that were included on both panel 1 and panel 2, showing PC1 and PC2. One point per plasma sample, per panel (198 points in total). (**B**) Heatmap of the same binary serostatus data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were run in panel 1 (red) or panel 2 (blue).

To further assess the reproducibility of the binary serostatus data, for the data for each antigen panel, the number of reactive plasma samples for each antigen were calculated, and the number of reactive antigens for each plasma sample were calculated. Like the reproducibility analysis from section 3.3.2, most antigens showed a similar number of reactive plasma samples in both antigen panels, and there was a significant correlation between the total number of reactive antigens for the plasma samples screened against both antigen panels (**Figure 3.16**). However, once again, there were some plasma samples which showed a very different number of reactive antigens on each panel.

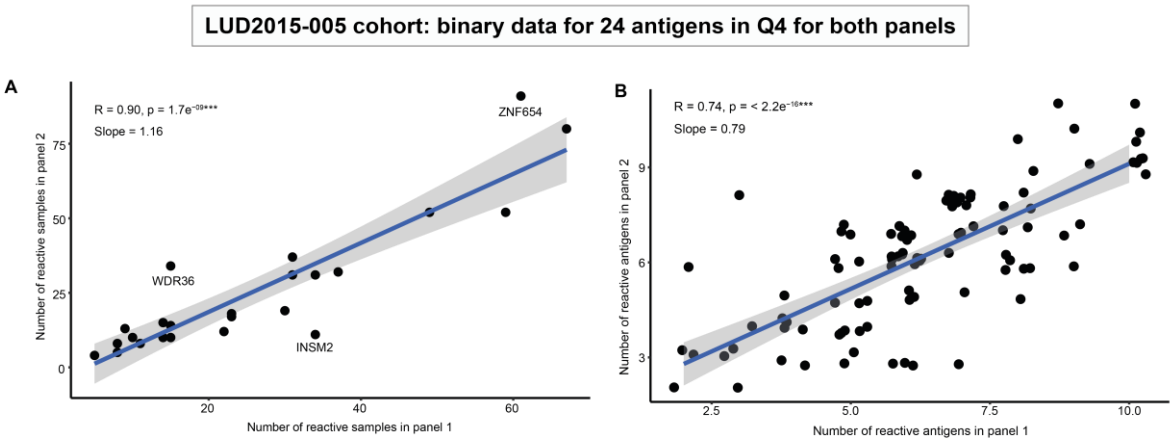


Figure 3.16. Most antigens and plasma samples show a similar number of reactivities between antigen panel 1 and antigen panel 2 when modelling on the LUD2015-005 cohort. (**A**) Scatterplot of number of reactive plasma samples, out of the 99 which were included in both antigen panel 1 and antigen panel 2, in antigen panel 1 and antigen panel 2. One point per antigen. 24 antigens in total. The 5 antigens which showed the largest discrepancy between the two antigen panels are labelled. (**B**) Scatterplot of total number of reactive antigens, out of a possible 24, for each of the 99 plasma samples included in both antigen panels. One point per sample per panel, 198 points in total. The R value, slope and p value for each plot was calculated by fitting a linear model using Pearson correlation method. Shaded area indicates the 95% confidence interval for the straight line fitted.

As the binary serostatus data for antigens and samples from both antigen panels were found to be largely reproducible, whereas the ln transformed net MFI data showed a clear batch effect, only the binary serostatus data was taken forward for further analyses. Although the reproducibility of the binary serostatus data was similar when modelling on all samples and when modelling on just the LUD2015-005 cohort, there appeared to be less of a visible batch effect in the binary data from the LUD2015-005 cohort. Therefore, the data was divided into two cohorts for further downstream analysis (**Figure 3.17**). The binary serostatus data from both panels for the 24 antigens and 99 samples

in the LUD2015-005 cohort, described in section 3.3.3, was taken forward for further analysis relating to clinical outcome of the LUD2015-005 trial patients. Whereas the binary serostatus data for all antigens from antigen panel 2 were taken forward for the BO, OC, and LC samples as an exploratory cohort to identify associations with clinical parameters which would need assessing for reproducibility and independently validating.

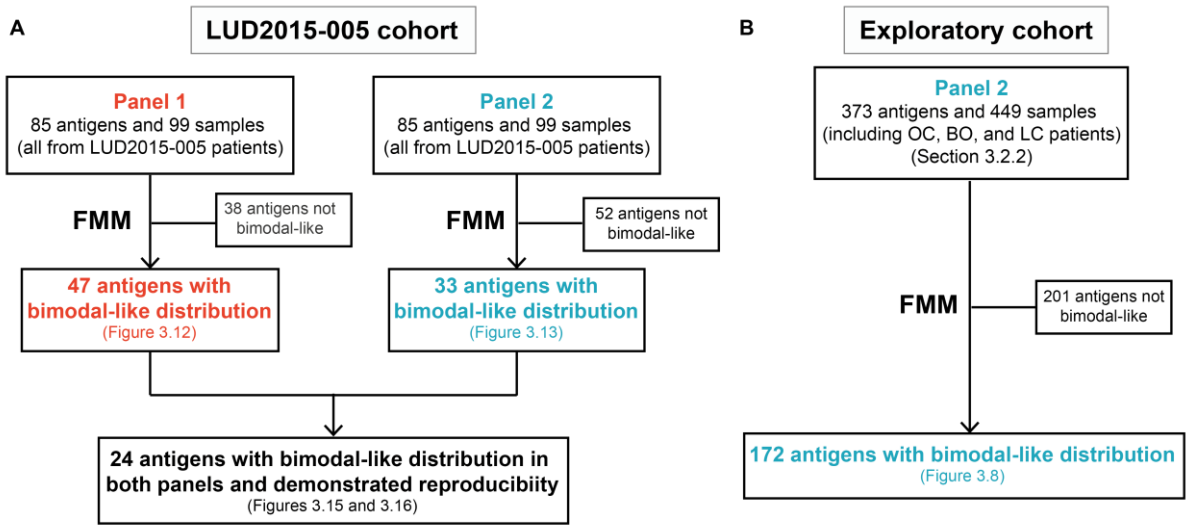


Figure 3.17. Summary of the LUD2015-005 and Exploratory cohorts used for downstream analysis. Following application of the finite mixture modelling (FMM), data was divided into two analysis cohorts, the LUD2015-005 cohort and the Exploratory cohort, for further downstream analysis. In the LUD2015-005 cohort, the binary serostatus data from both antigen panels for the 99 samples and 24 antigens which showed a bimodal-like distribution in both panels were taken forward for further analysis relating to clinical outcome of the LUD2015-005 trial patients. In the Exploratory cohort, the binary serostatus data for the 172 antigens from antigen panel 2 which had a bimodal-like distribution were taken forward for the BO, OC, and LC samples to identify associations with clinical parameters which would need assessing for reproducibility and independently validating.

3.4. Autoantibodies and clinical outcome in the LUD2015-005 cohort

3.4.1 Rationale

After establishing a workflow to generate reproducible binary serostatus data, the intent to comprehensively explore whether AAb profiling could predict clinical outcome to ICI+CTX treatment in OC patients was revisited. Subsequently, serostatus data from the LUD2015-005 cohort, from both the PreTx and ICI-4W timepoints, were analysed for associations with patient clinical outcomes. Once again, the number of reactive antigens was explored as a potential surrogate measurement for a higher baseline immune activation and thus a possible increased likelihood to overcome tolerance to the tumour. Serostatus data for individual antigens were also assessed to identify any differentially reactive antigens which, if independently validated, could have clinical utility for stratifying treatment of OC patients.

3.4.2 Associations between the total number of reactive antigens and clinical outcome

3.4.2.1 The pre-treatment number of reactive antigens shows a non-significant negative association with change in tumour size following four weeks of ICI treatment

Using the binary serostatus data derived from both antigen panel 1 and antigen panel 2, the number of reactive antigens, out of a possible 24, were calculated for each of the 29 PreTx OAC samples. The corresponding data from both antigen panels were then used to assess whether there was an association with clinical outcome. This analysis showed that there was no significant difference in the PreTx number of reactive antigens between CB patients and NCB patients (**Figure 3.18**).

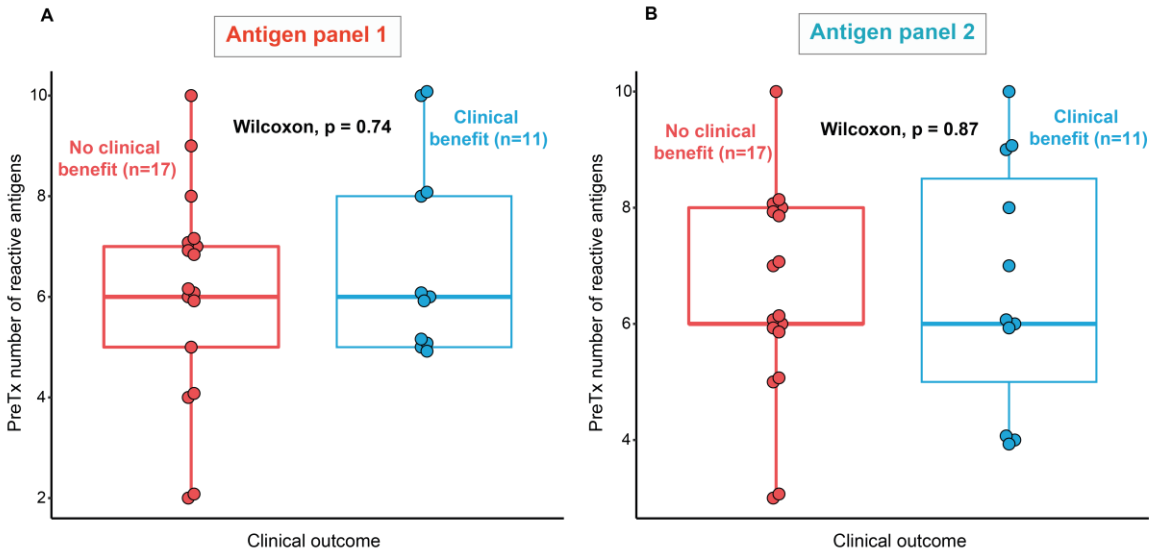


Figure 3.18. PreTx number of reactivities and clinical benefit in the LUD2015-005 cohort. Boxplots showing the number of reactive antigens between clinical benefit and no clinical benefit for the same 28 PreTx plasma samples screened against the same 24 antigens for (**A**) antigen panel 1 and (**B**) antigen panel 2. Statistical comparisons of group medians were carried out using a Wilcoxon test.

Given the original hypothesis that the PreTx number of reactive antigens could be a measure of greater baseline immune activation, and CTX treatment has been suggested to be detrimental to immune cells[380,381], it is possible that the number of reactive antigens is more relevant to patient outcome to ICI treatment alone, rather than the combination ICI+CTX treatment that the LUD2015-005 patients receive for most of their therapy. To test this, correlation analysis between the PreTx number of reactive antigens, from both antigen panels, and change in tumour size over the four-week period of treatment with ICI only was performed. For both antigen panels there was a trend towards an association between number of reactive antigens and decrease in tumour size (**Figure 3.19**).

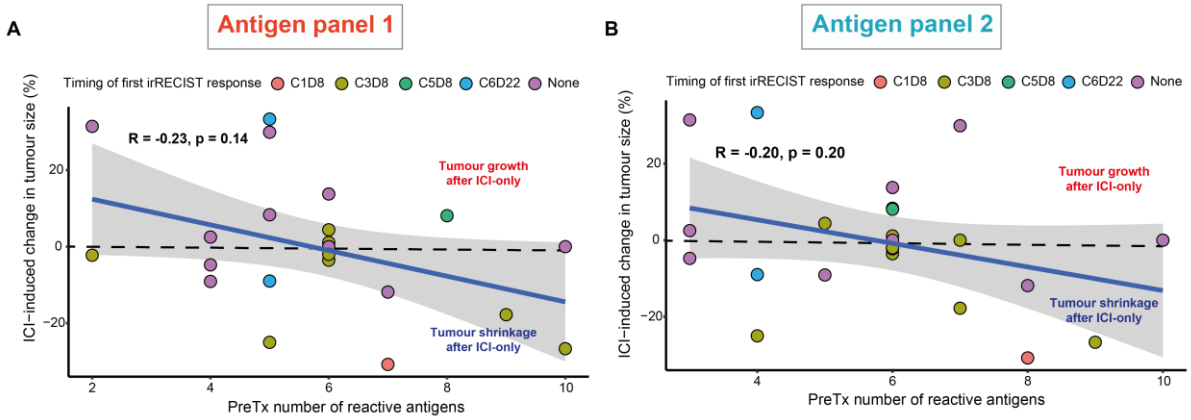


Figure 3.19. PreTx number of reactivities and ICI-induced tumour shrinkage in the LUD2015-005 cohort. Correlation plot of ICI-induced change in tumour size and number of reactive antigens PreTx for each of the 22 OAC patients who had assessable tumour lesions before and after the initial four weeks of ICI treatment, and who had AAb profiling performed as part of the LUD2015-005 cohort. (**A**) Correlation plot using the number of reactive antigens from antigen panel 1 (**B**) Correlation plot using the number of reactive antigens from antigen panel 2. One point per patient. Points are coloured based on the time at which the corresponding patient experienced their first response according to the irRECIST criteria, with C indicating the cycle number of CTX and D indicating the day during that cycle at which tumour size was assessed. Dashed line indicates zero change in tumour size. Shaded area indicates the 95% confidence interval for the line fitted (blue). R and p values for each correlation analysis were calculated by fitting a linear model with the Kendall correlation method.

Given this non-significant trend with ICI-induced tumour shrinkage, in combination with the association of irAE development with good clinical outcomes in the LUD2015-005 cohort, the potential link of PreTx number of reactive antigens with on-treatment development of an irAE was assessed. However, the data from both antigen panels showed no significant difference in the PreTx number of reactive antigens between patients who developed an irAE during treatment and those who did not develop an irAE (**Figure 3.20**).

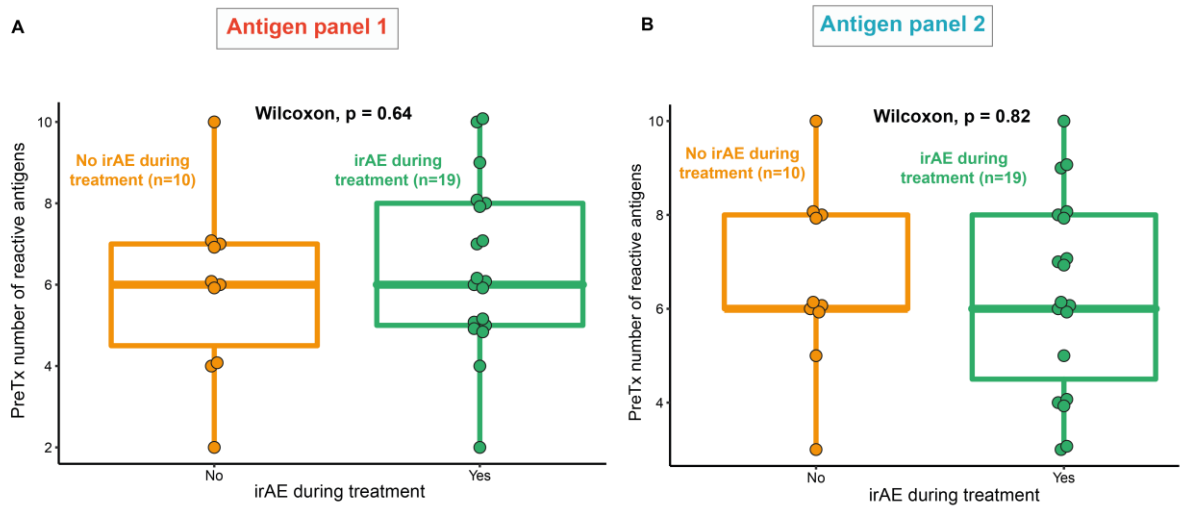


Figure 3.20. PreTx number of reactivities and irAE development in the LUD2015-005 cohort. Boxplots showing the number of reactive antigens between irAE and no irAE for the same 29 PreTx plasma samples screened against the same 24 antigens for (**A**) antigen panel 1 and (**B**) antigen panel 2. Statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant difference was observed.

3.4.2.2 Analysis of samples from the ICI-4W timepoint shows the same trends as the PreTx samples

Although no difference was observed between the PreTx number of reactive antigens and clinical outcome, it is possible that any differences in AAb profile are more evident after augmentation of immune activity by ICI treatment. Therefore, potential differences in number of reactive antigens at the timepoint after four week of ICI treatment (ICI-4W) were investigated. To this end, the binary serostatus data from both antigen panel 1 and antigen panel 2 were used to calculate the number of reactive antigens, out of a possible 24, for each of the 26 ICI-4W OAC samples. As with the data from the PreTx timepoint, associations of this data with clinical benefit status, ICI-induced tumour shrinkage and on-treatment irAE development were assessed. Once again, there was no significant difference in the number of reactive antigens at the ICI-4W timepoint between CB patients and NCB patients (**Figure 3.21**).

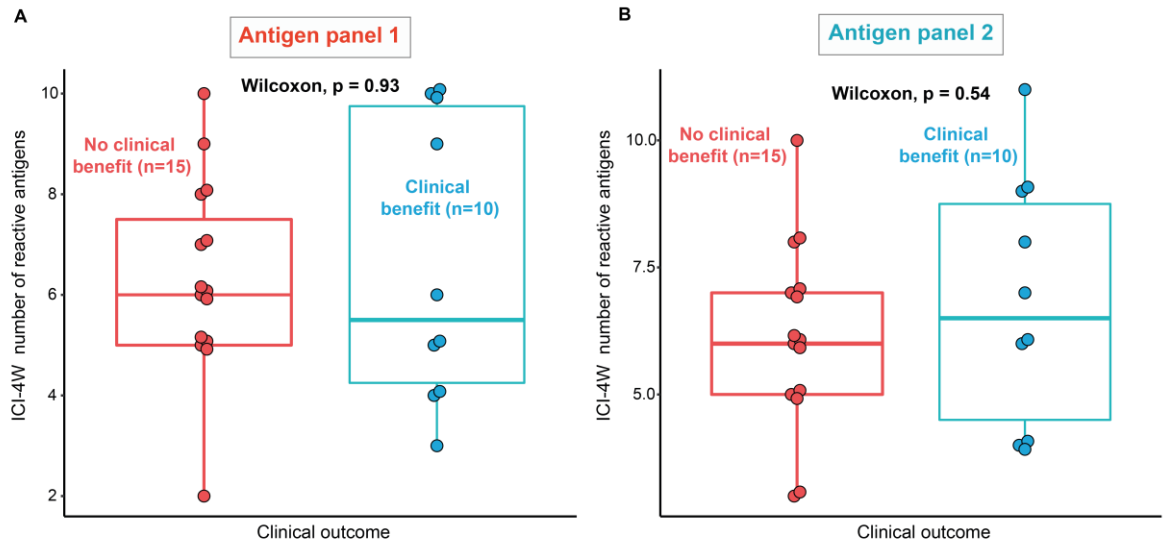


Figure 3.21. Number of reactive antigens at the ICI-4W timepoint and clinical benefit in the LUD2015-005 cohort. Boxplots showing the number of reactive antigens between clinical benefit and no clinical benefit for the same 25 ICI-4W plasma samples screened against the same 24 antigens for (**A**) antigen panel 1 and (**B**) antigen panel 2. Statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant difference was observed.

Correlation analysis between the number of reactive antigens at the ICI-4W timepoint, from both antigen panels, and change in tumour size over the four-week period of treatment with ICI only was performed and showed a non-significant negative association with change in tumour size following four weeks of ICI treatment (**Figure 3.22**).

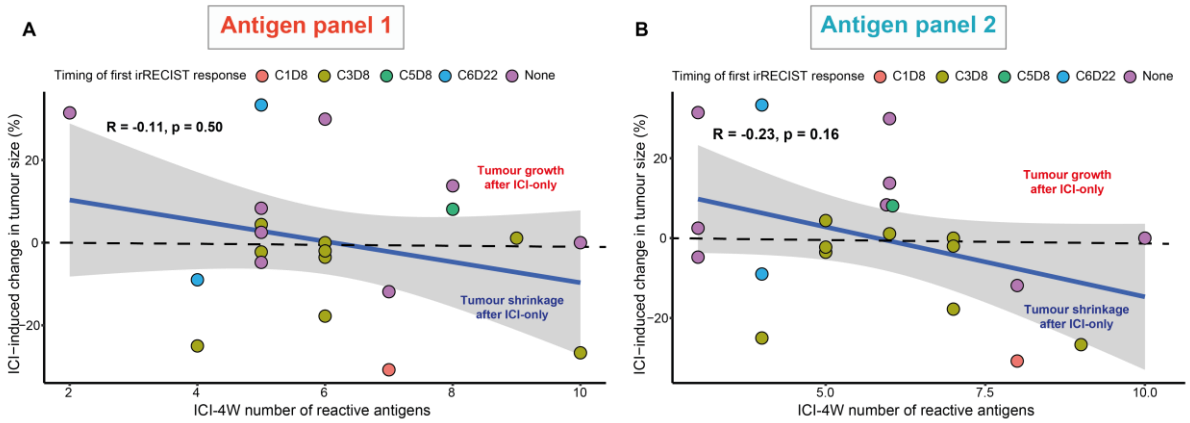


Figure 3.22. Number of reactive antigens at the ICI-4W timepoint and ICI-induced tumour shrinkage in the LUD2015-005 cohort. Correlation plots of ICI-induced change in tumour size and number of reactive antigens at the ICI-4W timepoint for each of the 21 OAC patients who had assessable tumour lesions before and after the initial four weeks of ICI treatment, and who had AAb profiling performed of their ICI-4W sample as part of the LUD2015-005 cohort. (**A**) Correlation plot using the number of reactive antigens from antigen panel 1 (**B**) Correlation plot using the number of reactive antigens from antigen panel 2. One point per patient. Points are coloured based on the time at which the corresponding patient experienced their first response according to the irRECIST criteria, with C indicating the cycle number of CTX and D indicating the day during that cycle at which tumour size was assessed. Dashed line indicates zero change in tumour size. Shaded area indicates the 95% confidence interval for the line fitted (blue). R and p values for each correlation analysis were calculated by fitting a linear model with the Kendall correlation method.

Finally, once again, the data from both antigen panels showed no significant difference in the number of reactive antigens at the ICI-4W timepoint between patients who developed an irAE during treatment and those who did not develop an irAE (**Figure 3.23**).

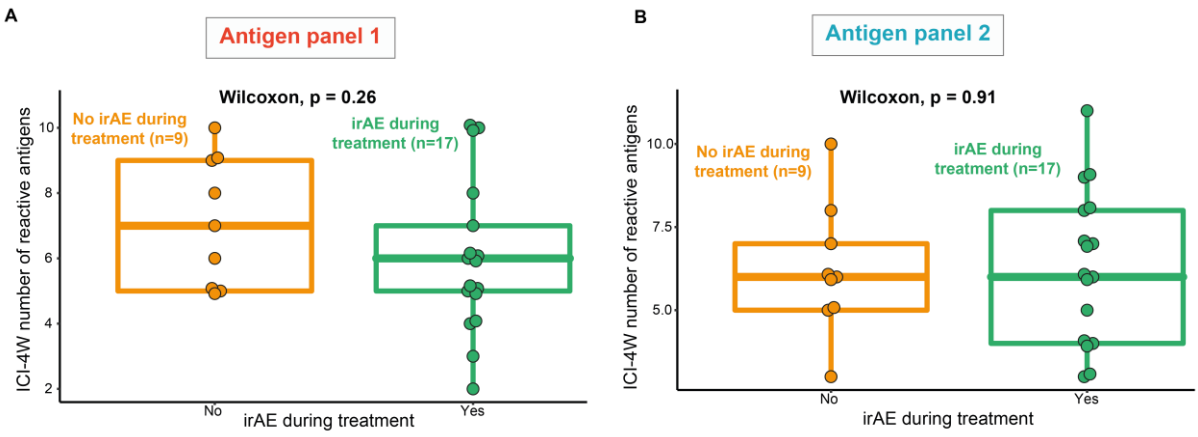


Figure 3.23. Number of reactive antigens at the ICI-4W timepoint and irAE development in the LUD2015-005 cohort. Boxplots showing the number of reactive antigens between irAE and no irAE for the same 26 ICI-4W plasma samples screened against the same 24 antigens for (**A**) antigen panel 1 and (**B**) antigen panel 2. Statistical comparisons of medians were carried out using a Wilcoxon test. No statistically significant difference was observed.

3.4.2.3 Autoantibody profiles remain relatively stable following four weeks of ICI treatment

Although no difference was seen when looking at the PreTx and ICI-4W timepoints in isolation, as was seen in studies of ICI treated melanoma[316–318], it is possible that ICI-induced change in antigen reactivity could be linked with clinical outcomes, potentially by providing an indication of successful circumvention of tumour associated immune tolerance.

To investigate this, the data from the LUD2015-005 cohort was subset to the 26 OAC patients with paired antigen reactivity data for their PreTx and ICI-4W samples. The change in number of reactive antigens, using data from both antigen panels, was then analysed with regards to CB status. Overall, the number of reactive antigens of a patient seemed to remain quite stable from PreTx to ICI-4W and no significant association was seen with CB status (**Figure 3.24**).

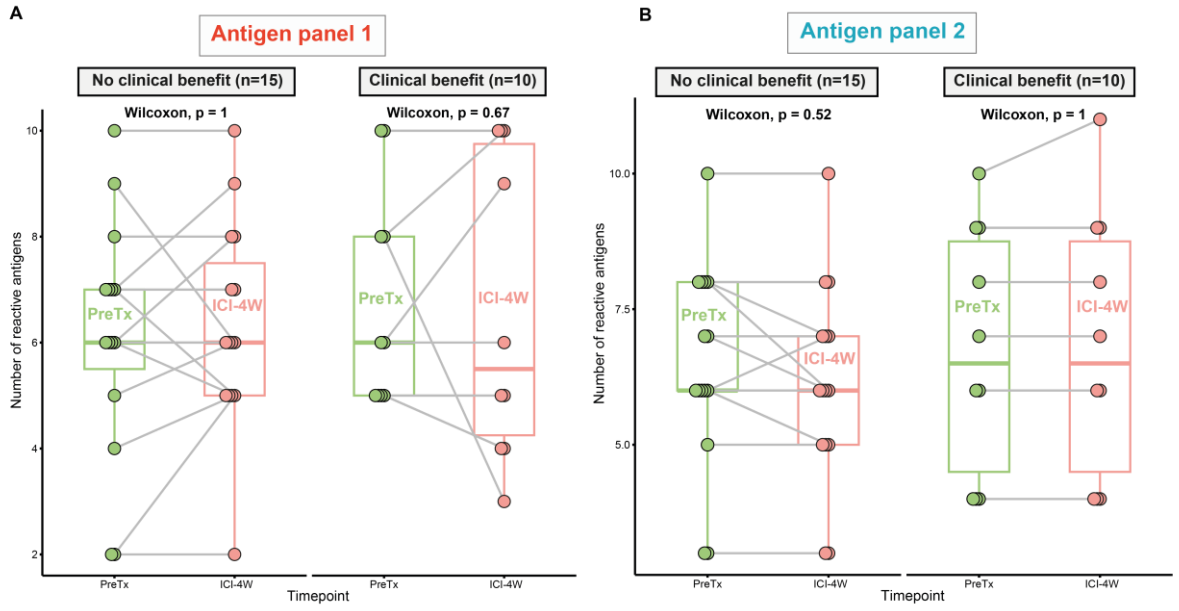


Figure 3.24. Change in the number of reactive antigens after four weeks of ICI treatment in the LUD2015-005 cohort. Boxplots showing the number of reactive antigens, out of a possible 24, for the 15 no clinical benefit patients and 10 clinical benefit who had both PreTx and ICI-4W plasma samples screened in the LUD2015-005 cohort. (**A**) antigen panel 1 and (**B**) antigen panel 2. One point per patient and points are coloured by timepoint. Grey lines connect PreTx and ICI-4W samples from the same patient. Statistical comparisons of medians were carried out using a Wilcoxon test. No statistically significant differences were observed.

3.4.3 Analysing reactivity to individual antigens as biomarkers

Although no association of number of reactive antigens with clinical outcome was observed, it is possible that the overall profile of antigen reactivity, or reactivity to a specific antigen could differ and predispose to CB.

3.4.3.1 Pre-treatment no antigens show differential reactivity between CB and NCB patients

The overall profile of antigen reactivity was assessed through PCA analysis of the binary serostatus data for the 24 antigens and 29 PreTx OAC plasma samples. However, this did not show clear clustering of samples by CB status (**Figure 3.25A-B**). Next, the association of reactivity to individual antigens with CB status was assessed. However, following correction for multiple testing, none of the 24 antigens were differentially reactive between CB and NCB patients (**Figure 3.25C-D**).

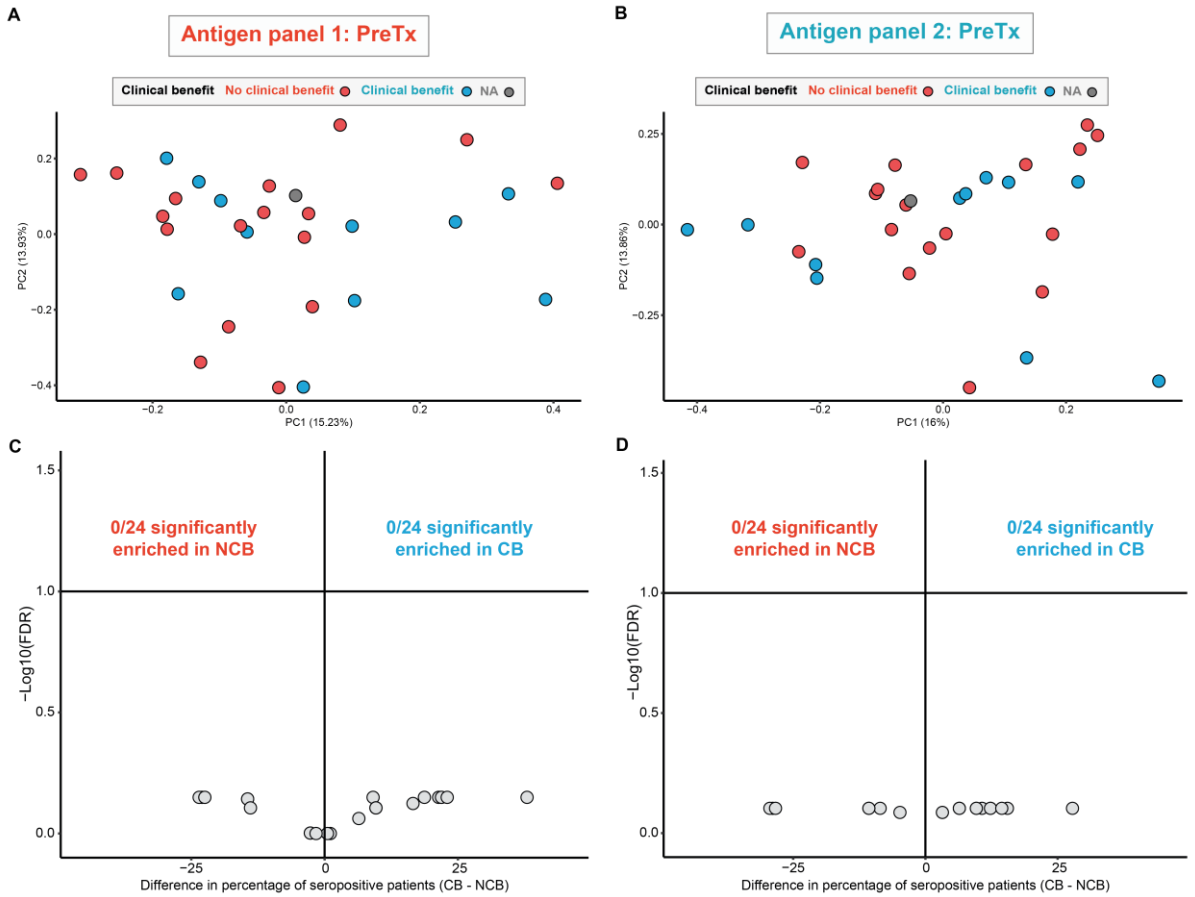


Figure 3.25. Reactivity to antigens PreTx in the LUD2015-005 cohort is not substantially different between clinical benefit and no clinical benefit patients. (**A-B**) PCA plots of binary serostatus data of the 24 antigens and 29 PreTx OAC plasma samples that were included in the LUD2015-005 cohort showing PC1 and PC2 for (**A**) antigen panel 1 and (**B**) antigen panel 2. One point per patient and points are coloured by clinical benefit status. (**C-D**) Volcano-like plots showing the difference in the percentage of seropositive patients between clinical benefit and no clinical benefit patients versus the -Log10(FDR) for association of reactivity to each of the 24 antigens with (**C**) antigen panel 1 and (**D**) antigen panel 2. p values were calculated using Wilcoxon comparison of group medians. False discovery rates (FDRs) were then calculated from these p values using the Benjamini–Hochberg method. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for FDR which was 0.1.

3.4.3.2 After four weeks of ICI treatment no antigens show differential reactivity between CB and NCB patients

As for the PreTx samples, the overall profile of antigen reactivity was assessed through PCA analysis of the binary serostatus data for the 24 antigens and 26 ICI-4W OAC plasma samples. However, once again, this did not show clear clustering of samples by CB status (**Figure 3.26A-B**). Next, the association of reactivity to individual antigens with CB status was assessed. However, as for the PreTx timepoint, following correction for multiple testing, none of the 24 antigens were differentially reactive between CB and NCB patients (**Figure 3.26C-D**).

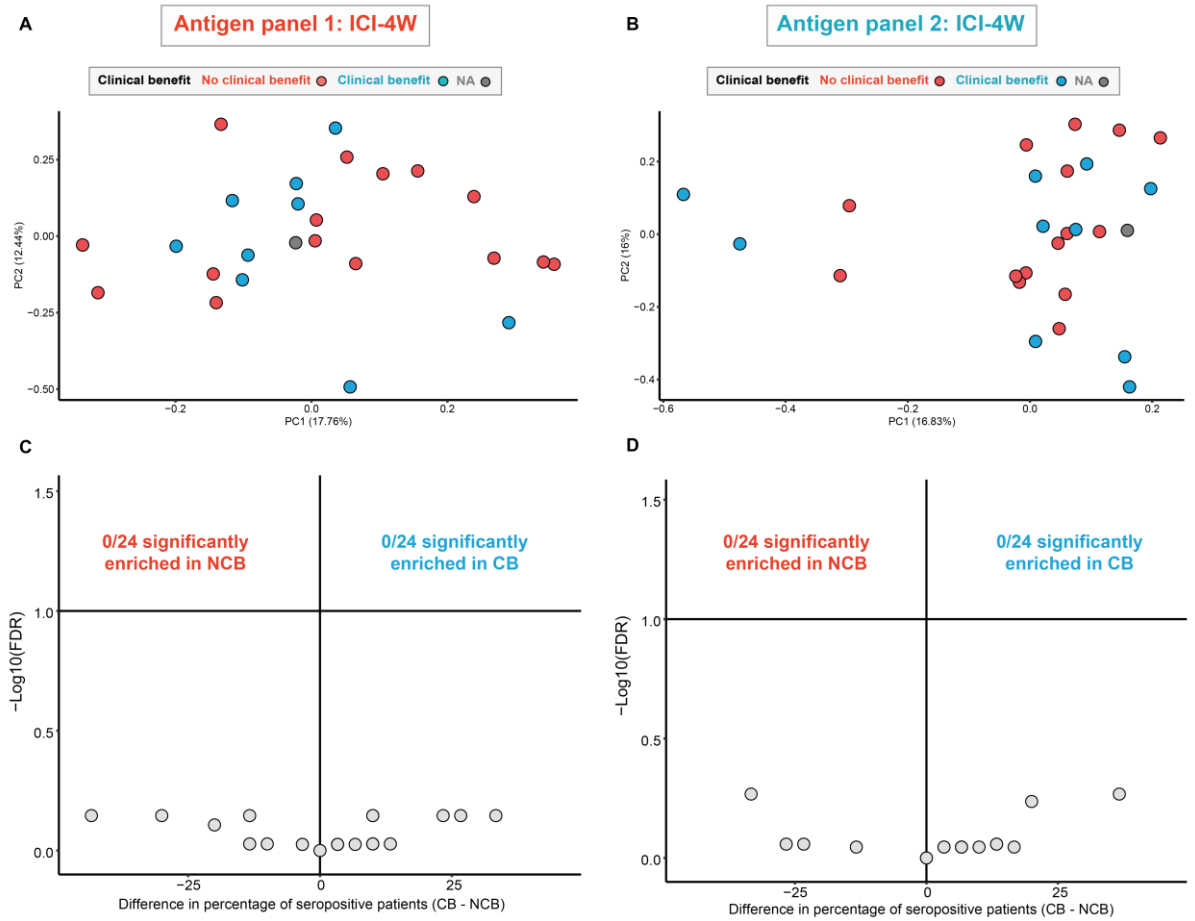


Figure 3.26. Reactivity to antigens at the ICI-4W timepoint in the LUD2015-005 cohort is not substantially different between clinical benefit and no clinical benefit patients. (**A-B**) PCA plots of binary serostatus data of the 24 antigens and 26 ICI-4W OAC plasma samples that were included in the LUD2015-005 cohort showing PC1 and PC2 for (**A**) antigen panel 1 and (**B**) antigen panel 2. One point per patient and points are coloured by clinical benefit status. (**C-D**) Volcano-like plots showing the difference in the percentage of seropositive patients between clinical benefit and no clinical benefit patients versus the -Log10(FDR) for association of reactivity to each of the 24 antigens for (**C**) antigen panel 1 and (**D**) antigen panel 2. p values were calculated using Wilcoxon comparison of group medians. False discovery rates (FDRs) were then calculated from these p values using the Benjamini–Hochberg method. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for FDR which was 0.1.

3.5. Exploratory cohort: autoantibodies in Lung cancer, Barrett's Oesophagus, and Oesophageal cancer

3.5.1 Rationale

Although not repeated and assessed for reproducibility, the serostatus data from the exploratory cohort provides a valuable opportunity to explore the potential for AAb profiling to predict clinical outcome to ICI in LC, and non-invasive distinction of BO and OC patients.

As additional plasma samples from ICI treated OAC patients were not obtainable for the validation cohort, plasma samples from 95 ICI treated LC patients were kindly shared with us by Professor Christian Ottensmeier (University of Southampton and University of Liverpool) and Dr Ioannis Karydis (University of Southampton) as part of a collaboration with the University of Southampton. Unlike the LUD2015-005 patients, these patients did not receive CTX at any point during treatment and so provided an opportunity to further explore the association of AAbs with clinical outcomes to ICI, without concurrent CTX.

This exploratory cohort also comprised plasma samples from 102 BO and 129 OC which allowed for evaluation of reactivity to 373 antigens as potential biomarkers to non-invasively distinguish BO from OC patients and potentially facilitate early cancer detection.

3.5.2 Biomarkers of response to immunotherapy in lung cancer

3.5.2.1 Clinical summary

Patient demographics for the 95 LC patients from which the plasma samples were taken are summarised in Table 3.4.

Cohort	*Full cohort (n=95)*	*First line ICI adenocarcinoma only (n=42)*
Age in years at study enrolment		
Median (range)	68 (47 - 86)	67 (47-85)
***Sex**: Number (%)*		
Male	63 (66%)	26 (62%)
Female	32 (34%)	16 (38%)
***Histology:** Number (%)*		
Adenocarcinoma	66 (70%)	42 (100%)
Squamous	26 (27%)	0 (0%)
Not specified	3 (3%)	0 (0%)
***Line of treatment ICI given as**: Number (%)*		
1st	57 (60%)	42 (100%)
2nd	26 (27%)	0 (0%)
3rd-5th	12 (13%)	0 (0%)
Best irRECIST response: Number (%)		
Complete/Partial Response (CR/PR)	49 (52%)	31 (74%)
Stable Disease (SD)	16 (17%)	4 (9%)
Progressive disease (PD)	26 (27%)	7 (17%)
Not assessable	4 (4%)	0 (0%)

Table 3.4. Summary of patient demographics and clinical benefit for immunotherapy treated lung cancer samples collected from University of Southampton. Summary metrics are shown for the full cohort of lung cancer patients, and for the subset of 42 lung adenocarcinoma (LAC) patients who received ICI as their first line of treatment. Best irRECIST response refers to the best tumour response, in terms of decrease in tumour size, during treatment according to the irRECIST criteria.

Unlike the LUD2015-005 cohort, in this cohort, there was variability with regards to the line of treatment at which ICI was given to the patient. To match these samples as closely as possibly to the LUD2015-005 cohort, this full cohort of 95 LC patients was subset to only the 42 patients who received ICI as first line treatment and who had cancer of adenocarcinoma histology. These 42 lung adenocarcinoma (LAC) patients showed a greater median observed overall survival (702 days) compared to the entire cohort of LC patients (592 days) (**Figure 3.27**).

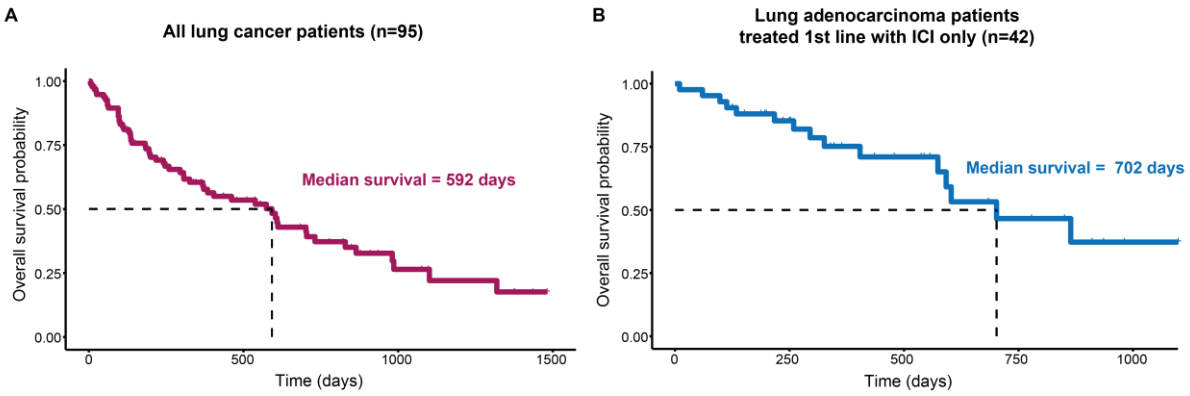


Figure 3.27 Overall survival outcomes in the ICI treated lung cancer samples from Southampton. Kaplan Meier plots of overall survival for **(A)** all 95 lung cancer patients (maroon) or **(B)** the 42 lung adenocarcinoma patients who received ICI as the first line of treatment (blue). The median observed overall survival for each cohort is marked on each plot with dashed lines.

3.5.2.2 irRECIST responders show an increased number of reactive antigens PreTx

As for the LUD2015-005 patients, the number of reactive antigens, out of a possible 172, were calculated for each of the 42 first line ICI treated LAC samples. The corresponding data was then used to assess whether there was an association with best irRECIST response and OS, which were two main measures of clinical outcome available for these patients. This analysis showed that irRECIST responders, defined as patients who showed clinically relevant shrinkage of their tumour as assessed by irRECIST criteria at any point during treatment, had a significantly greater number of PreTx reactive antigens on average than irRECIST non-responders (**Figure 3.28A**). Moreover, patients who had a high number of reactive antigens (>median) showed improved OS compared to patients who had a low number of reactive antigens (≤median), although this was not a significant difference (**Figure 3.28B**).

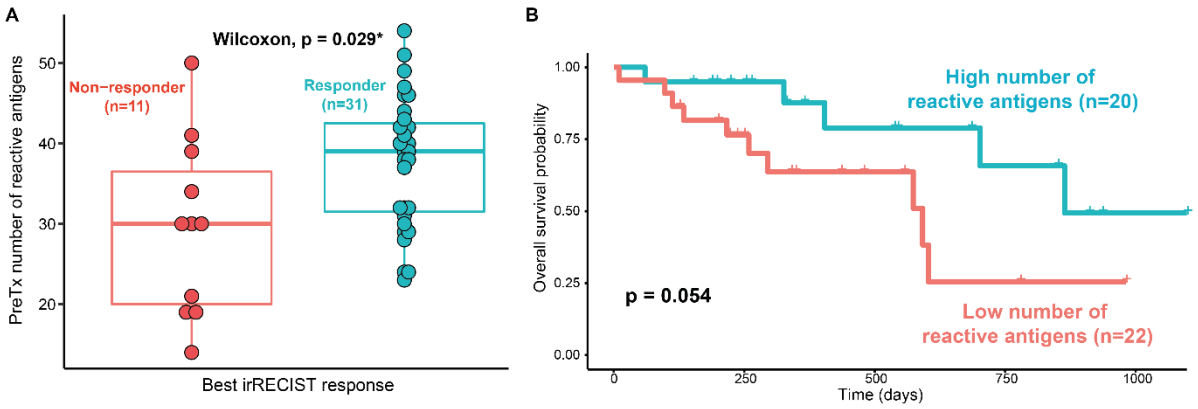


Figure 3.28 Number of reactive antigens and clinical outcome. (**A**) Boxplots showing the number of reactive antigens, out of a possible 172, between irRECIST responders (blue) and irRECIST non-responders (red) for the 42 PreTx plasma samples from lung adenocarcinoma patients who were treated 1st line with ICI. Statistical comparisons of medians were carried out using a Wilcoxon test. (**B**) Kaplan Meier plot of overall survival with patients split into those with a high number of reactive antigens (> median) (blue) and those who had a low number of reactive antigens (≤ median) (red). p value was calculated using the log rank method.

3.5.2.3 Pre-treatment no antigens show differential reactivity between irRECIST responders and non-responders

Next, I investigated the possibility that specific antigen reactivity profiles were associated with the likelihood of a beneficial irRECIST response. The overall profile of antigen reactivity was assessed through PCA analysis and unsupervised hierarchical clustering of the binary serostatus data for the 172 antigens for each of the 42 samples. No obvious clustering of samples by irRECIST response was seen using either approach (**Figure 3.29A-B**). The association of reactivity to individual antigens with patient irRECIST response and OS was also assessed.

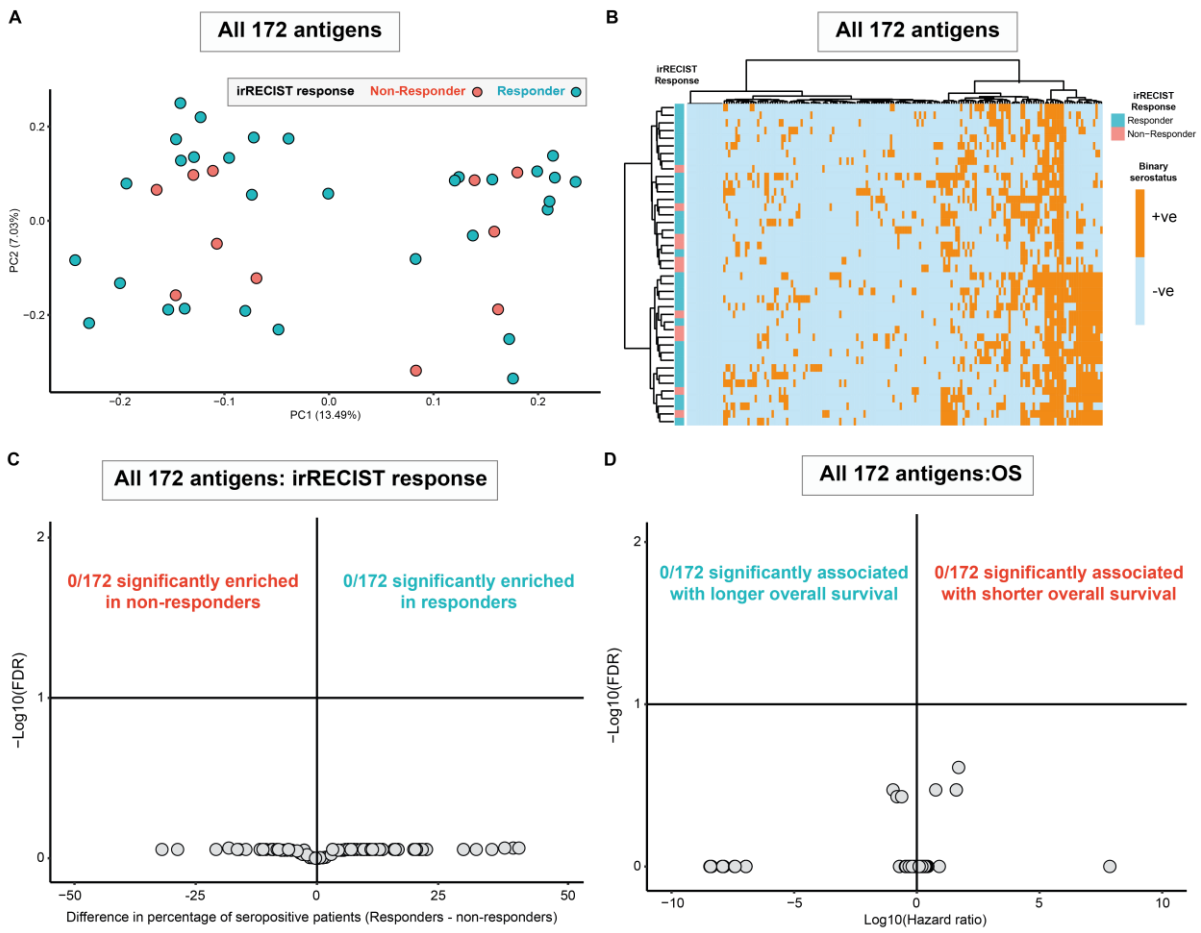


Figure 3.29. Reactive antigen profiles of PreTx samples from ICI treated lung adenocarcinoma patients are not substantially different between irRECIST responders and irRECIST non-responders. (**A**) PCA plot of binary serostatus data of the 172 antigens from the 42 PreTx lung adenocarcinoma plasma samples that were included in the exploratory cohort showing PC1 and PC2. One point per patient and points are coloured by irRECIST response status. (**B**) Heatmap of the same binary serostatus data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from irRECIST non-responders (red) or irRECIST responders (blue). (**C**) Volcano-like plot showing the difference in the percentage of seropositive patients between irRECIST responding patients (PR/CR) and irRECIST non-responding patients (SD/PD) versus the -Log10(FDR) for association of reactivity to each of the 172 antigens with irRECIST response. (**D**) Volcano-like plot showing the Log10(Hazard ratio) versus the -Log10(FDR) for association of reactivity to each of the 172 antigens with OS. p values for these volcano-like plots were calculated using Wilcoxon comparison of group medians and univariate cox regression respectively. False discovery rates (FDRs) were then calculated from these p values using the Benjamini–Hochberg method. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for FDR which was 0.1.

Although several antigens showed significant association by univariate analysis, following correction for multiple testing, none of the 172 antigens were either differentially reactive between irRECIST responders and non-responders or significantly associated with longer or shorter OS (**Figure 3.29C-D**).

3.5.3 Anti-p53 autoantibodies are a promising biomarker to non-invasively distinguish Oesophageal cancer patients from Barrett's oesophagus patients

As previously discussed, there is much scope to increase the number of OC cases which are detected at an early stage. In particular, a high throughput and non-invasive clinical test to identify those patients who are most likely to have, or develop OC, would allow for these patients to be stratified for monitoring by endoscopy, therefore increasing the likelihood detection of early cancer detection. The exploratory cohort, which included samples from 102 BO patients and 129 OC patients, provided a valuable opportunity to discover AAb based biomarkers which could help fulfil this need. To this end, the serostatus data for the 102 BO patients and 129 OC patients was analysed for differentially reactive antigens. As for the previous samples, the number of reactive antigens, out of a possible 172, was calculated for each of the 102 BO and 129 OC patient samples. Analysis of this data showed no significant difference between the number of reactive antigens between BO and OC patients (**Figure 3.30A**).

Differential reactivity to antigens between BO and OC patients was then assessed. Although several antigens showed significant association by univariate analysis, following correction for multiple testing, none of the 172 antigens were significantly differentially reactive, based on a false discovery rate (FDR) of 0.1 or lower, between BO and OC patients, although two antigens, P3H4 and TP53_NTerm had an FDR of 0.137 (**Figure 3.30B-D**).

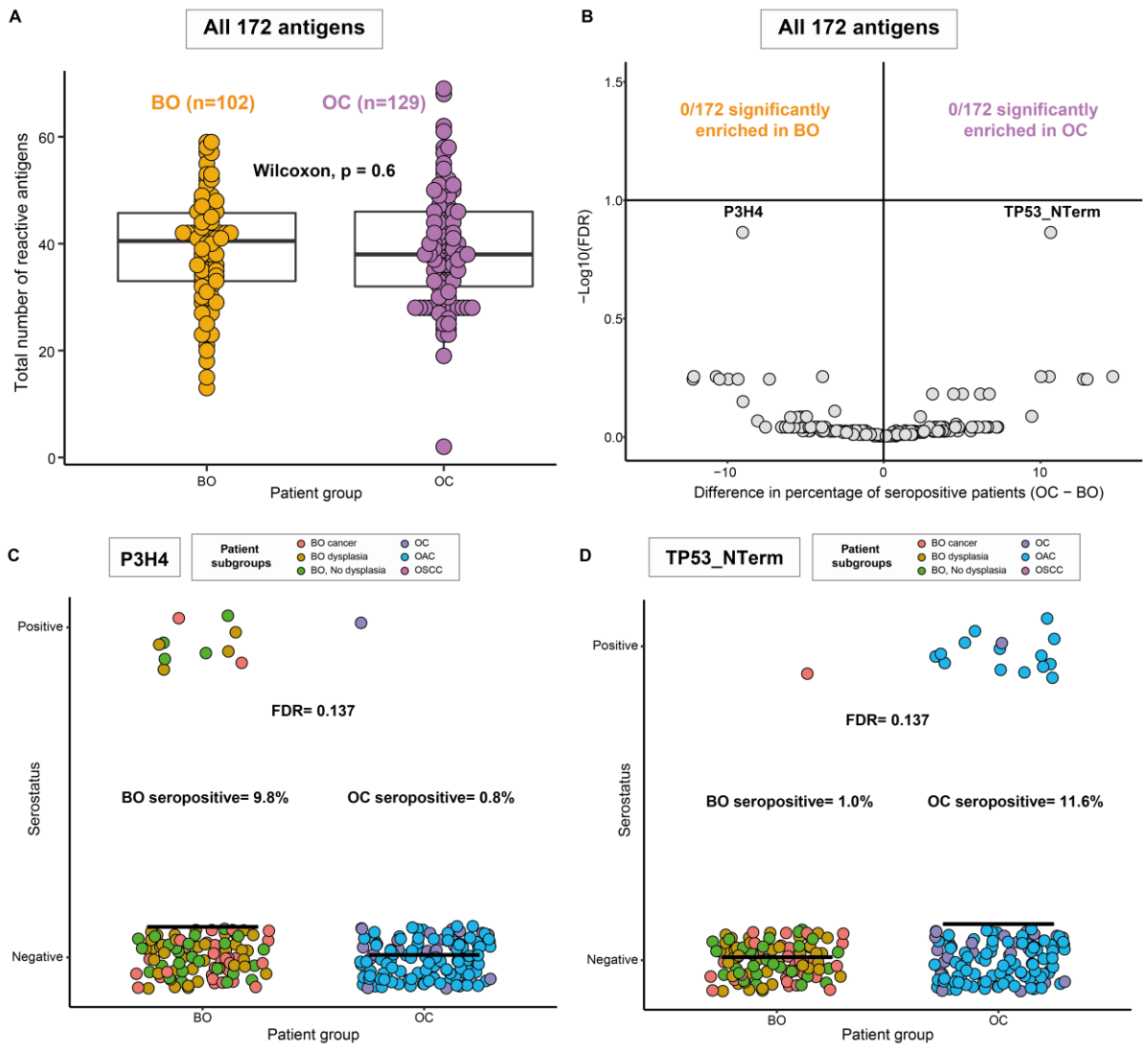


Figure 3.30. Reactive antigen profiles of BO and OC patients identifies reactivity to TP53_NTerm as a potential distinguishing biomarker. (**A**) Boxplots showing the number of reactive antigens, out of a possible 172, between 102 BO patients (orange) and 129 OC patients (purple). Statistical comparisons of medians were carried out using a Wilcoxon test. (**B**) Volcano-like plot showing the difference in the percentage of seropositive patients between OC patients and BO patients versus the -Log10(FDR) for association of reactivity to each of the 172 antigens with patient group. One point per antigen. (**C-D**) Boxplots showing the serostatus for (**C**) P3H4 and (**D**) TP53_NTerm, which were the two antigens with the lowest FDR values, in the 102 BO patients compared to the 129 OC patients. One point per patient and points are coloured by patient subgroup. BO cancer = BO with a history of cancer, BO dysplasia = BO with dysplasia, BO No dysplasia = BO with no dysplasia, OC = OC with unconfirmed subtype. p values were calculated using Wilcoxon comparison of group medians. False discovery rates (FDRs) were then calculated from these p values using the Benjamini–Hochberg method. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for FDR which was 0.1.

However, closer inspection of the patient subgroups revealed that the only TP53_Nterm seropositive BO patient had a history of non-oesophageal cancer (**Figure 3.30D**). Even with this patient, reactivity to TP53_Nterm showed an impressive specificity of 93.8% for OC, but relatively poor sensitivity of 11.6% for OC.

3.6. Autoimmune associated autoantibodies and clinical outcome in the LUD2015-005 cohort

3.6.1 Rationale

The previously mentioned studies of HIV infected individuals and ICI treated melanoma which reported associations of AAbs with development bnAbs or irAEs, focused on assessing reactivity to autoantigens associated with autoimmune disease[316,318], rather than the more generic DNA- and RNA-binding proteins which were the focus of the LUD2015-005 cohort. It is therefore not unreasonable to speculate that measurement of AAb based reactivity to autoantigens which are well established to be implicated in autoimmune diseases is a more pertinent assessment of baseline immune activation. To address this possibility, PreTx plasma samples from the LUD2015-005 were screened for reactivity to autoimmune disease associated autoantigens and this data was assessed for association with patient clinical outcomes.

3.6.2 Data overview and workflow

77 plasma samples, including 31 of the PreTx samples from the LUD2015-005 cancer patients, were screened for 33 different autoimmune associated AAbs using the commercially available slide-based Human Autoimmune Disease IgG Autoantibody Array G1 kits (RayBiotech) (henceforth referred to as RayBiotech AAb array). The other 46 blood plasma samples screened were 6 from healthy volunteers, 25 from BO patients, and 13 from ICI-4W timepoint of the LUD2015-005 patients. Data extraction and processing was performed based on the manufacturers recommendations and are briefly summarised below (**Figure 3.31**). Full details regarding these arrays, including processing of images and data, are described in section 2.3.2.

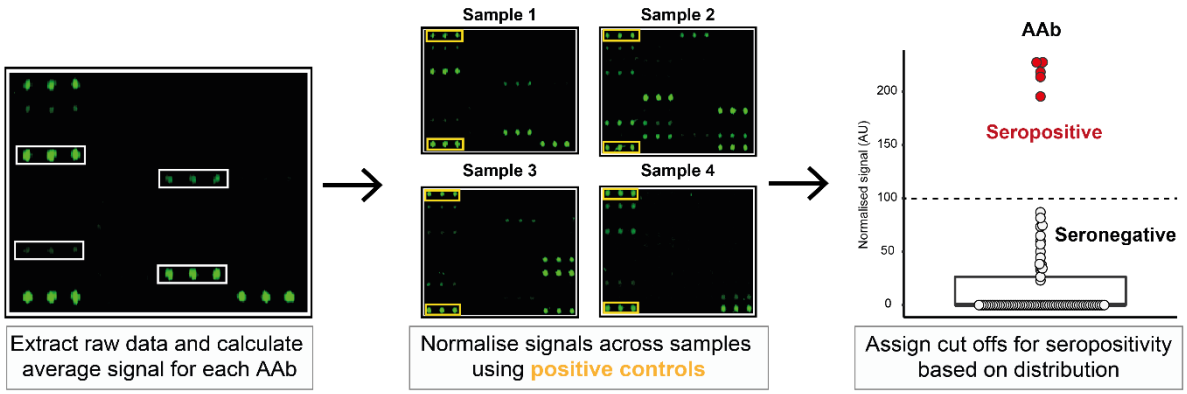


Figure 3.31. Data extraction and processing of RayBiotech AAb array. Fluorescent signal from array images were extracted for each AAb which were detected in triplicate (white boxes) and the average signal for each AAb within each sample was calculated. Signals from the positive controls (yellow boxes) were then used to normalise signals across all 77 samples. A cut off (dashed line) for each AAb was then assigned based on the distribution of normalised signal from all 77 samples.

Cut-offs and seropositivity rates for 32 of the 33 AAbs were assigned and calculated based on the distribution of signal from the 77 samples (**Figure 3.32**). No cut-off was assigned for gliadin due to due to the high signal in almost all samples and subsequent lack of dichotomy in the signal distribution (top right panel of **Figure 3.32**).

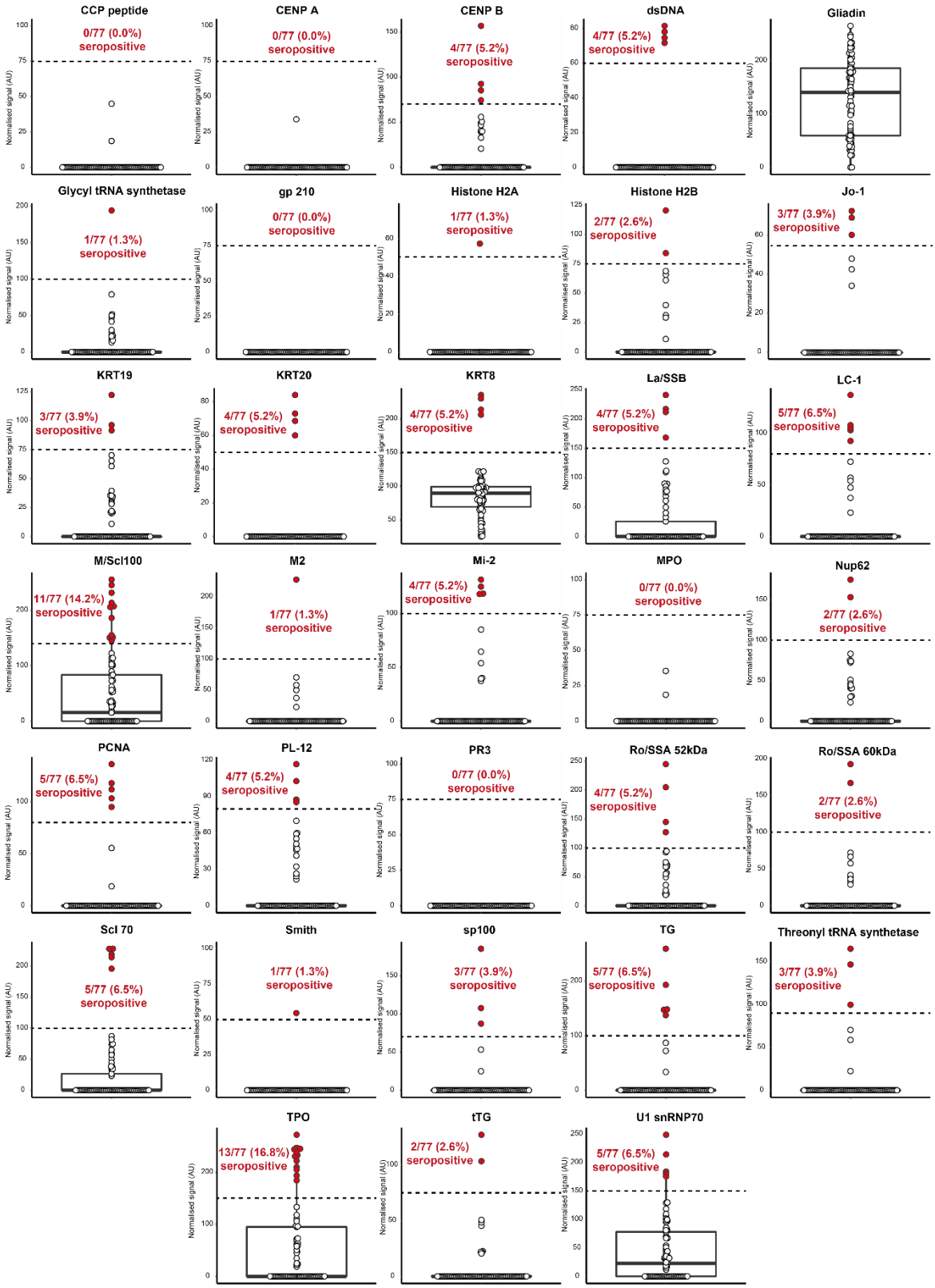


Figure 3.32. Cut-offs for serostatus for normalised data from RayBiotech AAb array. Boxplots of normalised signals from all 77 samples for all 33 AAbs detected by the RayBiotech AAb array. For 32 out of 33 AAbs cut offs (dashed lines) and seropositivity rates (in red) were assigned and calculated based on the distribution of signal from the 77 samples. No cut off was assigned, and hence no serostatus was calculated, for gliadin (top right box plot) due to the high signal in almost all samples and subsequent lack of dichotomy in the signal distribution.

3.6.3 No association of number of autoantibodies PreTx with clinical benefit or irAEs

The binary serostatus for the remaining 32 AAbs in the 31 PreTx OAC samples from the LUD2015-005 patients were used to assess whether there was an association with clinical outcome. Firstly, the number AAbs, out of a possible 32, was calculated for each of the 31 PreTx OAC samples. Analysis of this data showed that there was no significant difference in the number of AAbs based on clinical benefit status (**Figure 3.33A**). Given that many patients were seronegative for all 32 AAbs, analysis of CB status based on seropositivity to at least one of the 32 autoimmune associated AAbs was also performed. However, this analysis also showed no significant association with CB status (**Figure 3.33B**).

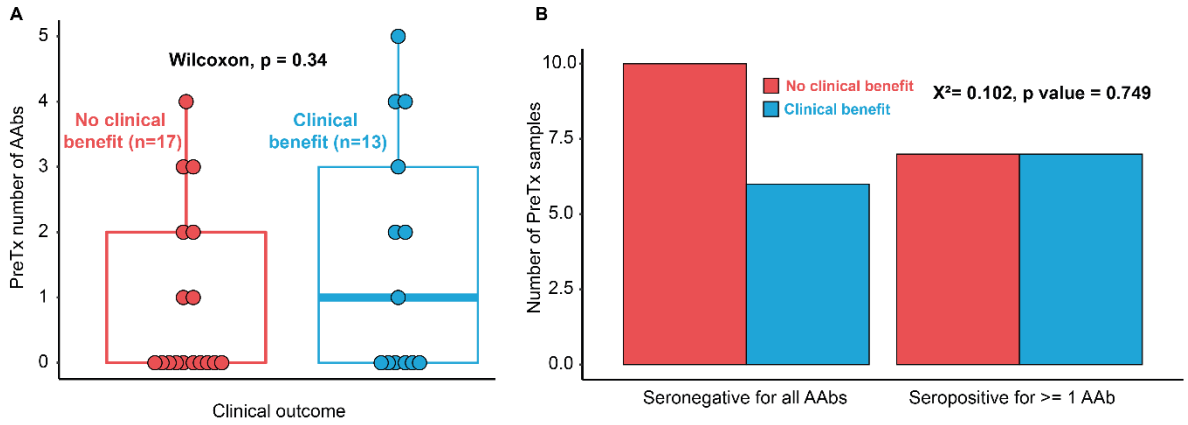


Figure 3.33. No association between number of autoimmune associated AAbs PreTx and clinical benefit. (**A**) Boxplots showing the number of AAbs, out of a possible 32, by clinical benefit status for the 30 PreTx OAC plasma samples. Statistical comparison of the medians of each group was carried out using a Wilcoxon test. (**B**) Bar plots showing the comparing the frequency of individuals in the clinical benefit and no clinical benefit groups which were seropositive for at least one AAb or seronegative for all AAbs. Statistical comparison of expected observations for each group was carried out using a Chi-squared test.

Similar analysis of both number of AAbs and seropositivity to at least one AAb also showed no significant association with development of an irAE during treatment (**Figure 3.34**).

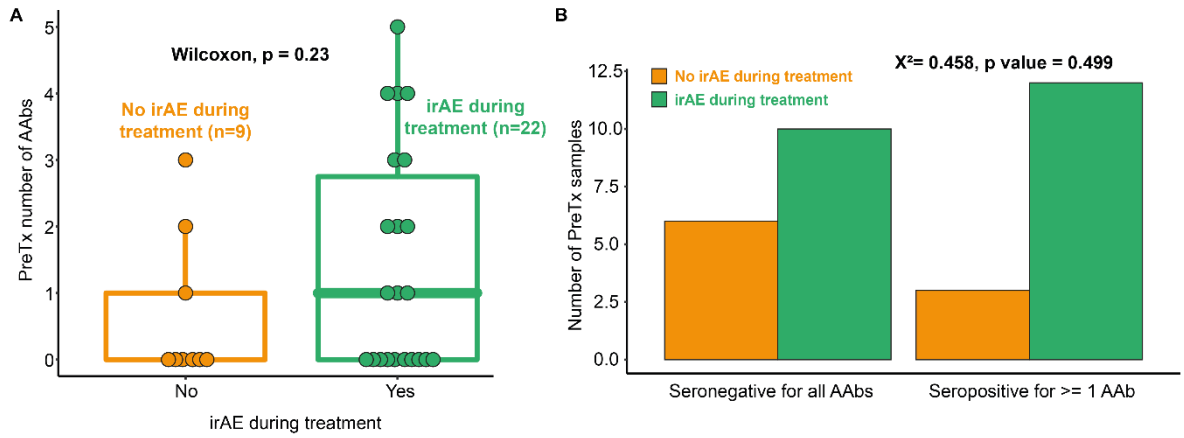


Figure 3.34. No association between number of autoimmune associated AAbs PreTx and irAE development. (**A**) Boxplots showing the number of AAbs, out of a possible 32, between irAE and no irAE for the 31 PreTx OAC plasma samples. Statistical comparison of the medians of each group was carried out using a Wilcoxon test. (**B**) Bar plots showing the comparing the frequency of individuals in the irAE and no irAE groups which were seropositive for at least one AAb or seronegative for all AAbs. Statistical comparison of expected observations for each group was carried out using a Chi-squared test.

The overall profile of AAbs for these samples was also assessed through PCA analysis and unsupervised hierarchical clustering of the binary serostatus data but no obvious clustering of samples by CB status or irAE development was seen using either approach (**Figure 3.35**).

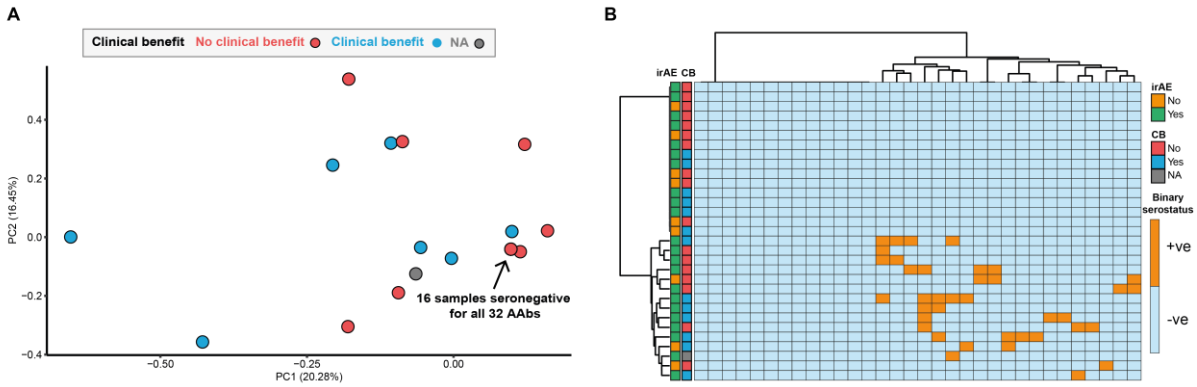


Figure 3.35 Autoimmune associated AAb profiles of PreTx samples from ICI treated OAC are not substantially different between clinical benefit and no clinical benefit patients. (**A**) PCA plot of binary serostatus data of the 32 AAbs from the 30 PreTx OAC plasma samples that were included analysed using the RayBiotech AAb array, showing PC1 and PC2. Arrow indicates point at which 16 samples are all superimposed due to identical profile of being seronegative for all 32 AAbs. Otherwise, one point per patient and points coloured by clinical benefit status. (**B**) Heatmap of the same binary serostatus data with unsupervised hierarchical clustering of both antigens (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating clinical benefit (red and blue) and irAE (yellow and green) status.

3.7 Exploring the aetiology of autoantigen reactivity

3.7.1 Rationale

Although there are suggested features which increase immunogenicity of autoantigens, it is still not clear why certain autoantigens are more commonly reactive in cancer patients[301,364]. The LUD2015-005 cohort provides a valuable opportunity to further understanding of this due to the in-depth multi-omics data that has been generated for these patients (as described in section 1.4 of chapter 1).

3.7.2 Mutations

As mentioned in section 3.2.1, mutations have been reported to change the structural conformation of proteins which can increase immunogenicity of the protein by generating or revealing neoepitopes[364]. In some cases, such as for p53, a mutation in the corresponding gene can also stabilise the protein and increase its half-life, which could potentially increase its immunogenicity[364,382]. To determine whether mutated *TP53* increases the likelihood of p53-AAb development, the *TP53* mutation status was determined from WGS data of DNA extracted from tumour biopsies that were sampled prior to treatment from the LUD2015-005 inoperable cohort. DNA extraction from these biopsies was performed by Phil Xie and Jaeho Chan, and analysis of the subsequent WGS data to determine *TP53* mutation status was performed by Professor Benjamin Schuster-Böckler, Dr Iliana Peneva, and Dr Joseph Kaplinsky.

Based on the binary serostatus data for reactivity to p53, there was no association of tumour *TP53* mutation status and increased prevalence, or number, of anti-p53 AAbs (**Figure 3.36**).

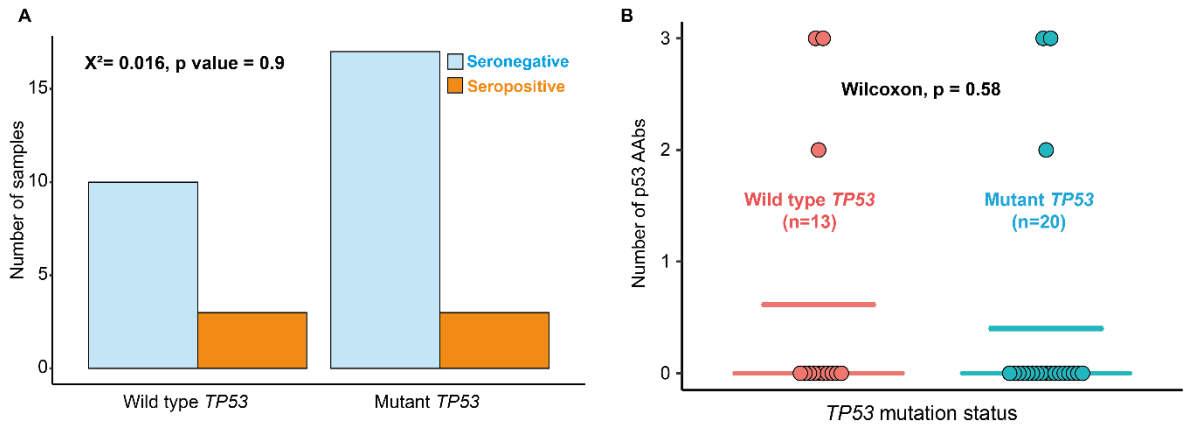


Figure 3.36. No association between tumour *TP53* mutation status and serostatus of anti-p53 AAbs. (**A**) Bar plots showing the comparing the frequency of individuals in the irAE and no irAE groups which were seropositive for at least one AAb or seronegative for all AAbs. Statistical comparison of expected observations for each group was carried out using a Chi-squared test. (**B**) Dot plots showing the number of p53-AAbs, out of a possible 3, between patients with wild type tumour *TP53* and mutant tumour *TP53* for the 33 PreTx OAC plasma samples analysed in antigen panel 2. Statistical comparison of the medians of each group was carried out using a Wilcoxon test.

3.8.2 Cross-reactivity with infectious agents

Molecular mimicry is a phenomenon in which foreign antigens, usually those expressed by pathogens, share identical or similar epitopes or structures with self-antigens[383]. This means that lymphocytes which react with foreign antigens via their BCR/TCR can also react with self-antigens, which, in some cases, is believed to elicit autoimmunity[383]. The most recent demonstration of this phenomenon was the association between the development of the autoimmune disease multiple sclerosis (MS) and seropositivity to EBV derived antigens[384,385]. In this case, B cells which expressed BCRs with reactivity to EBNA1 were found to also be reactive to the self-antigen glial cell adhesion molecule (GlialCAM)[385].

A recent meta-analysis of AAb data showed that a group of 21 commonly reactive autoantigens each share an amino acid sequence with proteins expressed by infectious agents[386]. One of these autoantigens was the krüppel C2H2-type zinc finger protein family member ZNF688 which, in the data discussed in this chapter, showed high percentages of seroprevalence in samples tested in both antigen panel 1 and antigen panel 2. ZNF688 was shown to share an amino acid sequence, GAQPPAP at amino acids 164-170, with the EBV enzyme thymidine kinase (KITH-EBV)[386]. However, despite high

seropositivity rates, the antigen panels described in this chapter used a protein fragment of ZNF688 which encompassed amino acids 95-166 and so only included the first three amino acids of the shared seven amino acid sequence (**Figure 3.37**).

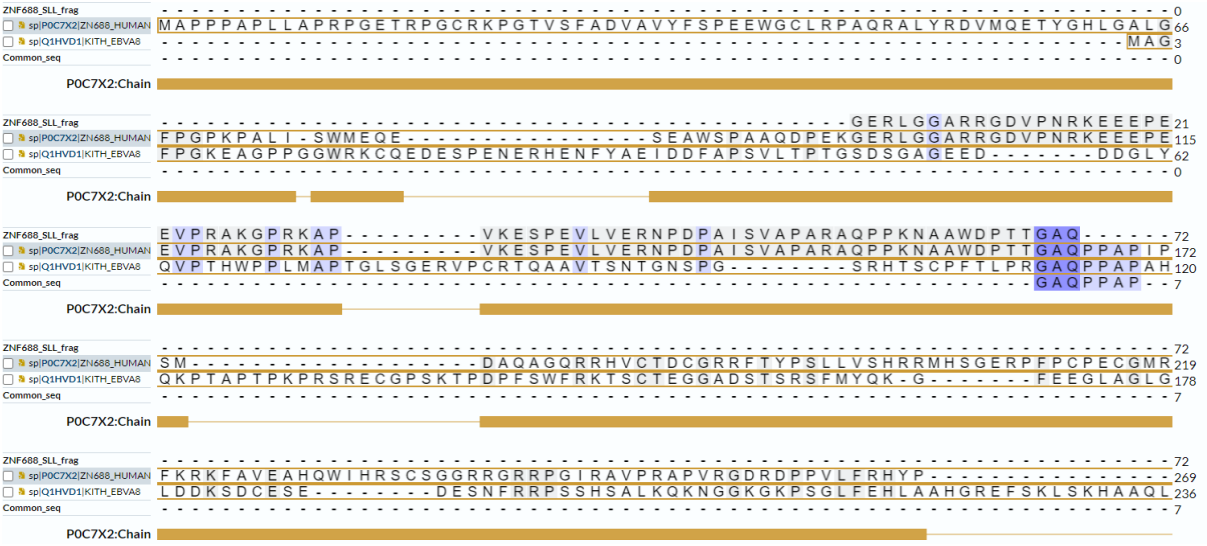


Figure 3.37. UniProt alignment tool shows a shared amino acid sequence between the common autoantigen ZNF688 and the EBV enzyme KITH. UniProt alignment result comparing the amino acid sequences of the ZNF688 protein fragment, corresponding to acids 95-166 and which was included in both antigen panel 1 and two, the full length ZNF688 protein, the EBV enzyme thymidine kinase (KITH_EBV), and the reported common sequence of seven amino acids GAQPPAP. Exact matches across all four sequences are shown in purple, whilst matches across three of four sequences are shown in lilac.

3.9 Discussion

The first aim of this thesis was to establish whether peripheral profiling of AAb based antigen reactivity, as a surrogate for baseline immune activation, could predict clinical outcome to ICI+CTX treatment in OC patients. Whilst profiles of antigen reactivity did not show significant associations with clinical outcome of ICI+CTX treated OC patients, a significant association was observed from similar analysis of LAC patients treated with 1st line ICI, suggesting that antigen reactivity may be a more appropriate predictor of good clinical outcome in patients treated with ICI in the absence of concurrent CTX.

The clinical outcomes from the inoperable arm of the LUD2015-005 clinical study were similar to those reported by the recent phase III clinical trials[30–32,36,81,320], and as expected there was considerable heterogeneity in clinical outcome between patients. As reported in other studies of ICI treated cancer, development of an irAE during treatment was significantly associated with improved clinical outcome.

3.9.1 Development of a novel analysis workflow to evaluating AAb based antigen reactivity to an array of autoantigens

Given the reported links between AAbs and irAE development, two Luminex bead based custom antigen panels were designed to profile AAb based antigen reactivity of plasma samples from ICI+CTX treated OC patients from the LUD2015-005 clinical study. However, preliminary analysis of the data generated showed poor reproducibility when using suggested analysis approach suggested by the manufacturer. Hence, a Gaussian FMM workflow was established to identify antigens with reactivity data that showed a bimodal-like distribution and then subsequently calculate reproducible binary serostatus data for these antigens. This is the first instance, to my knowledge, that a FMM workflow has been applied to process reactivity data from a multiplexed panel of multiple autoantigens. This workflow also differs from other reported FMM approaches by incorporating a step to identify antigens with a bimodal-like distribution before calculating binary serostatus data[376–379].

Net MFI data from the two Luminex based arrays (antigen panel 1 and antigen panel 2) showed a clear batch effect for the samples and antigens of the LUD2015-005 cohort. The poor reproducibility of net MFI data for some samples could be due to many factors. For example, degradation of antigen or poor coupling to the Luminex beads could contribute to poor reproducibility by reducing the capacity for AAb binding. Reassuringly, the binary serostatus data generated by applying the FMM workflow to this net MFI data did not show an obvious batch effect for the same samples. However, there were still some antigens and samples which showed differences in reactivity data between the two replicates, demonstrating that some variability between batches remained.

Despite generating serostatus data with considerably improved reproducibility compared to the net MFI data, there are limitations of the FMM workflow described. The most notable of which is that it may overestimate seropositivity for some antigens due to the combination of the skew of net MFI data, which is often seen for autoantigens, and the approach of calculating seropositivity based on a probability of 0.5 or greater of belonging to the upper distribution[376,379]. This could potentially be improved through the use of FMM which uses a skewed normal distribution, rather than a Gaussian distribution, and/or increasing the probability for assignment as seropositive from 0.5, perhaps on a per antigen basis[376,379].

This binary serostatus data provided a means, with demonstrated reproducibility, for assessing links between AAb profiling and clinical outcomes of the patient samples profiled.

3.9.2 AAb profiling and clinical outcome to ICI treatment

The PreTx number of reactive antigens showed no association with clinical outcomes of the ICI+CTX treated OAC patients from the inoperable arm of the LUD2015-005 study. This finding was supported by the use of the RayBiotech AAb array as an additional method for profiling AAb based reactivity to a different panel of autoantigens which are associated with autoimmune disease. However, there was a significant association of PreTx number of reactive antigens with likelihood of positive irRECIST responses in LAC patients treated first line with ICI. This difference in observation between OAC and LAC patients could possibly be due to differences in cancer type or, more likely, the treatment received. For example, it is not unreasonable to hypothesise that the PreTx number of antigen reactivities may be more relevant to tumour shrinkage in patients treated with ICI without concurrent CTX. This could also explain the non-significant positive association of PreTx number of reactive antigens and ICI-induced tumour shrinkage in the patients of the LUD2015-005 cohort.

Another possibility for a lack of association of number of reactive antigens with clinical outcome in the LUD2015-005 cohort could be due to differences in the platform of antigens used. For example, the study of SARS-CoV-2 infected individuals which showed that the number of reactivities was associated with disease severity may be biased towards measuring the number of strong AAb responses due to the nature of the REAP antigen platform used which, like phage display, utilises pull down of human AAbs. Another key difference between the assays is their focus on more than 2,000 antigens which are part of the exoproteome rather than the approach described in this chapter which focused primarily on 380 DNA- and RNA- binding proteins which are almost exclusively of an intracellular location[270,355].

The limited statistical power provided by the relatively small cohort size used in this chapter is another factor that cannot be discounted, particularly given the use of binary serostatus data rather than continuous data which is more conducive to demonstrate large differences in signal. For example, the

cohort sizes in the melanoma studies which showed a link of AAb development with irAE development were at least twice the size of the LUD2015-005 cohort evaluated in this chapter[316–318].

Another important consideration is that all assays described in this chapter have detected IgG based reactivity to antigens but not IgA reactivity, which may be more relevant in this cohort given the mucosal location of the tumour in OC. Further, there is no distinction between the isotypes of IgG AAb which are detected, which is an important consideration given that different isotypes of IgG can have differing effector functions[257]. For example, there could be a skew towards increased IgG4 AAbs, which have reduced effector functions[299], in patients who exhibit poor clinical outcomes, but this is not assessable by many antigen reactivity screening approaches which detect AAbs by the secondary antibody used in these assays which bind to all IgG subclasses.

3.9.3 Anti-p53 autoantibodies are a highly specific biomarker for distinguishing OC patients from BO patients

The binary serostatus data generated in this chapter was also utilised to determine whether AAb profiling could non-invasively distinguish OC patients from BO patients. Firstly, the number of reactive antigens were assessed based on the hypothesis that BO patients could have a greater baseline immune activation than OC patients and so better control growth of malignant cells. However, in contrast to the reported reduced cancer risk of TIF1-γ AAb seropositive DM patients who had a greater number of reactivities to other antigens[313], there was no significant difference in the number of reactive antigens of BO and OC patients.

Differential reactivity analysis for 172 antigens demonstrated the potential of reactivity to a N-terminal protein fragment of p53 as a highly specific biomarker for distinguishing OC patients from BO patients, although sensitivity for OC was only 11.6%. Caution should also be observed for BO patients who are p53 seropositive and have a history of cancer, regardless of the location of the malignancy.

Nonetheless, this finding should be validated with a large independent cohort, ideally with an alternative method of AAb detection which utilises full length and post translationally modified p53 such as yeast display. If validated, despite low sensitivity, a clinical test based on detecting anti-p53 AAbs could be adopted as a high-throughput and non-invasive monitoring approach for BO patients which would aid stratification of patients for endoscopic examination and detection of OC at an early stage.

3.9.4 Aetiology of autoantigen reactivity

As demonstrated in figure 3.36, in patients from the LUD2015-005 study, the mutation status of *TP53* did not associate with increased likelihood of anti-p53 AAb seropositivity[364,382,387]. As mentioned for the anti-p53 AAbs in BO and OC patients mentioned above, this finding should be validated using a detection method which utilises full length and post-translationally modified p53 protein as antigen rather than the unmodified protein fragments of p53 that were used in this chapter. If confirmed, the lack of association of *TP53* mutation status with anti-p53 AAb seropositivity would suggest that other factors such as high expression or host defects in tolerance may be more important for driving development of anti-p53 AAbs.

Contrastingly, as demonstrated by ZNF688, high antigen reactivity may be mediated by molecular mimicry which drives subsequent development of B cell clones with cross-reactive BCRs. It would be interesting to see if this is the case for other commonly reactive antigens identified in this chapter. This could be done using an approach similar to that which identified the shared epitope between ZNF688 and EBV-KITH[386]. However, this work is outside of the scope of this thesis.

3.10 Conclusion

Analysis of IgG AAb reactivity to two custom panels of autoantigens did not reveal significant associations with clinical outcome of ICI+CTX treated OC patients. However, a positive association with number of reactive antigens and good irRECIST responses in LAC patients treated with 1st line ICI was observed. This suggests that the number of reactive antigens may be a more pertinent predictor of good clinical outcome in patients treated with ICI in the absence of concurrent CTX. Similar analysis of IgG AAb reactivity in BO and OC patients led to the identification of anti-p53 AAbs as a promising peripheral blood biomarker of OC with high specificity but reasonably low sensitivity.

In the next chapter, data from profiling of PBMCs of ICI treated OC will be analysed to determine whether peripheral signatures of increased immune activation prior to ICI treatment are associated with good clinical outcome. This data will also be analysed, in combination with the antigen reactivity data described in this chapter, to identify features of PBMC profiles which may contribute to reactivity to a greater breadth of autoantigens.

4 Peripheral immune cell profiling to assess signatures of immune activation

4.1 Rationale

In the last chapter, although the expected association of development of on treatment irAE(s) with good clinical outcome to ICI+CTX was seen in the LUD2015-005 inoperable cohort, no significant association of IgG AAb reactivity with either patient clinical outcome or irAE development was observed. This could be because my hypothesis that AAb reactivity as a marker of baseline immune activation is incorrect, or that a higher baseline immune activation does not increase the likelihood of an OC patient showing good clinical outcome to ICI+CTX. To determine which of these possibilities is more likely, an alternative approach to measuring the baseline immune activation of the patients in the LUD2015-005 study was taken.

As briefly overviewed in chapter one, peripheral blood immune cells have been studied in multiple different cancer types to identify signatures that are present PreTx and which associate with good clinical outcome to ICI treatment[120,125,126,129,131,133–136]. Many of these studies identified signatures of increased activation and lower exhaustion of CD8+ T cells[120,125,126,131,135], which could reflect a higher baseline immune activation, as markers of increased likelihood of good clinical outcome to ICI. However, most of these studies only focused on T cells and so prognostic signatures from other immune cell types may have gone undetected.

Fortunately, recent advances in high dimension cytometry have made it feasible to obtain protein expression data for over 30 markers at a single cell resolution for millions of cells across tens and hundreds of samples[388,389], which allows for comprehensive *ex vivo* immunophenotyping of PBMCs. Indeed, studies of ICI treated melanoma and NSCLC patients have utilised such approaches to identify PreTx signatures related to activation of myeloid and NK cells which associate with good clinical outcome[133–136]. However, similar studies in OC have not been performed, even though such studies would be highly valuable for increasing understanding of mechanisms underlying good clinical outcome and development of irAEs during ICI treatment.

Similarly, comprehensive *ex vivo* phenotyping studies of PBMCs from BO and OC patients would be valuable for discovering signatures which distinguish BO patients from OC patients. Like the increased prevalence of anti-p53 AAbs in OC patients that was discussed in the previous chapter, any such signatures would be valuable for both early detection of OC and further understanding the mechanisms involved in transition of BO to OC.

This chapter addresses my aim to analyse whether peripheral signatures of increased immune activation are associated with improved clinical outcome and greater AAb numbers in OC patients treated with ICI+CTX. I hypothesise that patients that experience CB have an increased baseline immune activation which can be measured as immune signatures in the peripheral blood. Secondary aims of this chapter are to discover peripheral immune signatures that 1) are associated with high risk of irAE development in ICI+CTX treated OAC patients, and 2) can non-invasively distinguish OC patients from BO patients, which could be used in combination with anti-p53 AAb detection to stratify endoscopic screening of BO patients who are at high-risk of developing OC.

4.2 Samples, experimental approach, and data processing

To analyse peripheral signatures of immune activation, *ex vivo* immunophenotyping by high dimensional spectral flow cytometry of PBMCs samples using a Cytek Aurora was performed in collaboration with Professor Persephone Borrow and Dr Isabela Pedroza-Pacheco (NDMRB).

4.2.1. Samples

In total 98 PBMC samples were immunophenotyped, including samples from 60 different Individuals. 12 of these individuals were HVs aged 50-60 years old, 12 of these individuals were BO patients, and 36 of these were OC patients from the inoperable arm of the LUD2015-005. Out of the 74 total samples from the LUD2015-005 patients, 36 were collected PreTx, 33 were collected at the ICI-4W timepoint,

and five samples were collected at least 12 months after completion of treatment from LTRs (**Figure 4.1**). Full details regarding these samples are described in chapter 2 sections 2.1 and 2.2.

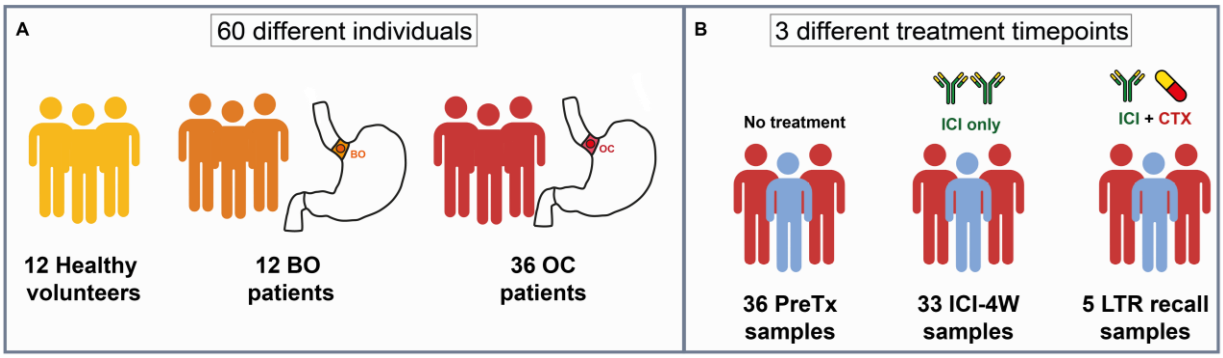


Figure 4.1. Summary of peripheral blood mononuclear cell samples used for the *ex vivo* immunophenotyping experiments described in this chapter. Overall, 98 PBMC samples were used. (**A**) These included samples from 60 different individuals. 12 of these individuals were healthy volunteers aged 50-60 years old, 12 of these individuals were patients with diagnosed Barrett's Oesophagus (BO), and 36 of these individuals were patients with diagnosed oesophageal cancer (OC). (**B**) In total there were 74 samples from the 36 OC patients, all from the LUD2015-005 study inoperable arm, which corresponded to one of three different treatment timepoints. 36 of the samples were collected prior to treatment (PreTx), 33 samples were collected after four weeks of ICI treatment (ICI-4W), and five samples were collected from long-term responders, based on them achieving at least 18 months from treatment initiation without clinical progression, (LTR recall). Created using Adobe illustrator and including icons from Biorender.com.

4.2.2. Experimental approach and data processing

In order to sufficiently characterise the major cell types present in the PBMCs of these samples, three panels, each of 33 to 35 fluorophores, were designed, optimised and validated by Dr Isabela Pedroza-Pacheco. The first of these panels was primarily focused on immunophenotyping T cells, referred to as Panel 1 (T cell centric), and was composed of 33 fluorophores which corresponded to markers of lineage, activation, inhibition, exhaustion, and function of T cells. The second of these panels was primarily focused on immunophenotyping NK cells, referred to as Panel 2 (NK cell centric), and was composed of 33 fluorophores which corresponded to markers of lineage, activating receptors, inhibitory receptors, and function of NK cells. The third and final of these panels was primarily focused on immunophenotyping B cells and myeloid cells, referred to as Panel 3 (B cell and myeloid centric), and was composed of 35 fluorophores which corresponded to markers of myeloid cell lineage, myeloid cell function, B cell lineage, B cell activation, and B cell BCR/antibody isotype class. Each of these three panels also included a cell viability dye which was used to identify live cells (**Figure 4.2**). Full details of

the fluorophores, markers, and antibodies used in each of these panels are described in chapter 2 section 2.10.1.

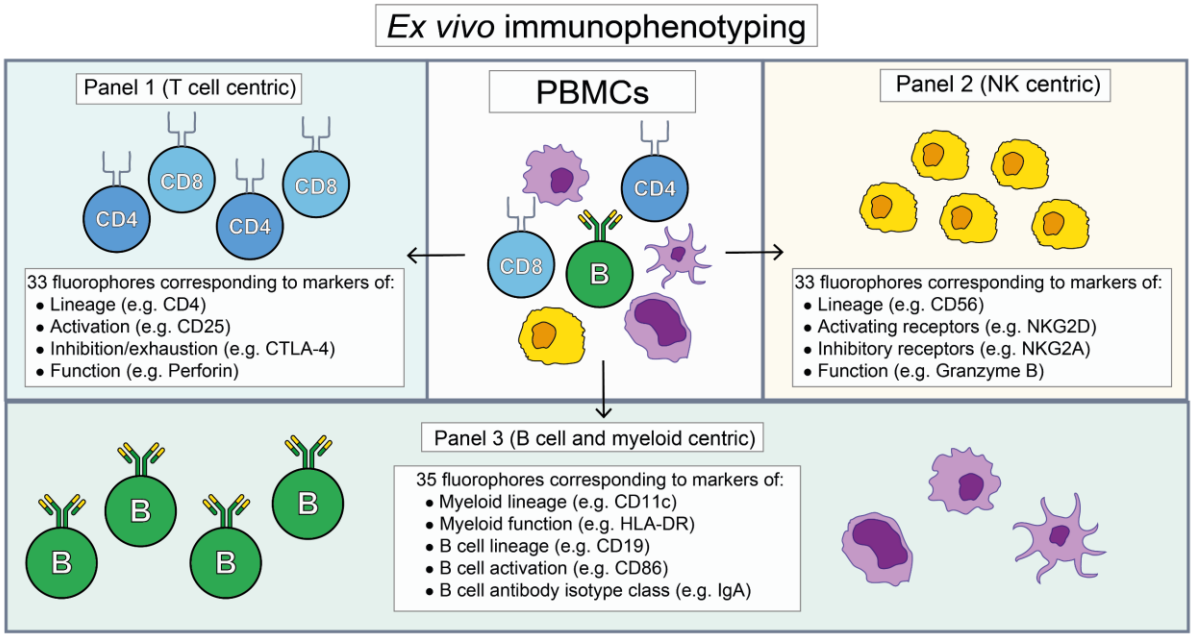


Figure 4.2. Overview of panels used for *ex vivo* immunophenotyping of PBMCs. Each PBMC sample was divided into three and then stained for one of three panels of markers. Panel one consisted of 33 markers which were primarily focused on T cells. Panel two consisted of 33 markers primarily focused on natural killer (NK) cells. Panel three consisted of 35 markers primarily focused on B cells or myeloid cells. One marker within each panel was a cell viability dye which was used to gate live cells.

To identify cell subsets and perform differential expression analysis between groups of samples, a semi-supervised approach based on the published CyTOF workflow[334,390] was taken to analyse the combined data from these three spectral flow cytometry panels. Firstly, unmixed flow cytometry data was imported into FlowJo™ Software[330] and all of the live mononuclear cells (MNCs) from each file were gated and exported as a new file. For these files of live MNCs, variance was stabilised on a channel-specific basis using the FlowVS package[331], then low quality events were removed using the PeacoQC package[332], next batch normalisation was performed using the CytoNorm package[333]. Full details are described in chapter 2 section 2.11.2. This quality-controlled batch normalised data was then imported into FlowJo™ Software[330] and a total of 15 major cell populations, across the three panels, were gated (**Figure 4.3**).

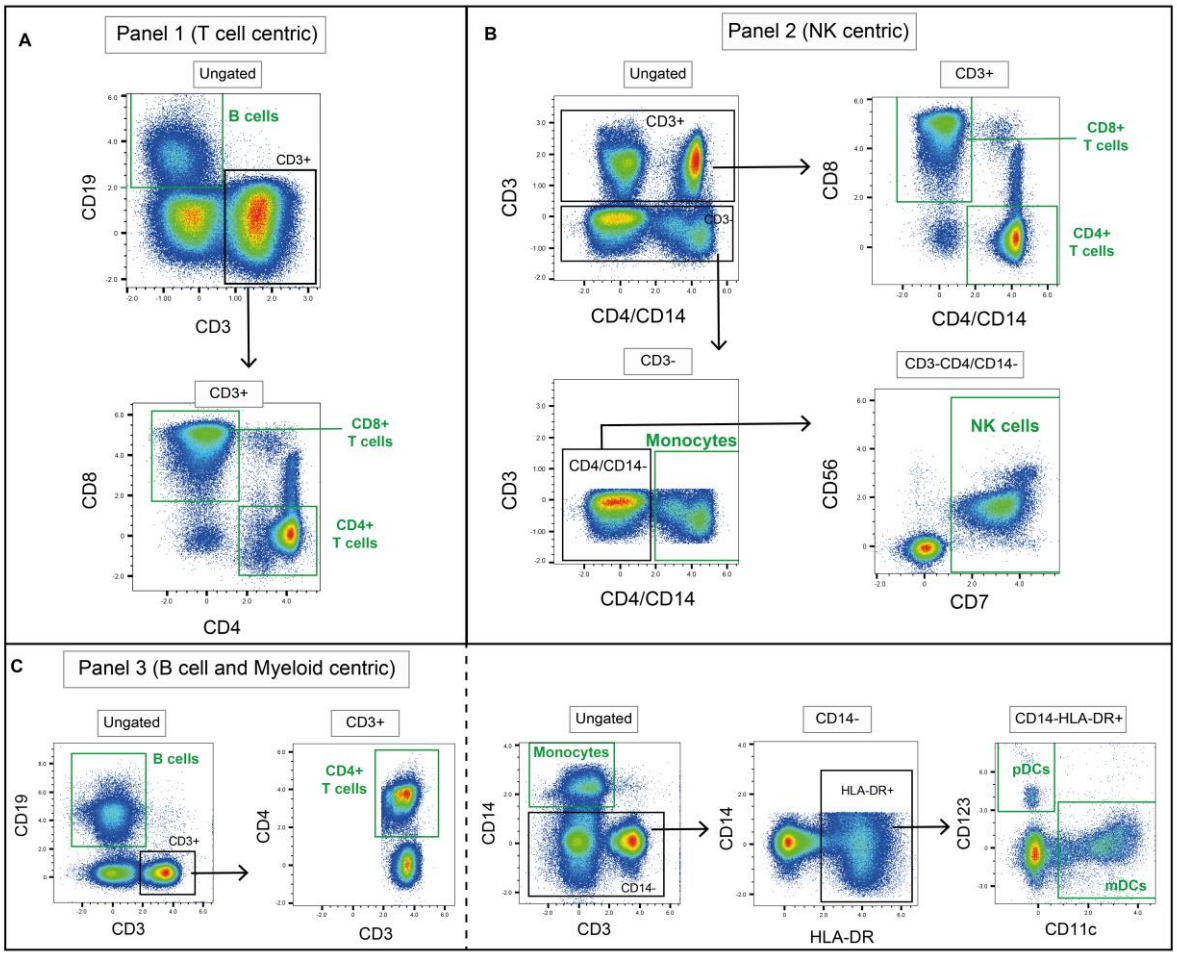


Figure 4.3. Gating strategies used for the major cell types identified in each of the three panels. (**A-C**) Gating of major cell populations from processed and normalised flow cytometry data was performed using FlowJo™ Software. Ungated refers to live singlet MNCs. The gating strategy to gate this population for each panel is shown in chapter 2. Gates and names used for down sampling of the major cell populations are shown in green. Gates and names for intermediate populations are shown in black. (**A**) For panel one, B cells were gated as CD19+CD3-, CD4+ T cells were gated as CD19-CD3+CD4+CD8-, and CD8+ T cells were gated as CD19-CD3+CD4-CD8+. (**B**) For panel two, CD4+ T cells were gated as CD3+CD4/CD14+CD8-, CD8+ T cells were gated as CD3+CD4/CD14-CD8+, monocytes were gated as CD3-CD4/CD14+, and NK cells were gated as CD3-CD4/CD14-CD7+. (**C**) For panel three, B cells were gated as CD19+CD3-, CD4+ T cells were gated as CD19-CD3+CD4+, monocytes were gated as CD3-CD14+, mDCs were gated as CD14-HLA-DR+CD11c+CD123-, and pDCs were gated as CD14-HLA-DR+CD11c-CD123+. Dashed line separates the gating strategies used for lymphoid and myeloid populations.

For each sample, a set number of cells from each of the 15 gated populations were then randomly down sampled using FlowJo™ Software and the median normalised expression of each marker was calculated using the R package CATALYST[334] (**Figure 4.4**).

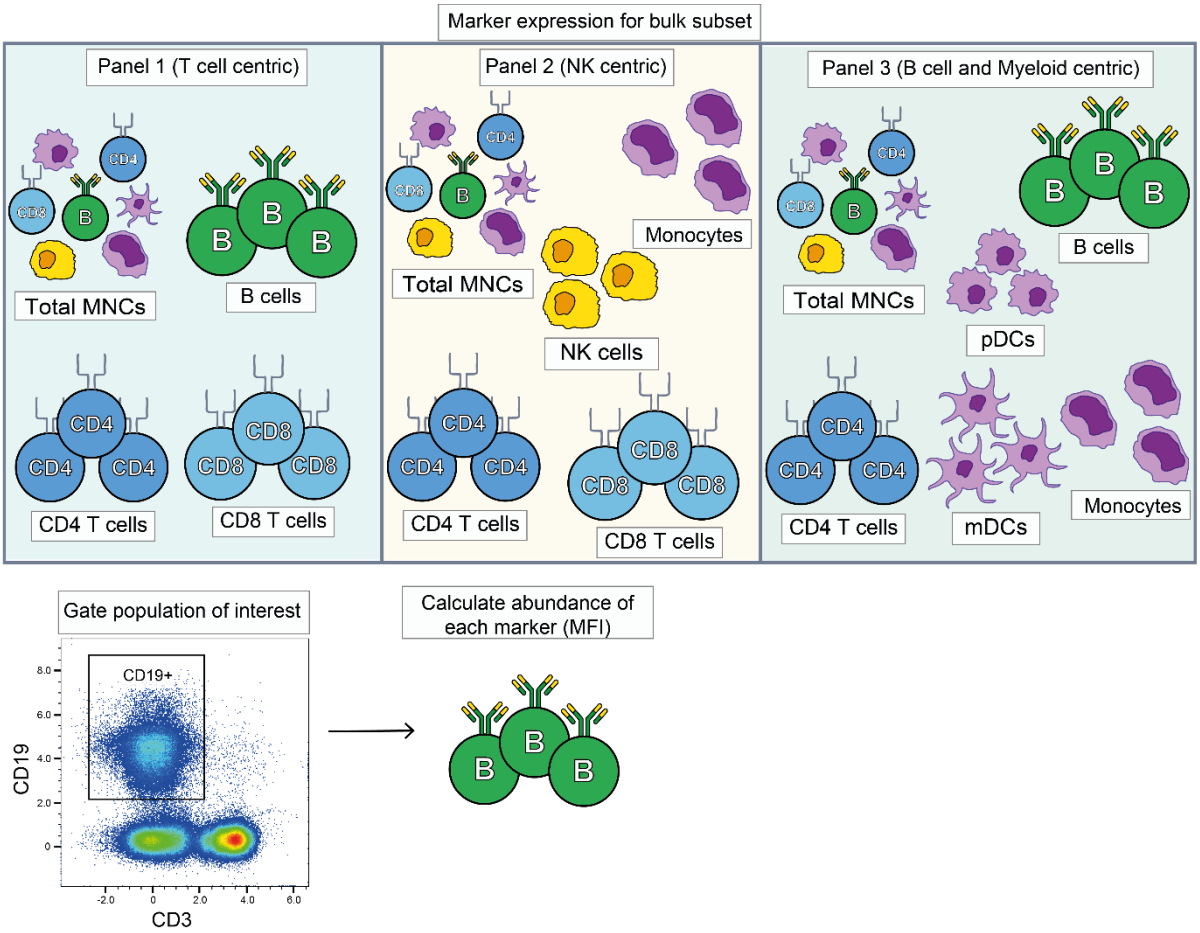


Figure 4.4. Marker expression was calculated for 15 major cell types which were gated and down sampled in FlowJo™ Software. Following processing and batch normalisation of flow cytometry data, a total of 15 major cell populations, across the three panels, were gated in FlowJo™ Software. For panel one, total mononuclear cells (MNCs), B cells, CD4-positive (CD4+) T cells, and CD8-positive (CD8+) T cells were gated. For panel two, total MNCs, monocytes, natural killer (NK) cells, CD4+ T cells, and CD8+ T cells were gated. For panel three, total MNCs, B cells, plasmacytoid dendritic cells (pDCs), myeloid dendritic cells (mDCs), monocytes, and CD4+ T cells were gated. For each sample, a set number of cells from each gated population were then randomly down sampled using FlowJo™ Software and the median normalised expression of each marker was calculated using the R package CATALYST.

To minimise the influence of sample specific variation, the same number of cells were down sampled per sample for analysis. The one exception to this was plasmacytoid DCs (pDCs), due to the very low frequency at which they are found in PBMCs, and so for which all cells were taken forward for analysis. The counts of pDCs per sample and per group are shown below (**Figure 4.5**).

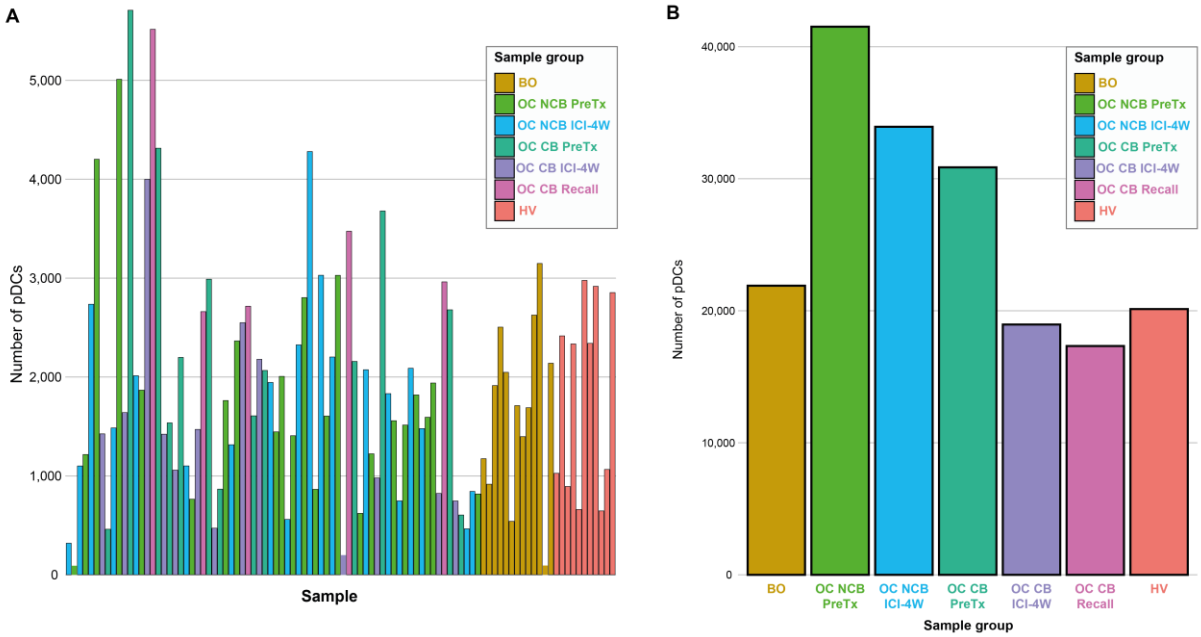


Figure 4.5. Cell counts for pDCs per sample and condition group. (**A-B**) Histograms of number of plasmacytoid dendritic cells (pDCs) for (**A**) each sample or (**B**) each sample group. Bars coloured by sample group. BO = Barrett's Oesophagus, OC NCB PreTx = Oesophageal Cancer No Clinical Benefit Pre-treatment, OC NCB ICI-4W = Oesophageal Cancer No Clinical Benefit post ICI treatment for four weeks, OC CB PreTx = Oesophageal Cancer Clinical Benefit Pre-treatment, OC CB ICI-4W = Oesophageal Cancer Clinical Benefit post ICI treatment for four weeks, OC CB Recall = Oesophageal Cancer Clinical Benefit long-term responder recall, HV = Healthy Volunteer.

In summary, for the PBMC samples stained with markers from the T cell centric panel 1, four major cell types were gated and taken forward for analysis: MNCs, B cells, CD4+ T cells, and CD8+ T cells (**Table 4.1**).

Major cell type	*MNCs*	*B cells*	*CD4 T cells*	*CD8 T cells*
Markers used for gating	NA	CD3-CD19+	CD19- CD3+CD4+CD8-	CD19- CD3+CD8+CD4-
Number of cells per sample used for analysis	80,000	21,000	50,000	21,000
Number of samples with at least the required number of cells	96/97	65/97	89/97	87/97

Table 4.1. Summary of down sampling performed in FlowJo™ Software from the samples stained with markers from panel one. Numbers of samples does not include the control samples which were included in each of the

5 batches and used for batch normalisation. N.B. One of the 98 samples did not have enough live cells to stain with this panel, so data for only 97 samples were acquired.

For the PBMC samples stained with markers from the NK cell centric panel 2, five major cell types were gated and taken forward for analysis: MNCs, monocytes, NK cells, CD4+ T cells, and CD8+ T cells (**Table 4.2**).

Major cell type	*MNCs*	*Monocytes*	*NK cells*	*CD4 T cells*	*CD8 T cells*
Markers used for gating	NA	CD3- CD4/CD14+	CD3- CD4/CD14- CD7+	CD3+CD4/CD14+	CD3+CD8+
Number of cells per sample used for analysis	73,000	20,000	20,000	50,000	22,000
Number of samples with at least the required number of cells	98/98	92/98	77/98	85/98	85/98

Table 4.2. Summary of down sampling performed in FlowJo™ Software from the samples stained with markers from panel two. Numbers of samples does not include the control samples which were included in each of the 5 batches and used for batch normalisation.

For the PBMC samples stained with markers from the B cell and myeloid centric panel 3, six major cell types were gated and taken forward for analysis: MNCs, B cells, CD4+ T cells, monocytes, mDCs, and pDCs (**Table 4.3**).

Major cell type	*MNCs*	*B cells*	*CD4 T cells*	*Monocytes*	*mDCs*	*pDCs*
Markers used for gating	NA	CD3-CD19+	CD3+CD4+	CD3-CD14+	CD14-HLA-DR+CD11c+CD123-	CD14-HLA-DR+CD11c-CD123+
Number of cells per sample used for analysis	78,000	20,000	52,000	22,000	20,000	All (91-5,709)
Number of samples with at least the required number of cells	96/98	78/98	91/98	73/98	81/98	98/98

Table 4.3. Summary of down sampling performed in FlowJo™ Software from the samples stained with markers from panel three. Numbers of samples does not include the control samples which were included in each of the 5 batches and used for batch normalisation.

Following down sampling, the median expression of each marker for each the major cell types gated were calculated. Each cell type showed the expected profile of marker expression (**Figure 4.6**).

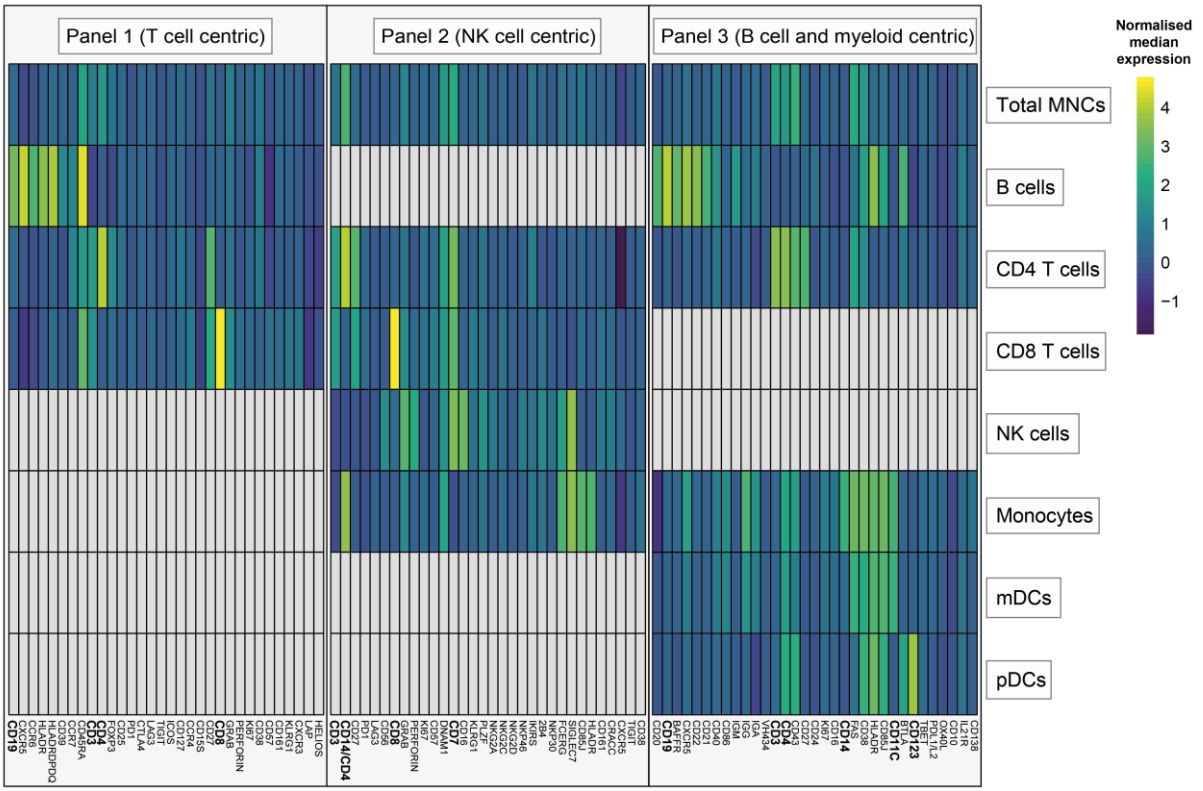


Figure 4.6. Expression of markers within each major cell type identified. Heatmap of median expression of each marker. Cell types as rows and markers, grouped by panel, as columns. Lineage markers in bold. Grey cells indicate where expression data for that marker was not available for that cell type.

To further characterise the cell types, dimensionality reduction and clustering of marker expression data from the major cell types gated into cell subsets was performed using the packages FlowSOM[335] and CATALYST[334]. This analysis was performed on the 10 of the 15 major cell subsets gated which were deemed to have sufficient markers to identify previously reported subsets (**Figure 4.7**).

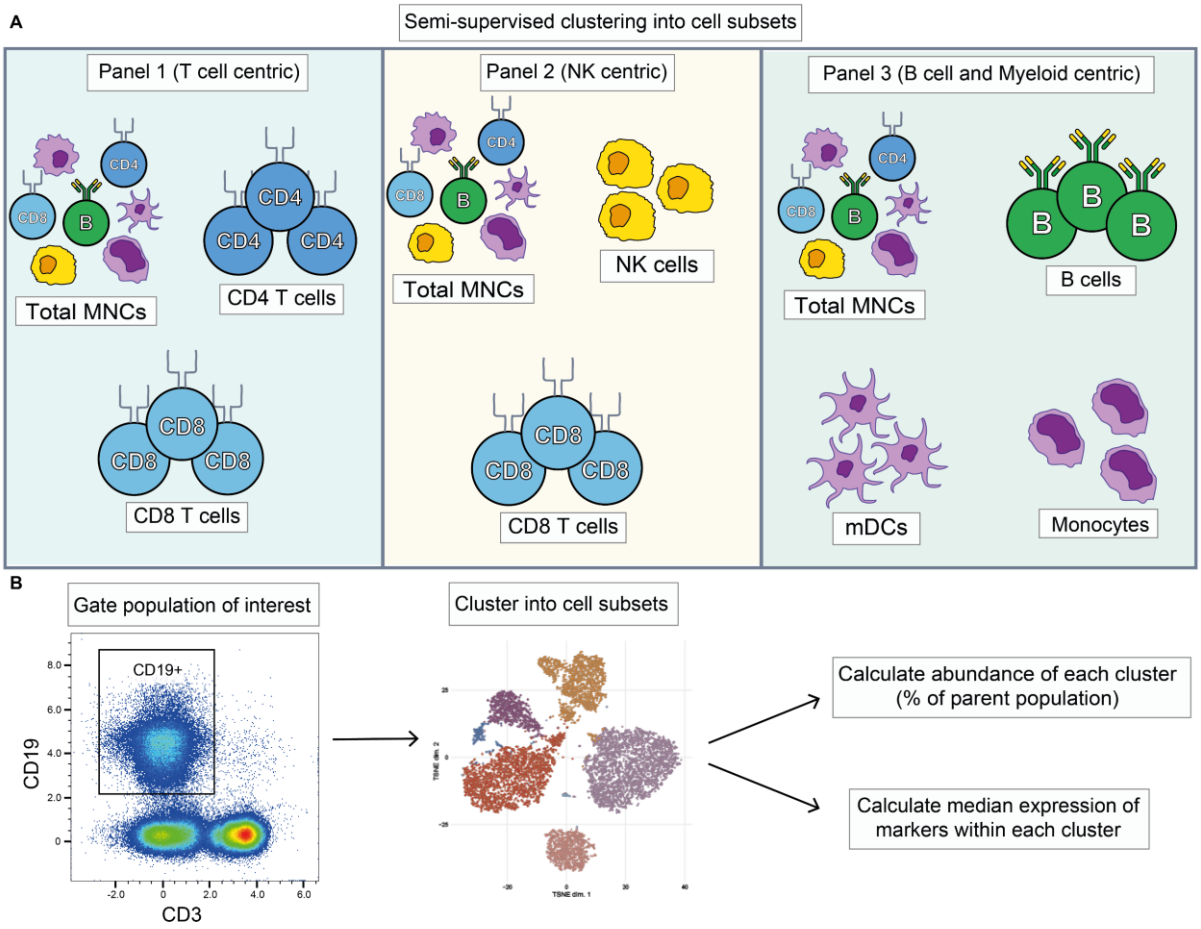


Figure 4.7. Cell populations FlowJo™ Software down sampling and clustering performed on for each of the three panels. Expression data for 10 of the 15 major cell populations, across the three panels, which were gated and down sampled in FlowJo™ Software, were subjected to dimensionality reduction and clustering in order to identify cell subsets. For panel one, total mononuclear cells (MNCs), CD4-positive (CD4+) T cells, and CD8-positive (CD8+) T cells were clustered. For panel two, total MNCs, natural killer (NK) cells, and CD8+ T cells were clustered. For panel three, total MNCs, B cells, myeloid dendritic cells (mDCs), and monocytes were clustered. For each sample, the abundance of cells within each cluster and the median normalised expression of markers within each cluster was calculated using the R package CATALYST.

4.3 Identifying cell subsets using unsupervised clustering

To aid robust identification of well-established cell subsets, for each major cell type, a subset of markers, henceforth referred to as 'type' markers, was chosen based on expression profiles previously reported in the literature and this subset of markers was used to cluster the cells into subsets. For the remaining markers, henceforth referred to as 'state' markers, their expression within each identified cluster was calculated.

4.3.1 Clustering of mononuclear cells identifies multiple cell subsets

For total MNCs, there were sufficient markers to identify cell subsets in all three panels and so three sets of dimensionality reduction and clustering were performed.

4.3.1.1 Seven subsets of MNCs identified by clustering with the markers from panel 1 (T cell centric)

96 of 97 PBMC samples which were stained with markers from panel one had at least 80,000 live MNCs. For these samples, 80,000 live MNCs were randomly down sampled using FlowJo™ Software, for a total of 7,680,000 cells, and taken forward for semi-supervised clustering into cell subsets.

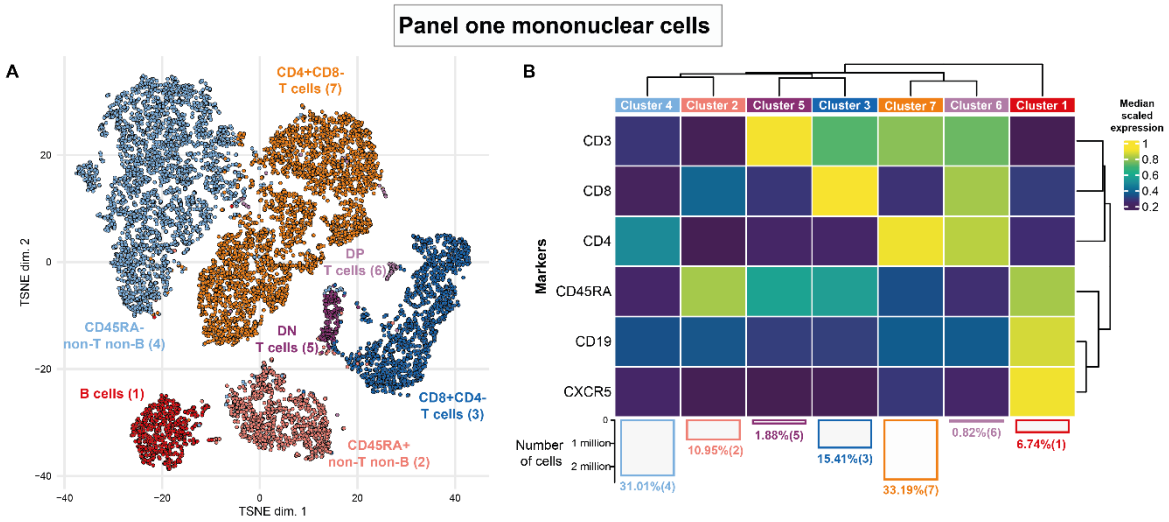


Figure 4.8. Clustering of total MNCs with the markers in panel one identified seven clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel one, by mononuclear cells (MNCs) identified seven major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the six markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the six markers used for clustering, for each of the seven clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the MNCs from the samples which were stained with panel one were the T cell lineage markers CD3, CD4 and CD8; CD45RA which is expressed by a range of MNCs[391–394]; and CD19 and CXCR5 which are highly expressed by most B cells[395]. Subsequent dimensionality reduction and clustering identified seven major clusters of cells (**Figure 4.8**). Cluster 1 was identified as a B cell cluster due to the high expression of the B cell marker CD19 and high expression of CXCR5

and CD45RA which are known to expressed by most circulating B cells[391,395]. Cluster 2 appeared to be CD45RA+ non-T non-B cells due to the lack of expression of CD3, CD19 but high expression of CD45RA. This cluster is likely to be made of NK cells and monocytes, which are known to express CD45RA when activated[392,393], and pDCs which constitutively express CD45RA[394]. Cluster 3 was identified as CD8+CD4- T cells due to the high expression of CD3 and CD8 in combination with lack of CD4 expression. Cluster 4 appeared to be CD45RA- non-T non-B cells due to the lack of expression of CD3, CD19 and CD45RA. This cluster is likely to be made of non-activated NK and myeloid cells, and mDCs which do not express CD45RA[394]. Cluster 5 seemed to be a cluster of double negative (DN) T cells, based on their expression of CD3 but lack of expression of either CD4 or CD8, which have been previously described[396]. Cluster 6 appeared to a cluster of double positive (DP) T cells, based on expression of high levels of CD3, CD4 and CD8, which have also been previously described[397]. Finally, cluster 7 was identified as CD4+CD8- T cells due to the high expression of CD3 and CD4 in combination with lack of CD8 expression.

4.3.1.3 Six subsets of MNCs identified by clustering with the markers from panel 2 (NK cell centric)

98 of 98 PBMC samples which were stained with markers from panel two had at least 73,000 live MNCs. For these samples, 73,000 live MNCs were randomly down sampled using FlowJo™ Software, for a total of 7,154,000 cells, and taken forward for semi-supervised clustering into cell subsets.

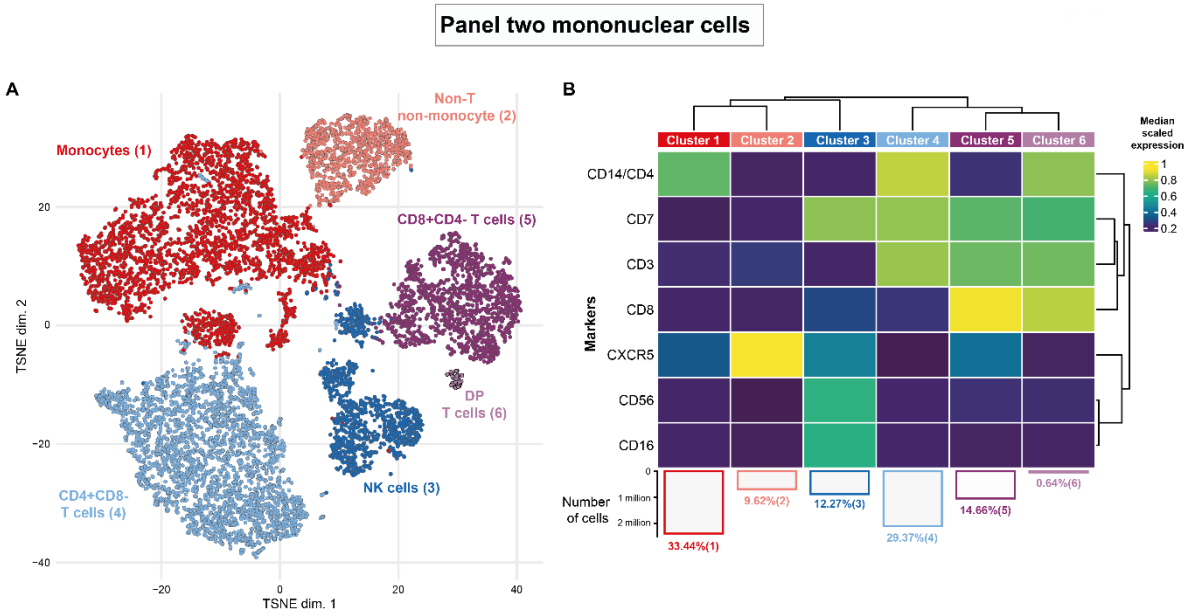


Figure 4.9. Clustering of total MNCs with the markers in panel two identified six clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel two, by mononuclear cells (MNCs) identified six major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the seven markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the six markers used for clustering, for each of the seven clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the MNCs from the samples which were stained with panel two were the T cell lineage markers CD3, CD8 and CD4 (although the anti-CD4 antibody used was conjugated to the same fluorophore as the anti-CD14 antibody used so their expression was indistinguishable); the myeloid cell lineage marker CD14; the NK cell lineage markers, CD7, CD16 and CD56[398]; and CXCR5. Subsequent dimensionality reduction and clustering identified six major clusters of cells (**Figure 4.9**). Cluster 1 appeared to be a cluster of monocytes, due to high expression of CD14/CD4 coupled with no CD3 or CD7 expression. This cluster could also include CD4+ T cells that have very low CD3 expression, although this population is believed to be very rare in *ex vivo* PBMCs[399]. Cluster 2 appeared to be a cluster of CXCR5+ non-T non-monocyte cells, so most likely predominantly mature B cells, based on high expression of CXCR5 coupled with no CD3, CD7, CD56 or CD14/CD4 expression. Cluster 3 was identified as NK cells due to the high expression of CD7, CD16 and CD56 in combination with lack of CD3 expression. Cluster 4 was identified as CD4+CD8- T cells due to the high

expression of CD3, CD4 and CD7 in combination with lack of CD8 expression. Cluster 5 was identified as CD8+CD4- T cells due to the high expression of CD3, CD7 and CD8 in combination with lack of CD4 expression. Finally, cluster 6 appeared to be a cluster of DP T cells similar to cluster 6 identified in the clustering of MNCs based on panel one markers.

4.3.1.3 Seven subsets of MNCs identified by clustering with the markers from panel 3 (B cell and myeloid centric)

96 of 98 PBMC samples which were stained with markers from panel three had at least 78,000 live MNCs. For these samples, 78,000 live MNCs were randomly down sampled using FlowJo™ Software, for a total of 7,488,000 cells, and taken forward for semi-supervised clustering into cell subsets.

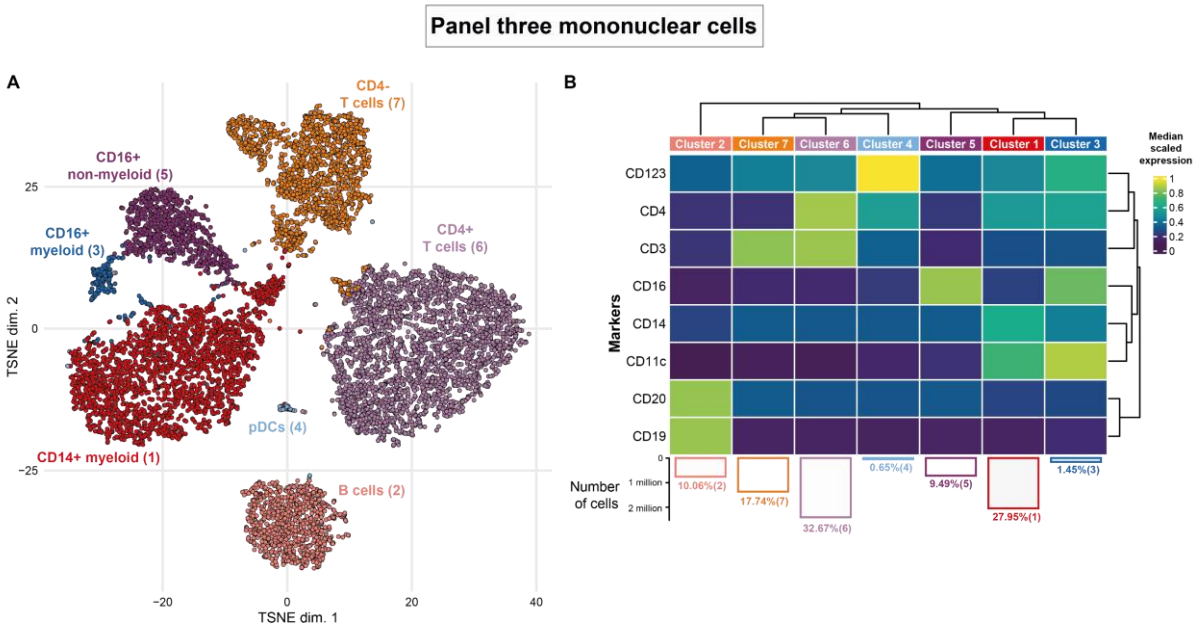


Figure 4.10. Clustering of total MNCs with the markers in panel three identified seven major clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel three, by mononuclear cells (MNCs) identified seven major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the seven markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the eight markers used for clustering, for each of the seven clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the MNCs from the samples which were stained with panel three were the T cell lineage markers CD3, and CD4; the B cell lineage markers CD19 and CD20; the myeloid cell lineage markers CD11c, CD14, CD16; and CD123 which is highly expressed by pDCs[394]. Subsequent

dimensionality reduction and clustering identified seven major clusters of cells (**Figure 4.10**). Cluster 1 was identified as a CD14+ myeloid cell cluster due to the high expression of the myeloid cell marker CD11c and high expression of CD14. Cluster 2 was identified as a B cell cluster due to the high expression of the B cell markers CD19 and CD20. Cluster 3 was identified as a CD16+ myeloid cell cluster due to the high expression of the myeloid cell marker CD11c and high expression of CD16. Cluster 4 was identified as a pDC cluster due to the lack of expression of CD3, CD19 and CD11c in combination with high expression of CD123. Cluster 5 appeared to be CD16+ non-myeloid cells due to high expression of CD16 but low expression of CD11c. Cluster 6 was identified as a CD4+ T cell cluster based on high expression of CD3 and CD4. This cluster is likely to be predominantly CD4+CD8- T cells but may also include DP T cells such as those seen in cluster 6 which was identified in the clustering of MNCs based on panel one markers. Finally, cluster 7 was identified as a CD4- T cell cluster based on high expression of CD3 but no expression of CD4. This cluster is likely to predominantly consist of CD8+CD4- T cells but can also include DN T cells such as those seen in cluster 5 which was identified in the clustering of MNCs based on panel one markers.

4.3.2. Nine subsets of B cells identified by clustering with the markers from panel 3 (B cell and myeloid centric)

For B cells, there were only sufficient markers to identify cell subsets in the B cell and myeloid centric panel (panel three) and so only one set of dimensionality reduction and clustering was performed.

Based on CD19 expression and lack of CD3 expression, 78 of 98 PBMC samples which were stained with markers from panel three had at least 20,000 live B cells. For each of these samples, 20,000 live B cells were randomly down sampled using FlowJo™ Software, for a total of 1,560,000 cells, and taken forward for semi-supervised clustering into cell subsets.

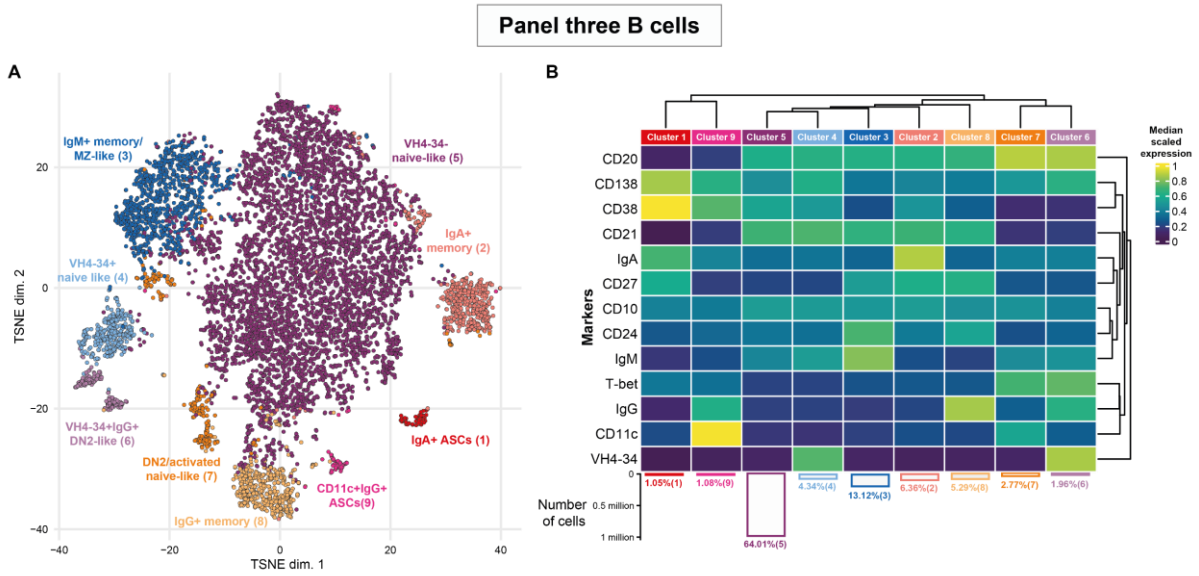


Figure 4.11. Clustering of CD19+ B cells with the markers in panel three identified nine clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel three, by CD19+ B cells identified nine major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the 13 markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the 13 markers used for clustering, for each of the nine clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the B cells from the samples which were stained with panel three were the B cell lineage markers CD10, CD20, CD21, CD24, CD27, and CD38[400]; the BCR antibody isotypes IgA, IgG, and IgM (N.B. IgD was not included for all of the batches due to issues with the antibody used and so all IgD data was excluded from analysis); markers which have been associated with autoreactive B cells in SLE, CD11c, T-bet, and VH4-34[348,400–403]. Subsequent dimensionality reduction and clustering identified nine major clusters of cells (**Figure 4.11**). Cluster 1 appeared to be a cluster of IgA+ ASCs due to high expression of CD38, CD138, and IgA in combination with lack of CD20 expression. Cluster 2 appeared to be a cluster of IgA+ memory B cells due to high expression of CD38, CD138, and IgA. Cluster 3 appeared to be a cluster of IgM+ memory/MZ-like B cells due to high expression of IgM, CD21, CD24 and CD27 in combination with lack of CD38 expression. Cluster 4 was identified as VH4-34+ naïve-like B cells due to the high expression of VH4-34 and CD21 in combination with low expression of CD24 and IgM, and lack of expression of CD27, IgG and IgA. Cluster 5 was identified as VH4-34- naïve-like B cells due to the similar expression profile as cluster 4, except with lack of VH4-34 expression. Cluster 6 appeared to be a cluster of VH4-34+IgG+ cells with a similar

expression profile to so-called double negative (DN) B cells which lack conventional markers of naïve and memory B cells[400]. Subsequently, this cluster was named as VH4-34+IgG+DN2-like B cells due to the high expression of VH4-34, IgG, and T-bet in combination with lack of expression of IgM, CD21, CD24, CD27 and CD38. Cluster 7 appeared to be a cluster of DN2/activated naïve-like B cells due to the high expression of T-bet and CD11c in combination with low expression of IgM, and lack of expression of CD21, CD24, CD27, CD38, IgG and IgA. Cluster 8 was identified as IgG+ memory B cells due to the high expression of IgG, CD21 and CD27 in combination with lack of expression of CD38. Finally, cluster 9 was identified as IgG+ CD11c+ ASCs due to the high expression of CD11c, IgG, CD38, and CD138 in combination with lack of CD20 expression.

4.3.3. Seven subsets of CD4 T cells identified by clustering with the markers from panel 1 (T cell centric)

For CD4+ T cells, there were only sufficient markers to identify cell subsets in the T cell centric panel (panel one) and so only one set of dimensionality reduction and clustering was performed.

Based on CD3 and CD4 expression, 89 of 98 PBMC samples which were stained with markers from panel one had at least 50,000 live CD4+ T cells. For each of these samples, 50,000 live CD4+ T cells, for a total of 4,450,000 cells, were randomly down sampled using FlowJo™ Software and taken forward for semi-supervised clustering into cell subsets.

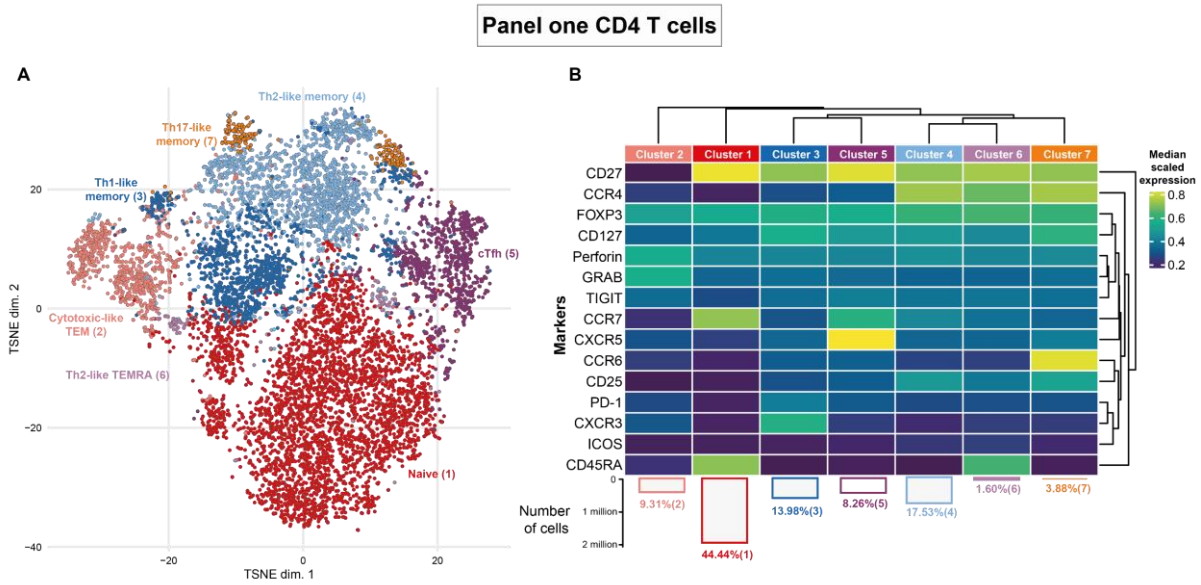


Figure 4.12. Clustering of CD4+ T cells with the markers in panel one identified seven clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel one, by CD4+ T cells identified seven major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the 15 markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the 15 markers used for clustering, for each of the seven clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the CD4 T cells from the samples which were stained with panel one were the T cell lineage markers CD45RA, CCR7, and CD27[404]; the markers of different 'helper' T cell types CXCR3, CXCR5, CCR4, CCR6, PD-1, TIGIT, and ICOS[404–407]; FOXP3, CD25 and CD127 to identify Tregs[404]; Perforin, and Granzyme B (GRAB) as markers of cytotoxic cells[408]. Subsequent dimensionality reduction and clustering identified seven major clusters of cells (**Figure 4.12**). Cluster 1 appeared to be a cluster of naïve CD4+ T cells due to high expression of CD45RA, CCR7 and CD27. Cluster 2 appeared to be a cluster of cytotoxic-like effector memory (TEM) cells due to expression of perforin and GRAB in combination with lack of expression of CD45RA and CCR7[408]. Cluster 3 appeared to be a cluster of Th-1-like memory cells due to high expression of CXCR3 in combination with lack of CCR6, CCR4 and CD45RA expression[404,406,409]. Cluster 4 was identified as Th2-like memory cells due to the high expression of CCR4 and lack of expression of CXCR3, CCR6 and CD45RA[406]. Cluster 5 was identified as circulating Tfh (cTfh) cells due to high expression of CXCR5[405]. Cluster 6 appeared to be a cluster of Th2-like memory T cells re-expressing CD45RA (TEMRA) cells due to the similar expression profile as

cluster 4, except with expression of CD45RA[410]. Finally, cluster 7 appeared to be a cluster of Th17-like memory cells due to high expression of CCR6 and CCR4 in combination with lack of expression of CXCR3 and CD45RA[406]. No obvious Treg cluster with the typical CD25+FoxP3+CD127- expression profile was identified, with consistently low FoxP3 expression seen across all clusters.

4.3.4. Clustering of CD8 T cells identifies multiple cell subsets

For CD8+ T cells, there were sufficient markers to identify cell subsets in both the T cell centric panel (panel one) and the NK centric panel (panel two). Therefore, two sets of dimensionality reduction and clustering were performed.

4.3.4.1 Nine subsets of CD8 T cells identified by clustering with the markers from panel 1 (T cell centric)

Based on CD3 and CD8 expression, 87 of 98 PBMC samples which were stained with markers from panel one had at least 21,000 live CD8+ T cells. For each of these samples, 21,000 live CD8+ T cells, for a total of 1,827,000 cells, were randomly down sampled using FlowJo™ Software and taken forward for semi-supervised clustering into cell subsets.

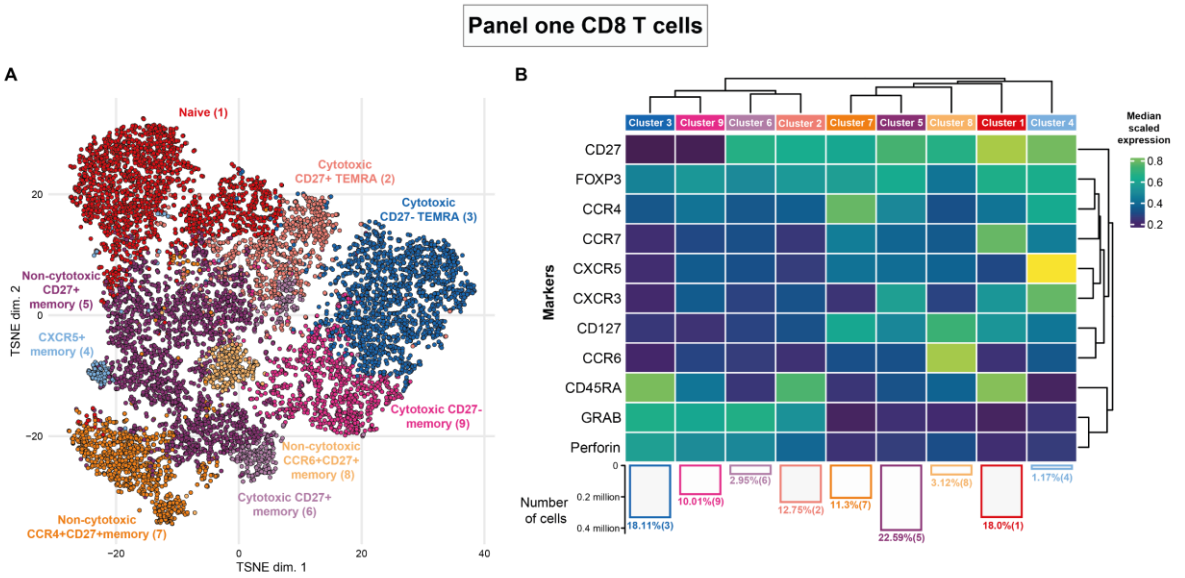


Figure 4.13. Clustering of CD8+ T cells with the markers in panel one identified nine clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel one, by CD8+ T cells identified nine major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the 11 markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the 11 markers used for clustering, for each of the nine clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the CD8 T cells from the samples which were stained with panel one were the T cell lineage markers CD45RA, CCR7, and CD27; FOXP3 and CD127 as markers of CD8 suppressor T cells[411]; the chemokine receptors CXCR3, CCR4, and CCR6 as markers of tissue homing[412]; CXCR5 as a marker of circulating follicular helper CD8 T cells[413]; and perforin, and GRAB as markers of cytotoxic cells.

Subsequent dimensionality reduction and clustering identified eight major clusters of cells (**Figure 4.13**). Cluster 1 was identified as naïve CD8+ T cells due to high expression of CD45RA, CCR7 and CD27. Cluster 2 was identified as Cytotoxic CD27+ TEMRA cells due to expression of CD45RA and CD27 in combination with low expression of perforin and GRAB, and lack of expression of CCR7[414–418]. Cluster 3 appeared to be a cluster of cytotoxic CD27- TEMRA cells due to the similar expression profile as cluster 2, except with increased perforin expression and lack of CD27 expression[414–418]. Cluster 4 was identified as CXCR5+ memory CD8 T cells due to high expression of CXCR5 in combination with low

expression of CCR7 and lack of CD45RA expression. This cluster of cells has a similar expression profile as previously described circulating follicular helper CD8 T cells which have been implicated as mediators of antibody production[413,419,420]. Cluster 5 was identified as non-cytotoxic CD27+ memory cells due to expression of CD27 and CXCR3 in combination with low expression of CCR7 and lack of expression of GRAB, perforin and CD45RA. Cluster 6 appeared to be a cluster of cytotoxic CD27+ memory cells due to the similar expression profile as cluster 2, except with increased expression of GRAB and perforin, and lack of CD45RA expression. Cluster 7 appeared to be a cluster of non-cytotoxic CD27+CCR4+ memory cells due to the high expression of CCR4, CD27 and CD127 in combination with low expression of CCR7 and GRAB, and lack of expression of CD45RA and perforin[421]. Cluster 8 appeared to be a cluster of non-cytotoxic CCR6+ CD27+ memory cells due to high expression of CCR6 and CD27 in combination with low expression of CCR7 and lack of expression of CD45RA, perforin and GRAB. Cells with a similar CCR6+CD8+ expression profile have been suggested to be associated with mucosal immunity, or have been found to express IL-17 and subsequently called Tc17 cells[422,423]. Finally, cluster 9 appeared to be a cluster of cytotoxic CD27- TEM cells due to the similar expression profile as cluster 3, except with lack of CD45RA expression.

4.3.4.2 Seven subsets of CD8 T cells identified by clustering with the markers from panel 2 (NK cell centric)

Given previous literature demonstrating that CD8+ T cells can acquire NK cell like features[424,425], down sampling and clustering of CD8+ T cells from PBMC samples which were stained with markers from panel two was also performed.

Based on positive CD3 and CD8 expression, 85 of 98 PBMC samples which were stained with markers from panel two had at least 22,000 live CD8+ T cells. For each of these samples, 22,000 live CD8+ T cells were randomly down sampled using FlowJo™ Software, for a total of 1,870,000 cells, and taken forward for semi-supervised clustering into cell subsets.

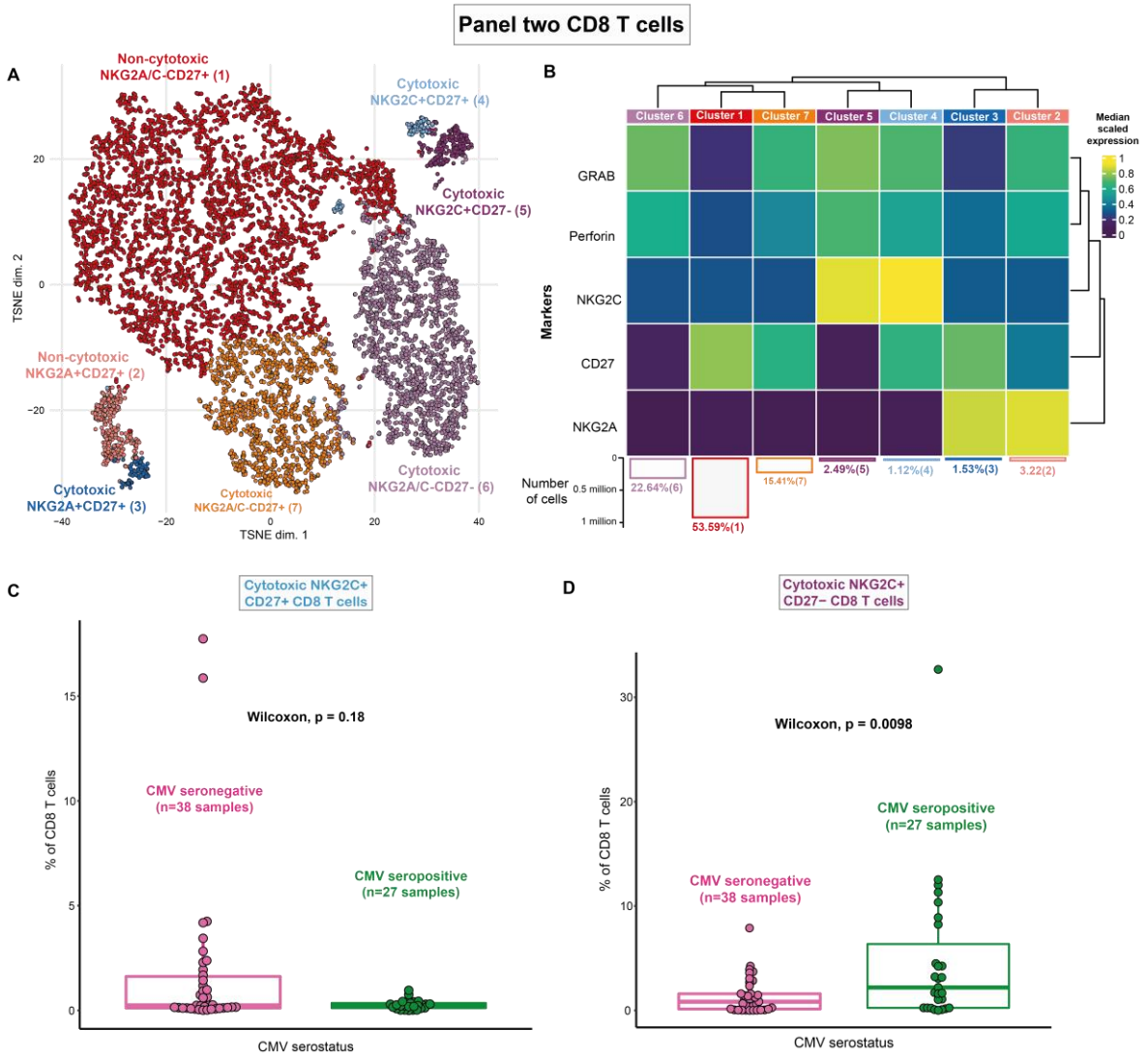


Figure 4.14. Clustering of CD8+ T cells with the markers in panel two identified seven clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel two, by CD8+ T cells identified seven major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the five markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the five markers used for clustering, for each of the seven clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage. (**C-D**) Boxplots of abundance of (**C**) Cytotoxic NKG2C+CD27- CD8 T cells and (**D**) Cytotoxic CD27+ memory CD8 T cells, expressed as a percentage of total CD8 T cells, between CMV seronegative and CMV seropositive samples. Points are coloured by CMV serostatus. Statistical comparisons of medians were carried out using a Wilcoxon test.

The 'type' markers used to cluster the CD8 T cells from the samples which were stained with panel two were the T cell lineage marker CD27; the NK cell receptors NKG2A and NKG2C which bind to HLA-E and are expressed by some CD8 T cells[424–426]; and perforin, and GRAB as markers of cytotoxic cells. Subsequent dimensionality reduction and clustering identified seven major clusters of cells (**Figure 4.14A-B**). The majority of cells composed cluster 1 which appeared to be a cluster of non-cytotoxic

NKG2A/C-CD27+ cells due to high expression of CD27 in combination with lack of expression of GRAB, perforin, NKG2A and NKG2C. Cluster 2 appeared to be a cluster of non-cytotoxic NKG2A+CD27+ cells due to expression of CD27 and NKG2A in combination with lack of expression of GRAB and perforin[427]. Cluster 3 appeared to be a cluster of cytotoxic NKG2A+CD27- cells due to expression of GRAB, perforin and NKG2A in combination with lack of expression of CD27[427]. Cluster 4 appeared to be a cluster of cytotoxic NKG2C+CD27+ cells based on expression of CD27, NKG2C, GRAB and perforin[428]. Cluster 5 was identified as cytotoxic NKG2C+CD27- cells due to the similar expression profile as cluster 4, except with lack of CD27 expression[428]. Cluster 6 appeared to be a cluster of cytotoxic NKG2A/C-CD27- cells due to high expression of GRAB and perforin in combination with lack of expression of CD27, NKG2A and NKG2C. Finally, cluster 7 appeared to be a cluster of cytotoxic NKG2A/C-CD27+ cells due to the similar expression profile as cluster 1, except with high expression of perforin and GRAB expression.

Given that NKG2C+ CD8 T cells are thought to expand following CMV infection[429,430], it is possible that the samples predominantly contributing to the two NKG2C+ clusters of CD8 T cells may be from individuals who are CMV seropositive. To evaluate this, the abundance of the NKG2C expressing clusters in the 65 samples for which CMV serostatus data was available were plotted with samples stratified by CMV serostatus (**Figure 4.14C-D**). This analysis showed that cytotoxic NKG2C+CD27- CD8 T cells were significantly more abundant in CMV seropositive samples, but there was no significant difference in abundance of NKG2C+CD27+ CD8 T cells in samples based on CMV serostatus.

4.3.5. 11 subsets of NK cells identified by clustering with the markers from panel 2 (NK cell centric)

For NK cells, there were only sufficient markers to identify cell subsets in the NK centric panel (panel two) and so only one set of dimensionality reduction and clustering was performed.

Based on CD3, CD4/CD14, CD19 and CD7 expression, 77 of 98 PBMC samples which were stained with markers from panel two had at least 20,000 live NK cells. For each of these samples, 20,000 live NK

cells were randomly down sampled using FlowJo™ Software, for a total of 1,540,000 cells, and taken forward for semi-supervised clustering into cell subsets.

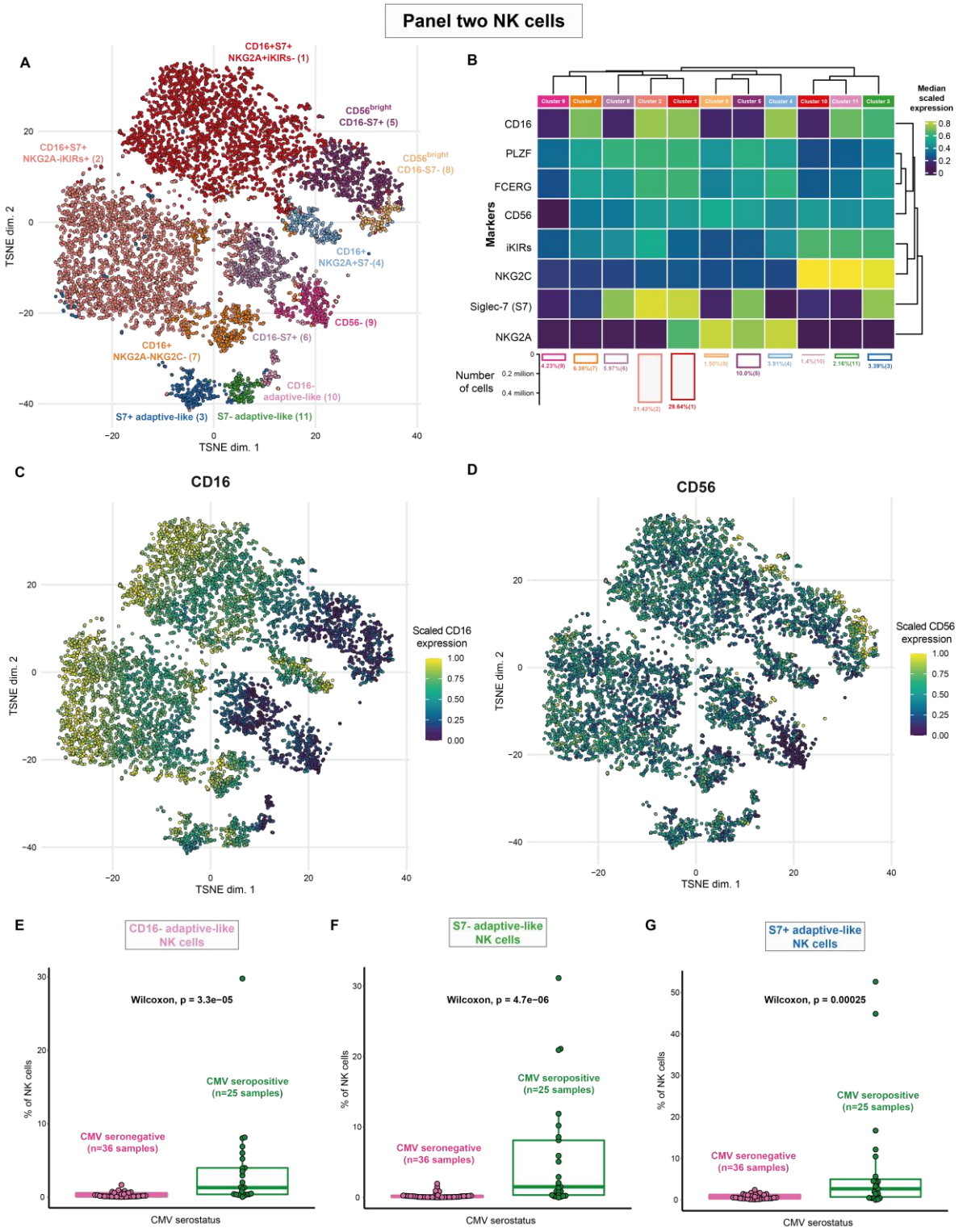


Figure 4.15. Clustering of natural killer cells with the markers in panel two identified 11 clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel two, by NK cells identified 11 major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the

eight markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the eight markers used for clustering, for each of the 11 clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage. (**C-D**). t-SNE visualisations showing scaled expression data per cell for (**C**) CD16 or (**D**) CD56. (**E-G**) Boxplots of abundance of (**E**) CD16- adaptive-like NK cells, (**F**) S7- adaptive-like NK cells, and (**G**) S7+ adaptive-like NK cells, expressed as a percentage of total NK cells, between CMV seronegative and CMV seropositive samples. Points are coloured by CMV serostatus. Statistical comparisons of medians were carried out using a Wilcoxon test.

The 'type' markers used to cluster the NK cells from the samples which were stained with panel two were the NK cell lineage markers CD16, CD56 and PLZF[431]; the NK cell signalling receptor FcεRIγ (FCERG)[432]; and the NK cell receptors NKG2A, NKG2C, Siglec-7 (S7), and inhibitory killer like immunoglobulin receptors (iKIRs)[433]. Subsequent dimensionality reduction and clustering identified twelve major clusters of cells (**Figure 4.15A-B**). Given the classical division of circulating NK cells by CD16 and CD56 expression[434–436], plots evaluating expression of these markers were used to help identify which subsets were present in each cluster (**Figure 4.15C-D**). NK cells did not seem to cluster based on CD56 expression, with most clusters composed of cells with varying CD56 expression. Although, cluster 9 looked to be almost entirely composed of cells lacking CD56 expression, and the cells expressing high levels of CD56 were predominantly concentrated in clusters 5 and 8.

Most NK cells were either in cluster 1 or cluster 2, which were similar in that they both showed high expression of CD16, FCERG, PLZF, and S7. However, cluster 1 showed expression of NKG2A and no iKIR expression whereas cluster 2 showed iKIR expression. This suggests that, although the cells in both of these clusters are likely mature NK cells, the cells in cluster 4 may be a more differentiated subset given their downregulation of NKG2A and upregulation of iKIRs[433,437,438]. Cells in cluster 3 were defined by their high expression of CD16, iKIRs, NKG2C in combination with low expression of PLZF and FCERG, which is a similar expression profile to the previously reported adaptive NK cells[434,439]. Clusters 10 and 11 also have a similar expression profile to cluster 3 but differ by expression of S7 and CD16. Cluster 4 appeared to be a cluster of CD16+NKG2A+S7- cells due to high expression of CD16 and NKG2A in combination with lack of expression of S7. As mentioned above, cluster 5 was $CD56^{bright}$CD16-S7+ and

cluster 8 was CD56^{bright}CD16-S7-. Cluster 5 and 8 both immature NK cells but cluster 5 likely to be slightly more mature due to S7 expression. Cluster 6 was CD16-S7+ due to high expression of S7 in combination with lack of expression of CD16. Cluster 7 was CD16+NKG2A-NKG2C- due to high expression of CD16 in combination with lack of expression of NKG2A and NKG2C. As mentioned above, cluster 9 was defined by its lack of CD56 expression and so is likely composed of the CD56- NK cells which are known to exist as a small population in peripheral blood[436].

Like CD8 T cells, NKG2C expressing NK cells are thought to expand following CMV infection[440–442]. Hence, it is possible that the samples predominantly contributing to the three adaptive-like clusters of NK cells, which are all NKG2C+, may be from individuals who are CMV seropositive. To evaluate this, the abundance of these adaptive-like and NKG2C expressing NK cell clusters in the 61 samples for which CMV serostatus data was available were plotted with samples stratified by CMV serostatus (**Figure 4.15F-G**). This analysis showed that all three adaptive-like and NKG2C expressing NK cell clusters were significantly more abundant in CMV seropositive samples.

4.3.6. Three subsets of monocytes identified by clustering with the markers from panel 3 (B cell and myeloid centric)

For monocytes, there were only sufficient markers to identify cell subsets in the B cell and myeloid centric panel (panel three) and so only one set of dimensionality reduction and clustering was performed.

Based on CD14 expression, 73 of 98 PBMC samples which were stained with markers from panel three had at least 22,000 live monocytes. For each of these samples, 22,000 live monocytes were randomly down sampled using FlowJo™ Software, for a total of 1,606,000 cells, and taken forward for semi-supervised clustering into cell subsets.

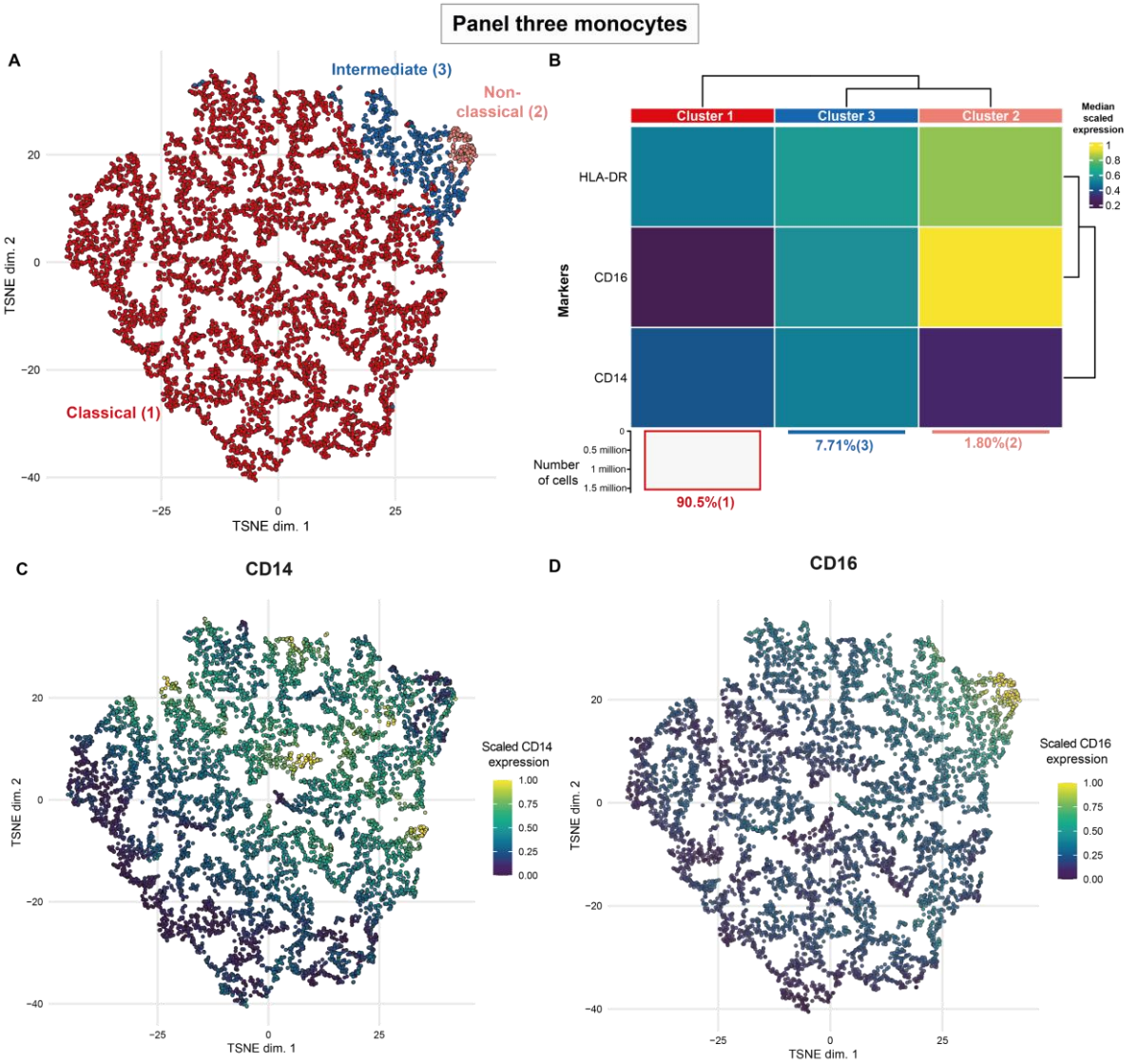


Figure 4.16. Clustering of CD14+ monocytes with the markers in panel three identified 3 clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel three, by CD14+ monocytes identified three major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the three markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the three markers used for clustering, for each of the three clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage. (**C-D**). t-SNE visualisations showing scaled expression data per cell for (**C**) CD14 or (**D**) CD16.

The 'type' markers used to cluster the monocytes from the samples which were stained with panel three were the monocyte lineage markers CD14 and CD16[443]; and the class II MHC molecule HLA-DR which is highly expressed by most monocytes but has been reported to be downregulated in certain subsets[444]. Subsequent dimensionality reduction and clustering identified three major clusters of cells (**Figure 4.16A-B**). Given the classical division of monocytes by CD14 and CD16 expression[443], plots evaluating expression of these markers were used to help identify which subsets were present in each

cluster (**Figure 4.16C-D**). Most cells composed cluster 1 which appeared to be a cluster of classical monocytes due to expression of CD14 in combination with lack of expression of CD16. Although, the level of CD14 expression by cells within this cluster was variable (**Figure 4.16C**). Cluster 2 appeared to be a cluster of non-classical monocytes due to expression of CD16 in combination with low expression of CD14. Finally, cluster 3 appeared to be a cluster of intermediate monocytes due to expression of both CD14 and CD16.

4.3.7. Four subsets of myeloid dendritic cells (mDCs) identified by clustering with the markers from panel 3 (B cell and myeloid centric)

For mDCs, there were only sufficient markers to identify cell subsets in the B cell and myeloid centric panel (panel three) and so only one set of dimensionality reduction and clustering was performed.

Based on expression of CD14, HLA-DR, CD11c and CD123, 81 of 98 PBMC samples which were stained with markers from panel three had at least 20,000 live mDCs. For each of these samples, 20,000 live mDCs were randomly down sampled using FlowJo™ Software, for a total of 1,620,000 cells, and taken forward for semi-supervised clustering into cell subsets.

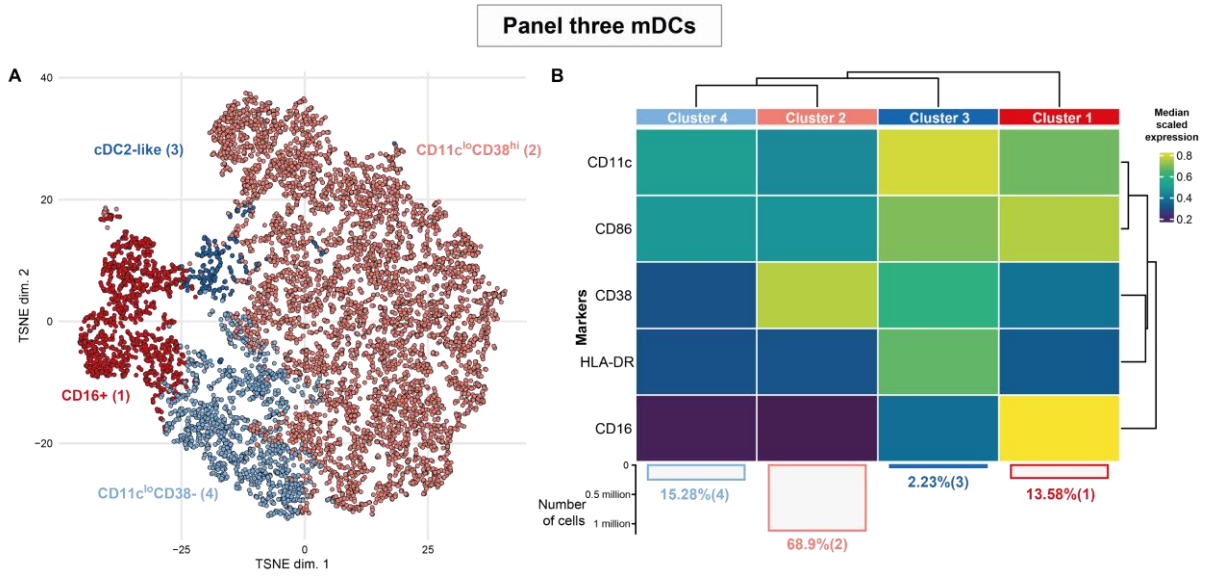


Figure 4.17. Clustering of CD14-HLA-DR+CD11c+CD123- mDCs with the markers in panel three identified four clusters of cells. Dimensionality reduction of normalised expression data, for a subset of markers from panel three, by monocytic dendritic cells (mDCs) identified four major clusters of cells. (**A**) t-SNE visualisation of the normalised expression data for each of the eight markers used for clustering. One point per cell with points coloured by the cluster they belong to. (**B**) Heatmap showing the median normalised expression for each of the five markers used for clustering, for each of the four clusters of cells identified. Hierarchical clustering of markers and cell clusters was calculated using the "average" linkage.

The 'type' markers used to cluster the mDCs from the samples which were stained with panel three were the myeloid lineage marker CD11c and CD16[445]; and CD38, CD86 and HLA-DR as markers of activation and potential to stimulate T cells[445,446]. The markers CD11c, CD16, CD38, CD86 and HLA-DR were used to cluster the CD123-CD11c+ mDCs from the PBMC samples stained with panel three markers. Dimensionality reduction and clustering identified four major clusters of cells (**Figure 4.17**). Cluster 1 appeared to be a cluster of CD16+ mDCs, similar to those previously described[445], due to high expression of CD16 and CD86 in combination with low expression of CD38 and HLA-DR Cluster 2, which was contained most cells, appeared to be a cluster of CD11c^{lo}CD38^{hi} cells due to expression of CD38 in combination with low expression of both HLA-DR and CD11c, and lack of expression of CD16. Cluster 3 appeared to be a cluster of DCs similar to the previously described conventional DC2 (cDC2) cells, except with high expression of CD16, due to high expression of HLA-DR, CD11c, CD38 and CD86. Finally, cluster 4 appeared to be a cluster of CD11c^{lo}CD38- cells due to a similar expression profile to cluster 2 except with a lack of CD38 expression. Given their low expression of CD11c, both cluster 2 and 4 could be composed of the cDC1 subset which have been previously described[445], although

expression of additional markers, such as C-type lectin-like receptor 9A (Clec9A) and cell adhesion molecule 1 (CADM1) would be needed to confidently label these cells as cDC1s[445]. Moreover, expression of CD38 has be proposed as a marker for mature monocyte derived DCs[447] and so the cells in clusters 2 and 4 could represent two different maturation states of the same mDC subset.

4.4 Associations of abundance and expression profiles of cell subsets with clinical factors

4.4.1. Rationale

Following identification of cell subsets in the previous section, the next step was to comprehensively explore whether differential abundance and/or expression patterns of these subsets associated with clinical factors of the patients, such as CB. To this end, differential analysis of abundances and states of the identified cell subsets from section 4.3 was performed using the R package diffcyt[336]. In total there were up to 2109 features per sample, with 70 of these features cell subset abundances, 389 features were expression of a marker by the bulk cell type (e.g. CD86 expression by all 20,000 B cells down sampled), and the remaining 1650 features were expression of markers by the cell subsets identified by clustering (e.g. CD86 expression by IgG+ memory B cells) (**Figure 4.18**).

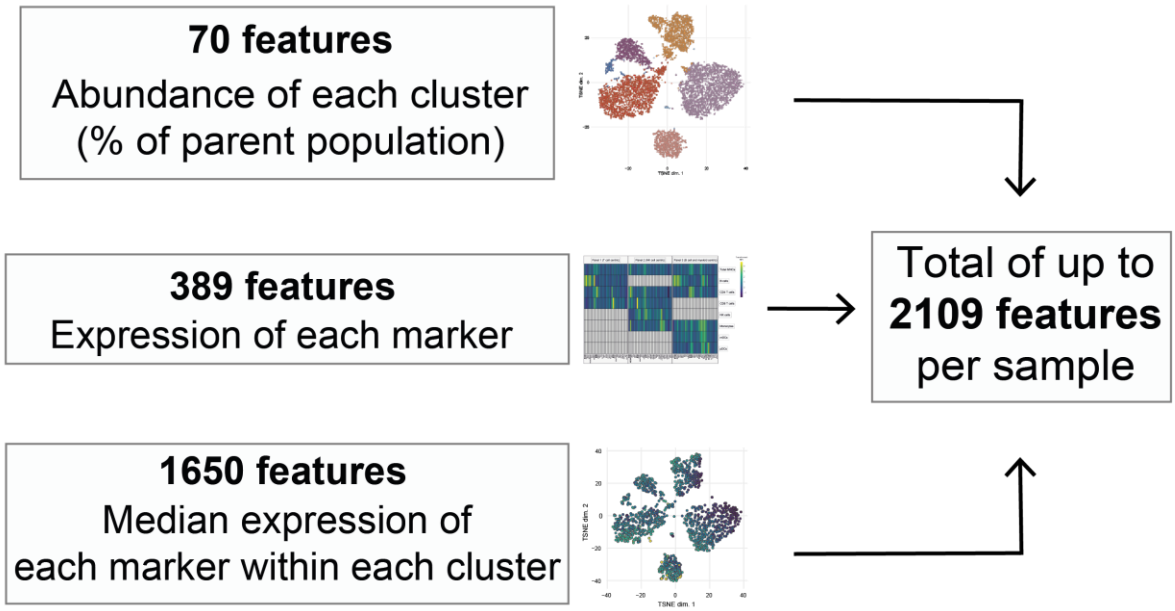


Figure 4.18. Summary of features calculated for each sample. Following gating, down sampling, and clustering into cell subsets, across the three panels, a total of up to 2109 features were calculated for each sample. 70 of these features were abundance data for the cell subsets identified by the described clusterings. 389 of these features were median expression data for each of the state markers for each major cell type. 1650 of these features were median expression data for each of the state markers within the cell subsets identified by the described clusterings.

4.4.2. Clinical benefit of patients from the LUD2015-005 inoperable cohort

As previously stated, the primary aim of this chapter is to determine whether peripheral signatures of increased immune activation are associated with improved clinical outcome of OC patients treated with ICI+CTX. Subsequently, to identify signatures of that could predict clinical outcome to ICI+CTX treatment in OAC, the abundance of cell subsets and cell state data of samples from OAC patients from the LUD2015-005 study, from both the PreTx and ICI-4W timepoints, were analysed for associations with patient clinical outcomes.

4.4.2.1. Differences in CD8 T cells and B cells are detectable pre-treatment

As it would be most beneficial to be able to identify patients who are likely to benefit from ICI treatment before they receive treatment, the PreTx samples were compared first.

Differential analysis of the abundances of the 70 cell subsets showed that two B cell subsets, VH4-34+IgG+DN2-like B cells and DN2/activated naïve-like B cells, and the cytotoxic NKG2C+CD27- subset of CD8 T cells were significantly increased in CB patients, whilst the cytotoxic CD27+ memory CD8 T cell subset was significantly enriched in NCB patients at the PreTx timepoint (**Figure 4.19**). However, performing hierarchical clustering on the abundance data for these four cell subsets did not cluster the PreTx OAC samples by CB status (**Figure 4.19B**).

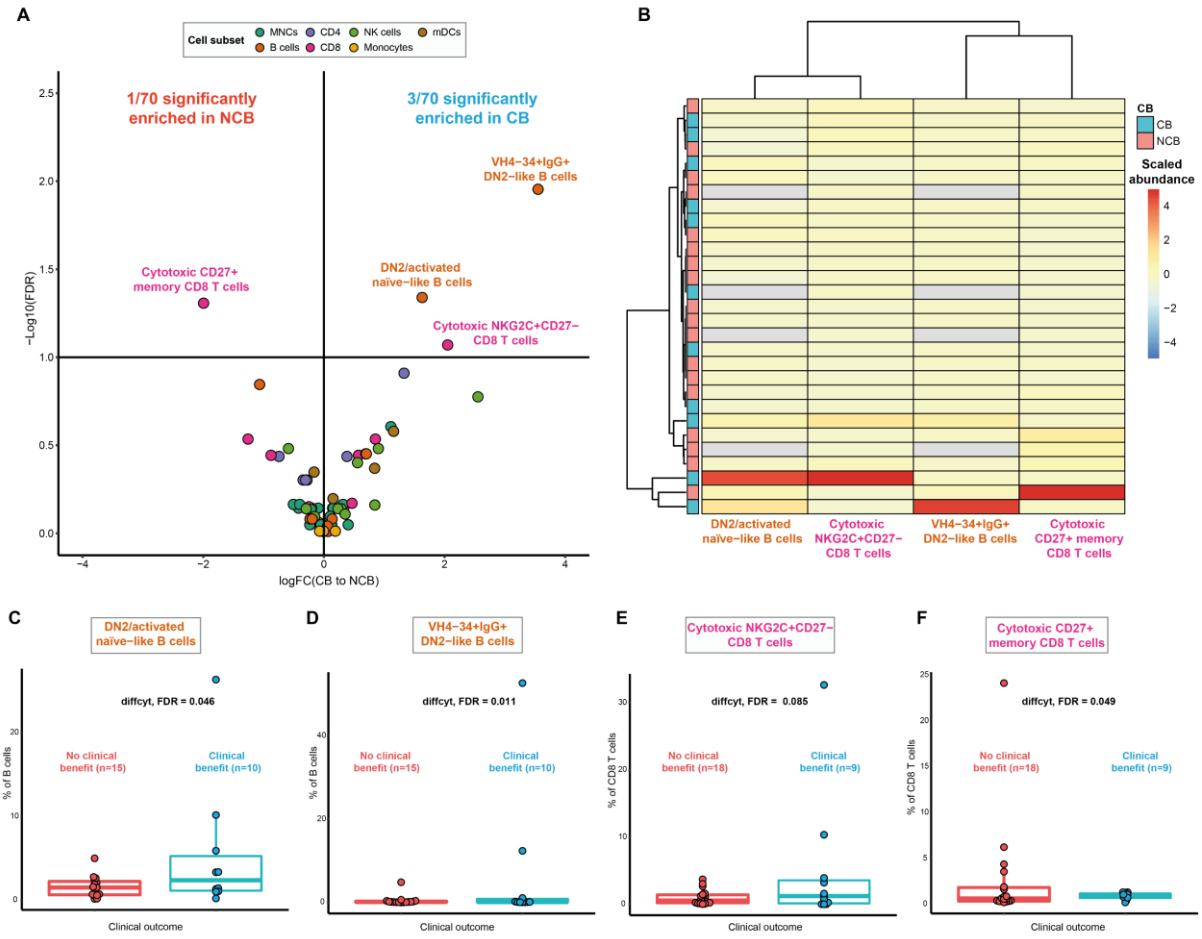


Figure 4.19. Subsets of B cells and CD8+ T cells showed significantly different abundances between PreTx samples from CB and NCB OAC patients. (**A**) Volcano-like plot showing the log fold change between CB and NCB OAC patients (logFC(CB to NCB)) versus the -Log10(FDR) for the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for the four cell subsets with an FDR<0.1, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from no clinical benefit (NCB) (red) or CB (blue) patients. Grey cells indicate where abundance data for that cell subset was not available for that sample. (**C-F**) Boxplots of cell subset abundance as a percentage of the parent major cell type between CB and NCB OAC patients for (**C**) DN2/activated naïve-like B cells, (**D**) VH4-34+IgG+ DN2-like B cells, (**E**) Cytotoxic NKG2C+CD27- CD8 T cells, and (**F**) Cytotoxic CD27+

memory CD8 T cells. Points are coloured by clinical benefit status. p values, FDRs, and logFC values were calculated using the diffcyt-DA-edgeR method from the diffcyt package.

Although subsets with a similar expression profile have been reported[400,448,449], given that the VH4-34+IgG+DN2-like B cell subset appeared to predominantly originate from a small number of patients (**Figure 4.19D**), the fcs files for down sampled B cells were inspected in FlowJo™ Software to gate the cells with the same expression profile. (**Figure 4.20**).

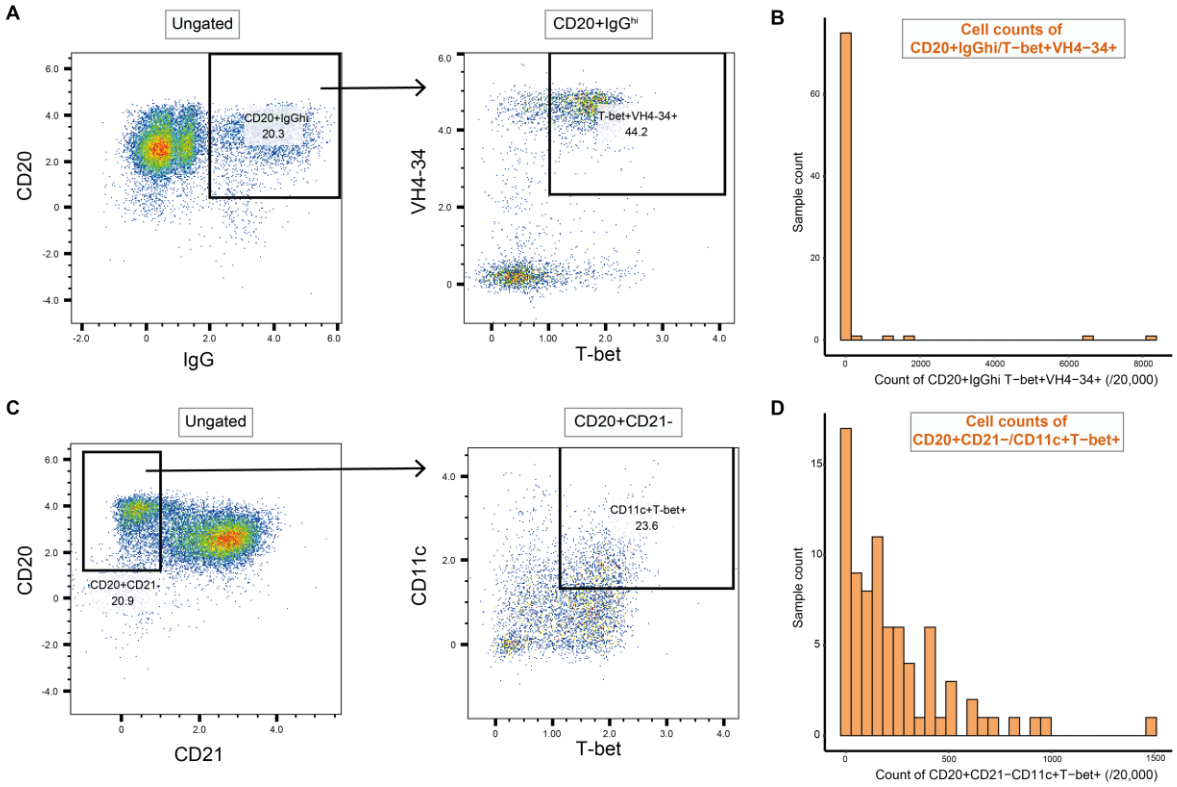


Figure 4.20. VH4-34+IgG+ DN2-like B cells are found in only a small number of samples whilst DN2/activated naïve-like B cells are more commonly present in samples. (**A**) Example FlowJo™ Software gating strategy for B cells with a similar expression profile to VH4-34+IgG+ DN2-like B cells, as CD20+IgG^{hi}T-bet+VH4-34+, in an individual sample. (**B**) Histogram of counts of CD20+IgG^{hi}T-bet+VH4-34+ cells, out of a total of 20,00 B cells, for each of the 78 samples which were used for B cell down sampling and clustering. (**C**) Example FlowJo™ Software gating strategy for B cells with a similar expression profile to DN2/activated naïve-like B cells, as CD20+CD21-CD11c+T-bet+, in the same individual sample as shown in (A). (**D**) Same as (B) except counts of CD20+CD21-CD11c+T-bet+ cells.

VH4-34+IgG+ DN2-like B cells could be identified in FlowJo™ Software as CD20+IgG^{hi}T-bet+VH4-34+ B cells but seemed to predominantly originate from a small number of samples (**Figure 4.20A-B**). Similarly, DN2/activated naïve-like B cells could be identified in FlowJo™ Software as CD20+CD21-

CD11c+T-bet+ B cells (**Figure 4.20C**). However, unlike VH4-34+IgG+ DN2-like B cells, they seemed to be present at reasonable cell numbers in most samples (**Figure 4.20D**).

To determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between OAC PreTx samples from CB and NCB patients. In total, based on an FDR cut off of 0.1, 50 features were found to be significantly different between these samples, with most of these features relating to CD8 T cells (**Figure 4.21A-B**). Subsequent hierarchical clustering on the data for these 50 significant features did not cluster the PreTx OAC samples by CB status (**Figure 4.21C**). However, similar clustering on the data for only the 39 significant features that were from CD8 T cells did largely cluster the PreTx OAC samples by CB status (**Figure 4.21D**).

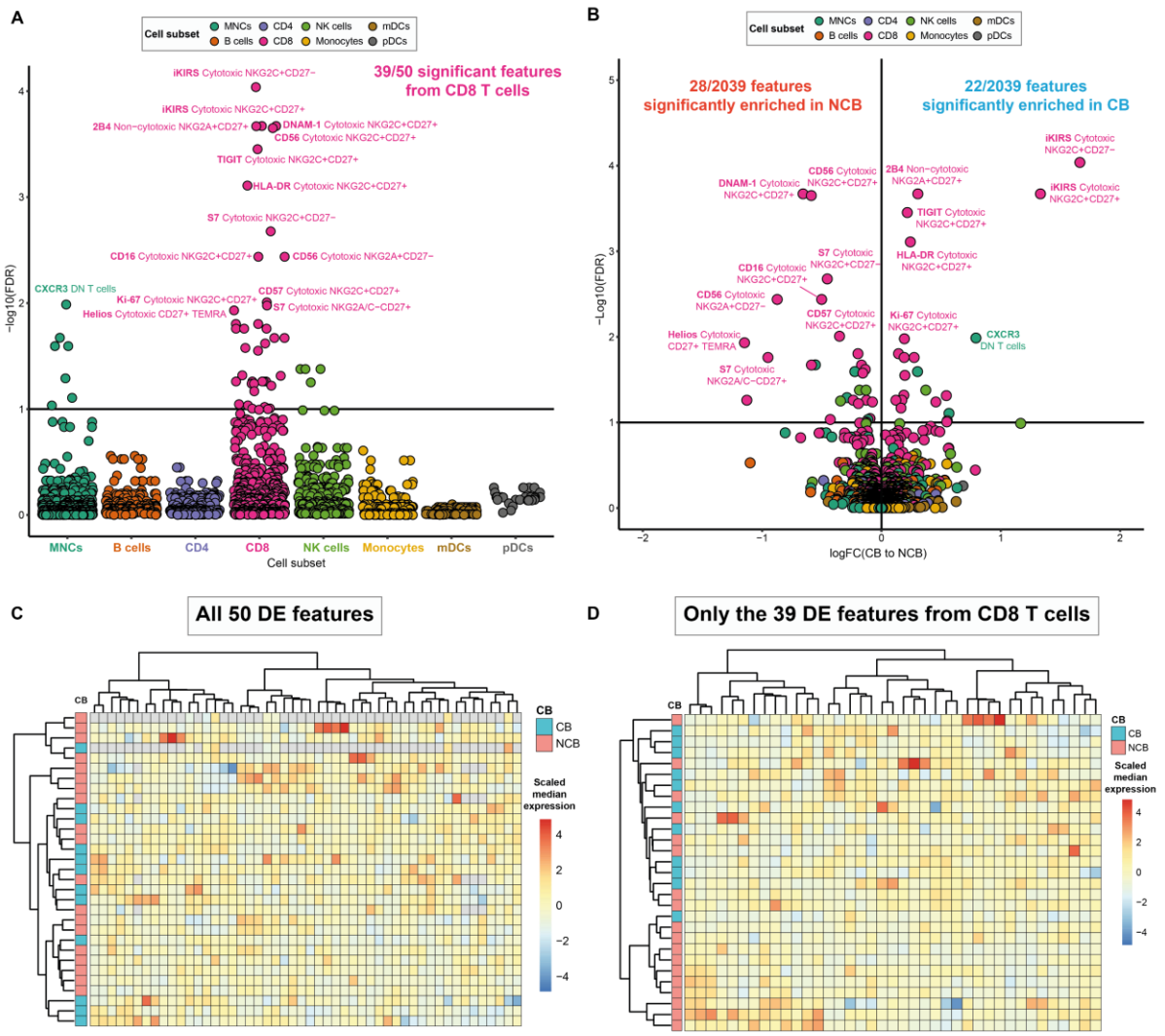


Figure 4.21. CD8+ T cells showed significantly different cell state features between PreTx samples from CB and NCB OAC patients. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between CB and NCB OAC patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between CB and NCB OAC patients (logFC(CB to NCB)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond to. (**C-D**) Heatmaps of the scaled cell state feature data for (**C**) all 50 features with an FDR<0.1 or (**D**) the 39 features from CD8 T cells. Data is scaled by column (feature), with unsupervised hierarchical clustering of both features (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from no clinical benefit (NCB) (red) or CB (blue) patients. Grey cells indicate where data for that feature was not available for that sample. For (D), two samples were excluded due to missing data for all of the 39 features. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

As iKIR expression by cytotoxic NKG2C+CD27+ and NKG2C+CD27- CD8 T cells (clusters 4 and 5 from section 4.3.4.2) were two of the features which were most significantly increased in CB patients, iKIR expression by CD8 T cells on an individual cell basis was analysed. Out of all of the CD8 T cells, the two

cytotoxic CD27- subsets seemed to show the highest expression of iKIRs, with cells within these subsets from CB patients showing more consistently high expression of iKIRs than cells of the same subsets from NCB patients (**Figure 4.22**).

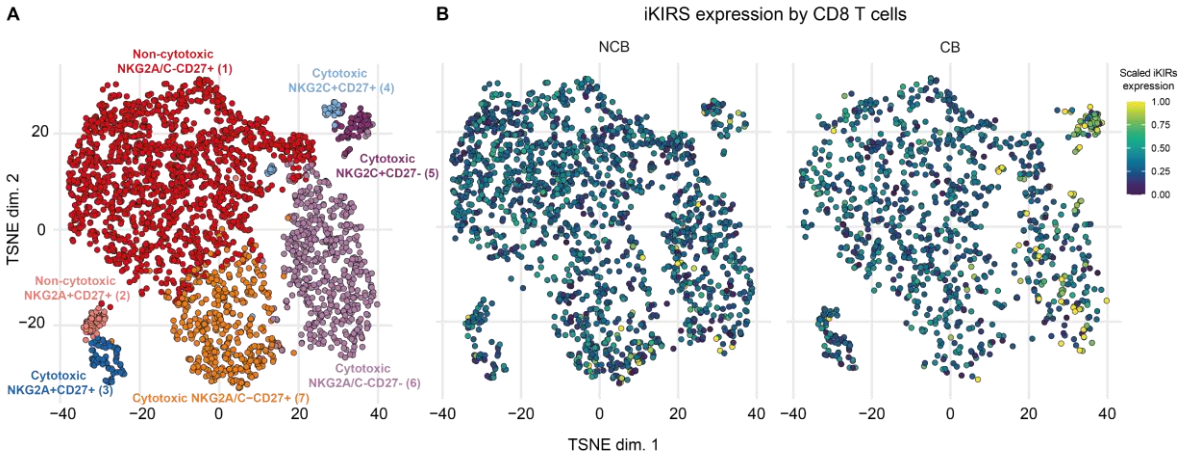


Figure 4.22. Expression of iKIRs by CD8 T cells is greatest in cytotoxic CD27- cells and expression of iKIRs in these cells is increased in PreTx samples from CB patients compared to NCB patients. (**A**) t-SNE plots of CD8 T cell subset clusters from the cells stained with panel two markers. One point per cell. Points coloured by the cell subset cluster they correspond to. (**B**) t-SNE plots of CD8 T cell subset clusters from the cells stained with panel two markers. One point per cell. Points coloured by expression of inhibitory killer Ig-like receptors (iKIRs). Left plot are cells from PreTx samples of CB OAC patients. Right plot are cells from PreTx samples of NCB OAC patients.

4.4.2.2. Increased abundance of NK cells after four weeks of ICI is associated with clinical benefit

Although it is of most interest to identify PreTx signatures that associate with good clinical outcome, several studies have shown that peripheral immune signatures that are associated with good clinical outcome can be detected during treatment[127,128,130,132]. Moreover, it would be clinically useful to monitor CB likelihood early on during treatment. Hence, next, differential analysis of the ICI-4W samples was performed.

At this timepoint, differential analysis of the abundances of the 70 cell subsets did not show the same enrichment of in CB patients that were observed in the PreTx samples (**Figure 4.19**). Instead it was found that two MNC cell subsets, NK cells and DN T cells, three subsets of adaptive-like NK cells, and the subset of VH4-34+IgG+DN2-like B cells were significantly increased in CB patients, whilst the IgM+

memory/MZ-like B cell subset and the CD16+S7+NKG2A+iKIRs- NK cell subset were significantly enriched in NCB patients at the ICI-4W timepoint (**Figure 4.23A**). Moreover, performing hierarchical clustering on the abundance data for total NK cells, and the three other NK cell subsets, that were found to be enriched in the CB patients largely clustered the ICI-4W OAC samples by CB status (**Figure 4.23B**). As all of the ICI-4W samples had abundance data for total NK cells (from clustering of MNCs), this data was more closely analysed for association with patient clinical outcome. CB patients had a significantly increased abundance of NK cells (as a percentage of MNCs) on average than NCB patients (**Figure 4.23C**). However, patients who had a high abundance of NK cells (>median) did not show significantly improved OS compared to patients who had a low abundance of NK cells (≤median) (**Figure 4.23D**).

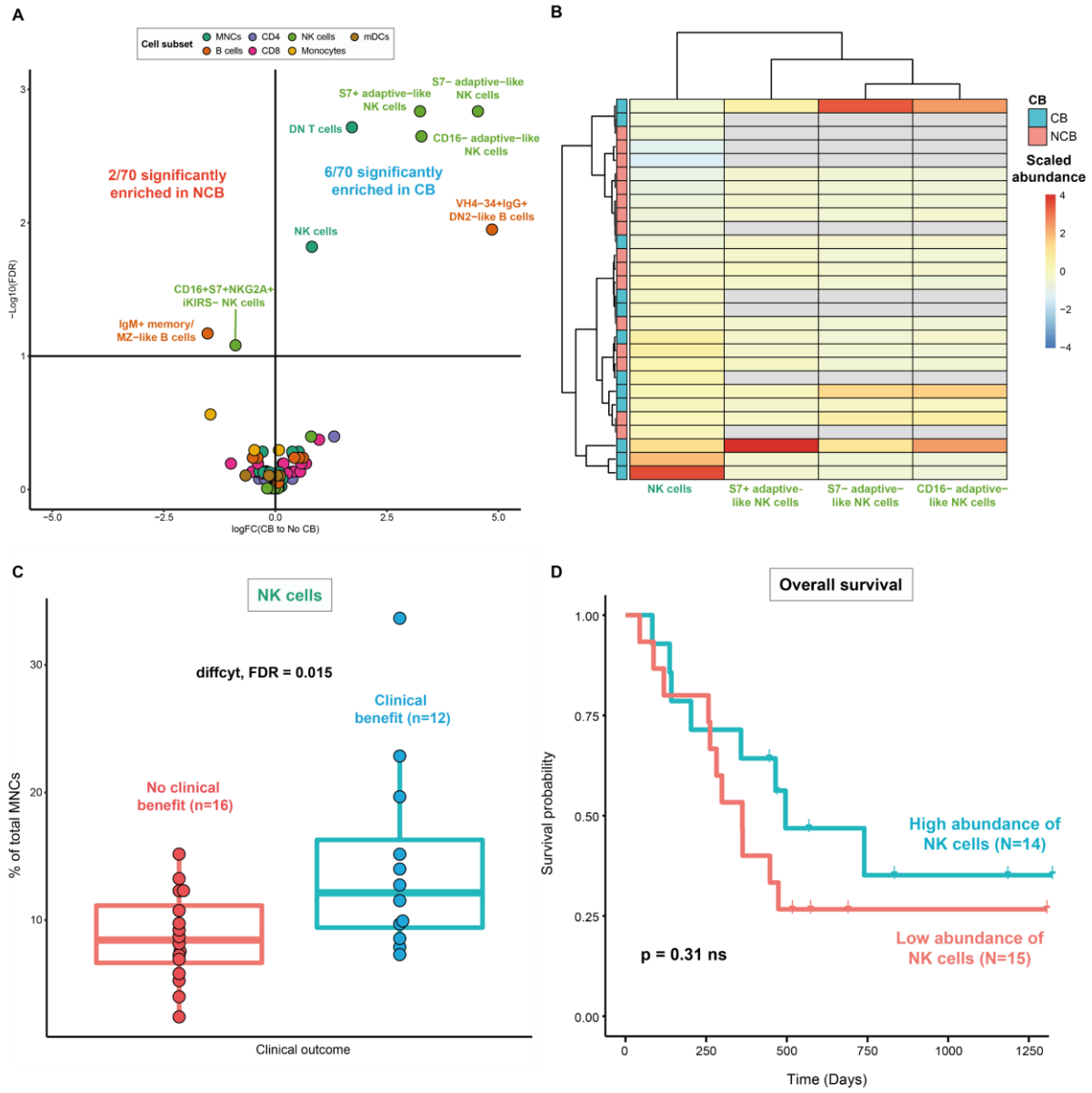


Figure 4.23. NK cells, and adaptive-like NK cell subsets, showed significantly different abundances between ICI-4W samples from CB and NCB OAC patients. (**A**) Volcano-like plot showing the log fold change between CB and NCB OAC patients (logFC(CB to NCB)) versus the -Log10(FDR) for the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for total NK cells and adaptive-like NK cell subsets, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from no clinical benefit (NCB) (red) or CB (blue) patients. Grey cells indicate where abundance data for that cell subset was not available for that sample. (**C**) Boxplots of NK cell abundance as a percentage of the total mononuclear cells (MNCs) between CB (blue) and NCB (red) OAC patients. p values, FDRs, and logFC values were calculated using the diffcyt-DA-edgeR method from the diffcyt package. (**D**) Kaplan Meier plot of overall survival with patients split into those with a high abundance of NK cells at the ICI-4W timepoint (> median) (blue) and those who had a low number of NK cells (≤ median) at the ICI-4W timepoint (red). p value was calculated using the log rank method.

As with the PreTx samples, to determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between OAC ICI-4W samples from CB and NCB patients. In total, based on an FDR cut off of 0.1, 67 features were found to be significantly different between these samples, with, once again, most of these features relating to CD8 T cells and increased iKIR expression by cytotoxic NKG2C+ CD8 T cells were the features most significantly associated with CB (**Figure 4.24**). In addition to the large number of features from CD8 T cells, expression of the pro-apoptotic receptor FAS by DN2/activated naïve B cells was a notable feature significantly enriched in NCB patients. FAS expression is upregulated upon B cell activation but can lead to apoptosis of the expressing cell if ligated by its ligand without concurrent survival signals[450].

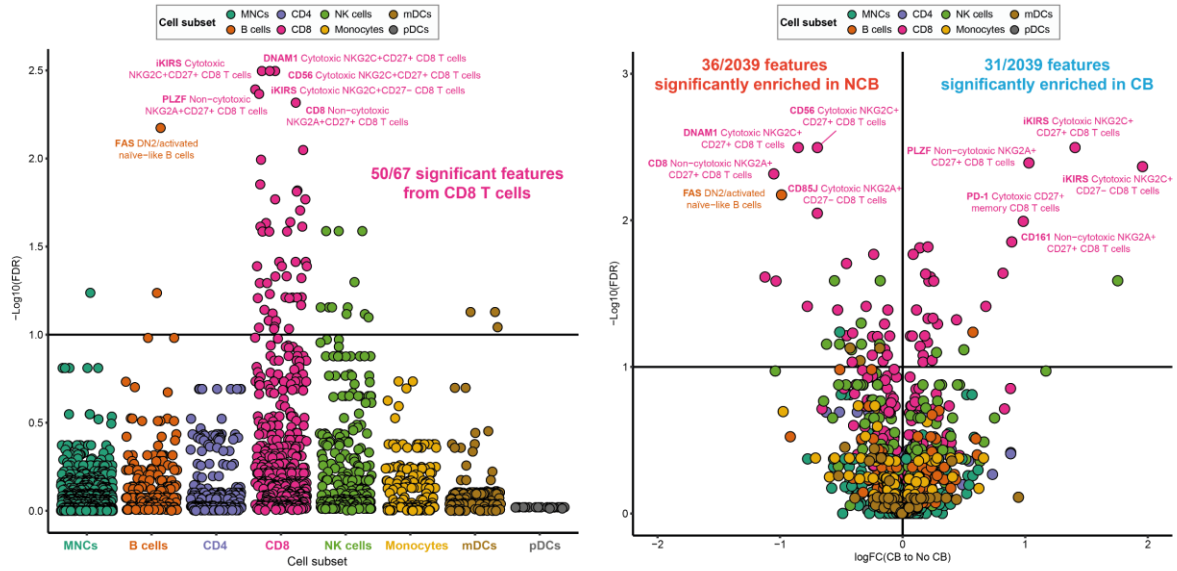


Figure 4.24. CD8+ T cells showed significantly different cell state features between ICI-4W samples from CB and NCB OAC patients. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between CB and NCB OAC patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between CB and NCB OAC patients (logFC(CB to NCB)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond to. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

4.4.3. Development of irAEs in patients from the LUD2015-005 inoperable cohort

Although irAE development has been associated with improved clinical outcome to ICI, there are still numerous patients who either develop irAEs during treatment but do not benefit clinically, or do not develop irAEs but show good clinical outcome[141–145]. This suggests that, although there may be overlap between signatures that predispose an individual to high risk of irAE development and good clinical outcome, there are likely to be signatures which predispose an individual for high risk of irAE development even though they are unlikely to show a good outcome to ICI treatment. Identifying these potential signatures could be highly valuable for understanding how to identify patients at high risk of irAE development and altering their treatment if necessary, especially if they are unlikely to experience clinical benefit.

Subsequently, to identify signatures of high risk of irAE development, the abundance of cell subsets and cell state data of samples from OAC patients who developed an irAE during treatment and OAC patients who did not develop an irAE during treatment were compared.

4.4.3.1. At the PreTx timepoint, irAE developers show signatures of reduced LAP expression and a skew towards a more immature and activated B cell compartment

As it would be most beneficial to be able to identify patients at high risk of irAE development prior to treatment with ICI, the PreTx samples were compared first.

Differential analysis of the abundances of the 70 cell subsets showed that a CD45RA+ non-T non-B cell subset of MNCs was significantly increased in irAE patients, whilst the similar CD45RA- non-T non-B cell subset was significantly enriched in no irAE patients, along with the IgM+ memory/MZ-like B cells, at the PreTx timepoint (**Figure 4.25**). However, performing hierarchical clustering on the abundance data for these three cell subsets did not cluster the PreTx OAC samples by irAE status (**Figure 4.25B**).

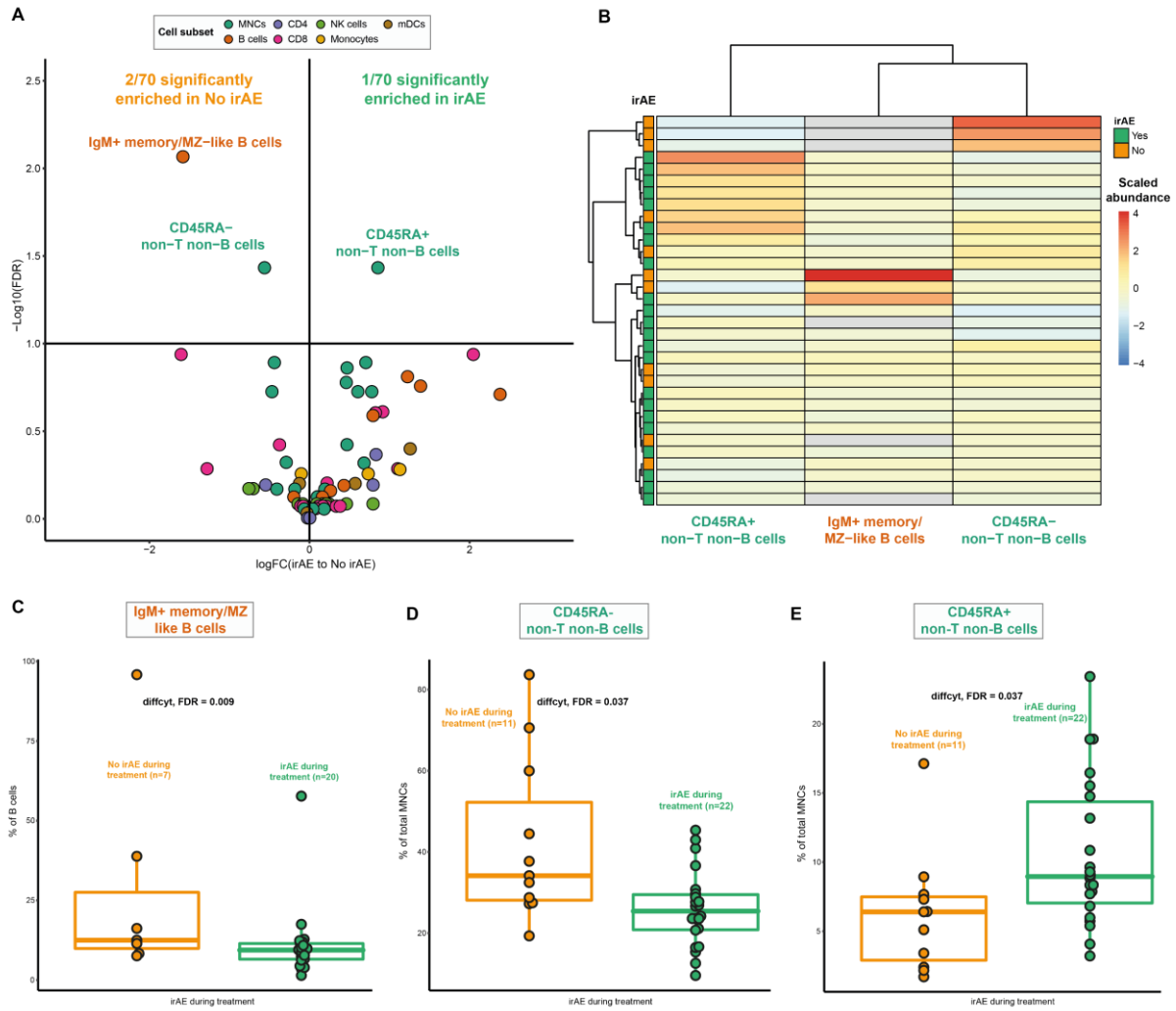


Figure 4.25. Subsets of mononuclear cells and B cells showed significantly different abundances between PreTx samples from OAC patients who do and do not develop irAEs during treatment. (**A**) Volcano-like plot showing the log fold change between irAE and no irAE OAC patients (logFC(irAE to No irAE)) versus the -Log10(FDR) for the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for the three cell subsets with an FDR<0.1, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from irAE (green) or no irAE (yellow) patients. Grey cells indicate where abundance data for that cell subset was not available for that sample. (**C-E**) Boxplots of cell subset abundance as a percentage of the parent major cell type between CB and NCB OAC patients for (**C**) IgM+ memory/MZ-like B cells, (**D**) CD45RA- non-T non-B cells and, (**E**) CD45RA+ non-T non-B cells. Points are coloured by irAE status. p values, FDRs, and logFC values were calculated using the diffcyt-DA-edgeR method from the diffcyt package.

To determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between OAC PreTx samples from patients who developed an irAE during treatment and patients who did not develop an irAE during treatment. In total, based on an FDR cut off of 0.1, 114 features were found to be significantly different between these samples, with significant features predominantly related to MNCs, B cells, and CD8 T cells (**Figure 4.26A-B**).

Subsequent hierarchical clustering on the data for these 50 significant features largely clustered the PreTx OAC samples by irAE status, with clustering of samples seeming to be predominantly driven by the left most cluster of features which were predominantly related to B cells (**Figure 4.26C**). Indeed, similar clustering on the data for only the 43 significant features that were from B cells also largely clustered the PreTx OAC samples by irAE status (**Figure 4.26D**).

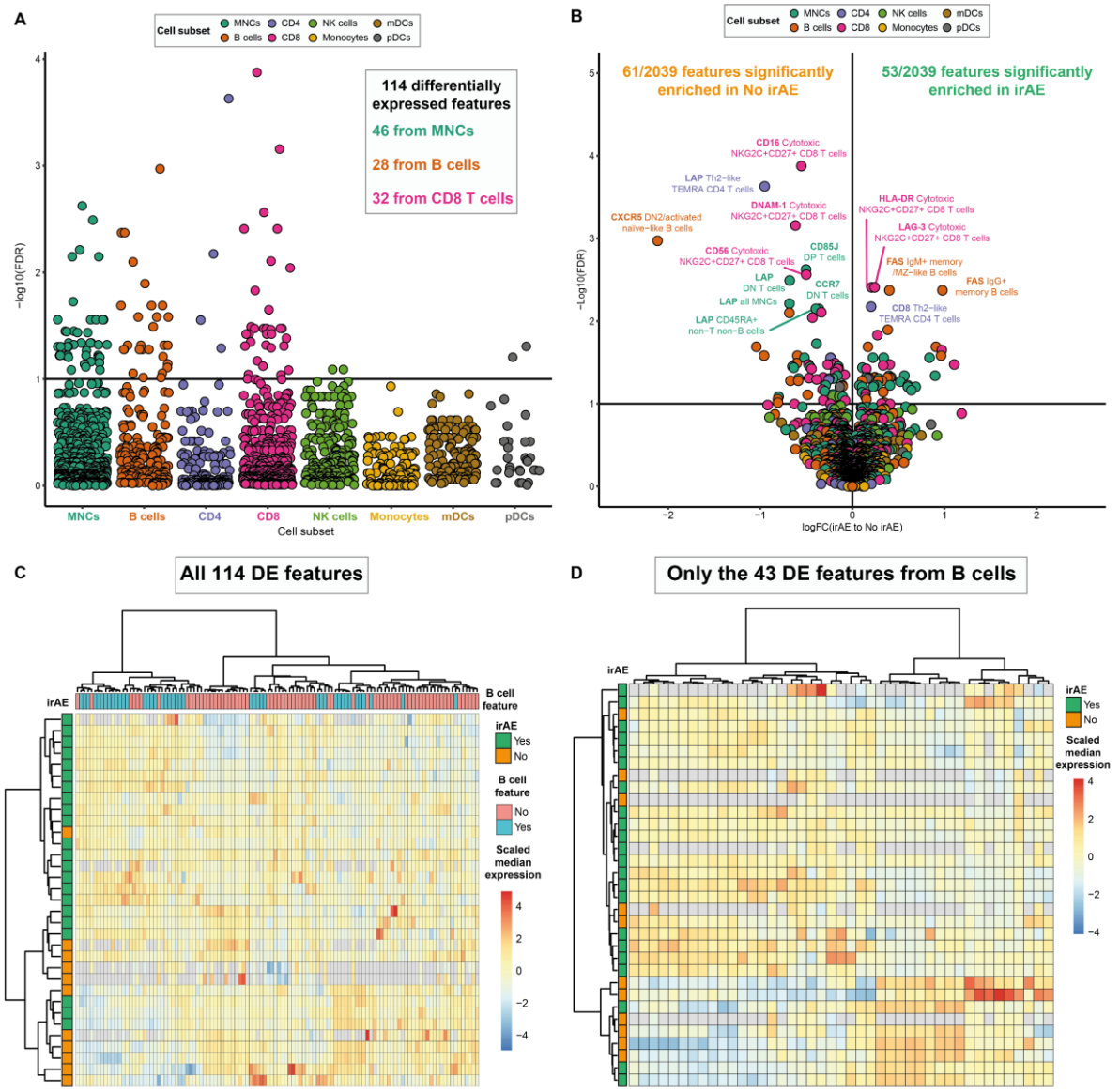


Figure 4.26. Mononuclear cells, B cells, and CD8+ T cells showed significantly different cell state features between PreTx samples from OAC patients who do and do not develop irAEs during treatment. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between between irAE and no irAE OAC patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between irAE and no irAE OAC patients (logFC(irAE to No irAE)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond

to. (**C-D**) Heatmaps of the scaled cell state feature data for (**C**) all 114 features with an FDR<0.1 or (**D**) the 43 features from B cells. Data is scaled by column (feature), with unsupervised hierarchical clustering of both features (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from irAE (green) or no irAE (yellow) patients. Grey cells indicate where data for that feature was not available for that sample. For (C) features (columns) are annotated with a key indicating whether the feature was related to B cells (blue) or not (red). p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

Given the clustering of samples by irAE status using just the B cell related features, the markers and subsets of B cells that these significant features were corresponding to were further analysed. The most prominent signal from this analysis was an increase in activation markers, such as FAS, CD38, Ki67, and CD86 in the total B cells from patients who developed an irAE during treatment (**Figure 4.27**). In contrast, the total B cells from patients that did not develop an irAE during treatment showed an increase in expression of the chemokine receptor CXCR5, which directs B cells to migrate to SLOs and is highly expressed on mature B cells[395], and CD24 and CD27 which are markers associated with transitional and memory B cells.

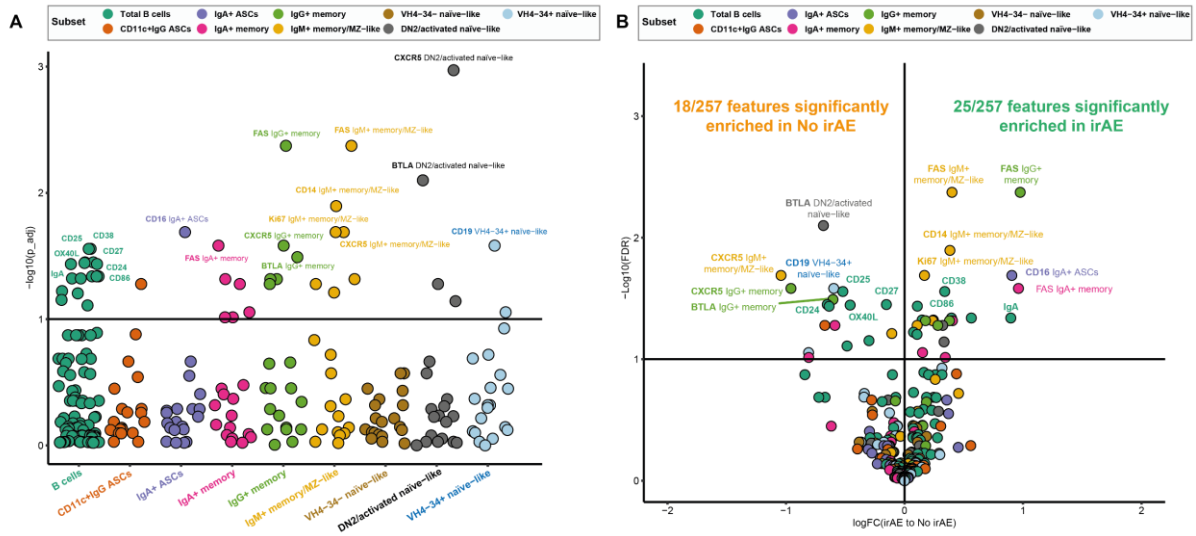


Figure 4.27. B cells from patients who developed an irAE during treatment showed increased expression of features associated with activation. (**A**) Dotplot showing -Log10(FDR) for the 257 B cell related cell state features identified by cell type down sampling and clusterings, when comparing between between irAE and no irAE OAC patients, versus the B cell subset that the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between irAE and no irAE OAC patients (logFC(irAE to No irAE)) versus the -Log10(FDR) for the same 257 B cell related cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the B cell subset that they correspond to. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

Aside from the large number of B cell features, a number of differentially expressed features were related to expression of latency associated peptide (LAP), which is non-covalently bound to the cytokine transforming growth factor-beta (TGF-β), were most significantly increased in no irAE patients. Therefore, LAP expression by MNCs was further analysed. Patients who developed an irAE during treatment showed a significantly lower expression of LAP by their MNCs than MNCs from patients who did not develop an irAE during treatment (**Figure 4.28A**). Analysis of LAP expression on an individual cell basis, showed that out of all of the MNCs, the CD45RA- non-T non-B cell subset showed the highest expression of LAP, although certain cell of other cell types including CD4+CD8- T cells showed high LAP expression (**Figure 4.28B**).

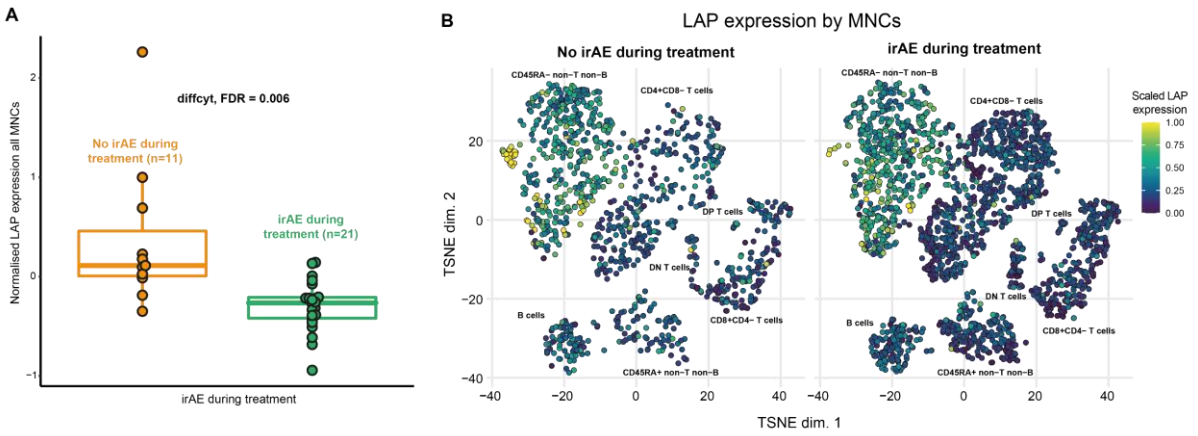


Figure 4.28. LAP expression by mononuclear cells is significantly lower in PreTx samples from patients who develop an irAE during treatment compared to patients who don't develop an irAE. (**A**) Boxplots of the normalised LAP expression by all MNCs for PreTx samples from either irAE patients (green) or no irAE patients (yellow). The p value and the FDR value were calculated using the diffcyt-DS-limma method from the diffcyt package. (**B**) t-SNE plots of mononuclear cells (MNCs) subset clusters from the cells stained with panel one markers. One point per cell. Points coloured by scaled latency associated peptide (LAP) expression. Left plot are cells from PreTx samples of OAC patients who developed an irAE during treatment. Right plot are cells from PreTx samples of OAC patients who did not develop an irAE during treatment.

4.4.3.2. After four weeks of ICI treatment, irAE developers no longer show signatures of reduced LAP expression and increased B cell activation

Although it is of most interest to identify PreTx signatures that associate with high risk of irAE development, as with CB, on-treatment monitoring of risk of irAE development would be clinically useful. Hence, next, differential analysis of the ICI-4W samples was performed.

At the ICI-4W timepoint, differential analysis of the abundances of the 70 cell subsets did not show the same enrichment of cell subsets in patients who developed an irAE that were observed in the PreTx samples, except for the increased abundance of IgM+ memory/MZ-like B cells in no irAE patients, although the difference in this subset appears to be primarily driven by one patient sample (**Figure 4.25**). Instead, it was found that specific subsets of mDCs, B cells, monocytes, CD8 T cells, and NK cells were significantly increased in irAE patients, whilst the subset of IgM+ memory/MZ-like B cells, the subset of cytotoxic NKG2C+ CD27+ CD8 T cells, and the subset of CD16-S7+ NK cells were significantly enriched in no irAE patients, at the ICI-4W timepoint (**Figure 4.29A**). Moreover, performing hierarchical clustering on the abundance data for these nine cell subsets largely clustered the ICI-4W OAC samples by irAE status (**Figure 4.29B**).

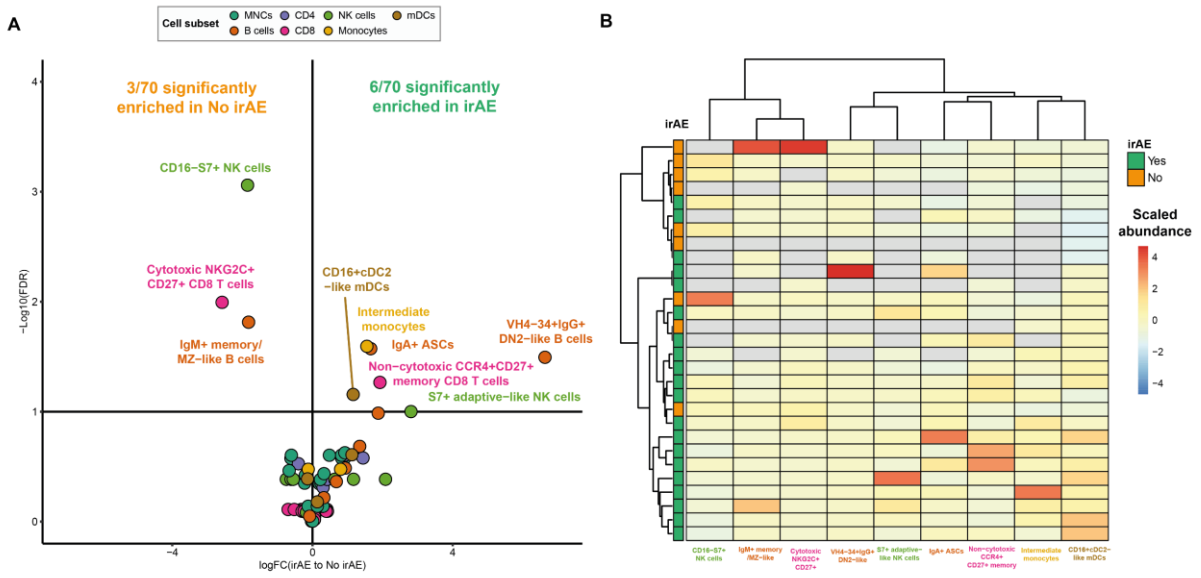


Figure 4.29. Lymphoid and myeloid cell subsets showed significantly different abundances between ICI-4W samples from OAC patients who do and do not develop irAEs during treatment. (**A**) Volcano-like plot showing the log fold change between irAE and no irAE OAC patients (logFC(irAE to No irAE)) versus the -Log10(FDR) for the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for the nine cell subsets with an FDR<0.1, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from irAE (green) or no irAE (yellow) patients. Grey cells indicate where abundance data for that cell subset was not available for that sample.

As with the PreTx samples, to determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between OAC ICI-4W samples from irAE

and no irAE patients. In total, based on an FDR cut off of 0.1, 40 features were found to be significantly different between these samples, with, once again, most of these features related to MNCs, B cells, or CD8 T cells, and increased FAS expression by IgM+ memory/MZ B cells were again significantly associated with irAE development (**Figure 4.30**). Although it should be noted that none of the features relating to LAP expression that were observed in the PreTx samples were not seen in these ICI-4W samples.

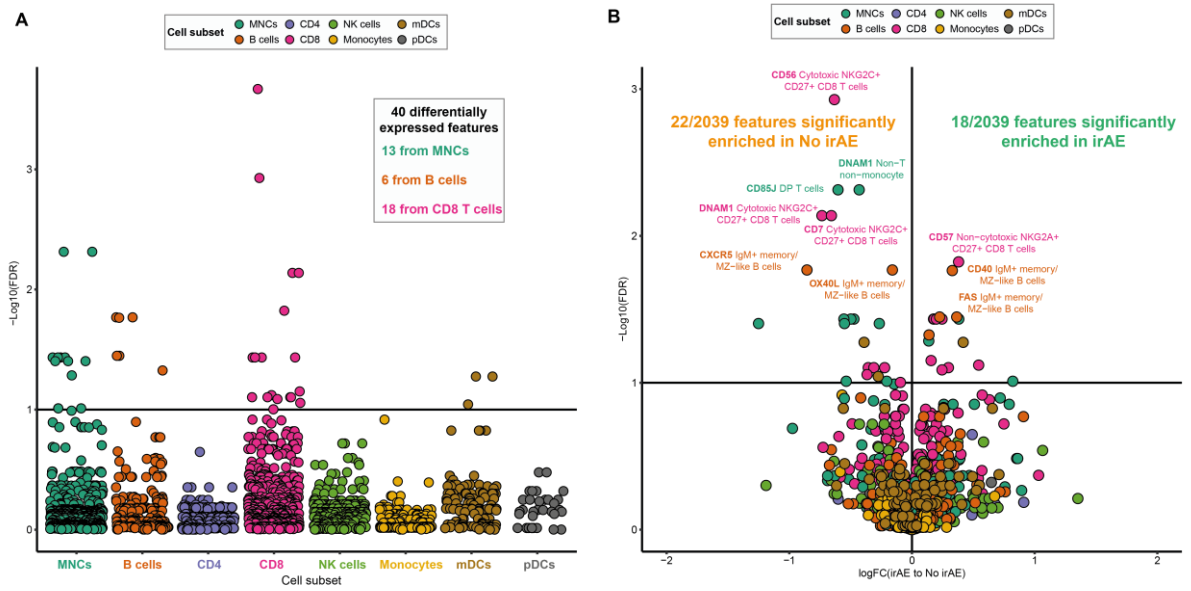


Figure 4.30. Subsets of mononuclear cells and CD8+ T cells showed significantly different cell state features between ICI-4W samples from OAC patients who do and do not develop irAEs during treatment. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between irAE and no irAE OAC patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between irAE and no irAE OAC patients (logFC(irAE to No irAE)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond to. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

4.4.4. Autoantibody seropositivity is associated with decreased CD22 expression by DN2/activated naïve-like B cells, and an increased IgA staining of myeloid cells

Although it was shown in the previous chapter that AAb seropositivity was not significantly associated with CB or irAE development in OAC patients from the LUD2015-005 study, this *ex vivo* immunophenotyping data from the LUD2015-005 patients nonetheless provides a valuable opportunity to identify peripheral immune signatures that are associated with AAb seropositivity in

OC patients. Such signatures would help to further understand what may drive AAbs formation in OC, and how these signatures might differ from those associated with CB and irAE, described above.

Subsequently, to identify signatures of AAb seropositivity, the abundance of cell subsets and cell state data of samples from OAC patients who were seropositive for at least one of the 33 RayBiotech autoimmune associated AAbs, and OAC patients who were seronegative for all of the 33 RayBiotech autoimmune associated AAbs were compared.

Differential analysis of the abundances of the 70 cell subsets showed that two subsets of similar cytotoxic CD27- memory CD8 T cells, which were identified from staining of PBMCs with two separate panels of markers (panel one and panel two) and so are likely include some of the same cells, were significantly increased in RayBiotech AAb seropositive patients, whilst the cytotoxic NKG2C+ CD27+ CD8 T cell subset was significantly enriched in RayBiotech AAb seronegative patients, at the PreTx timepoint (**Figure 4.31**). Moreover, performing hierarchical clustering on the abundance data for these three cell subsets largely clustered the PreTx OAC samples by RayBiotech AAb status (**Figure 4.31B**).

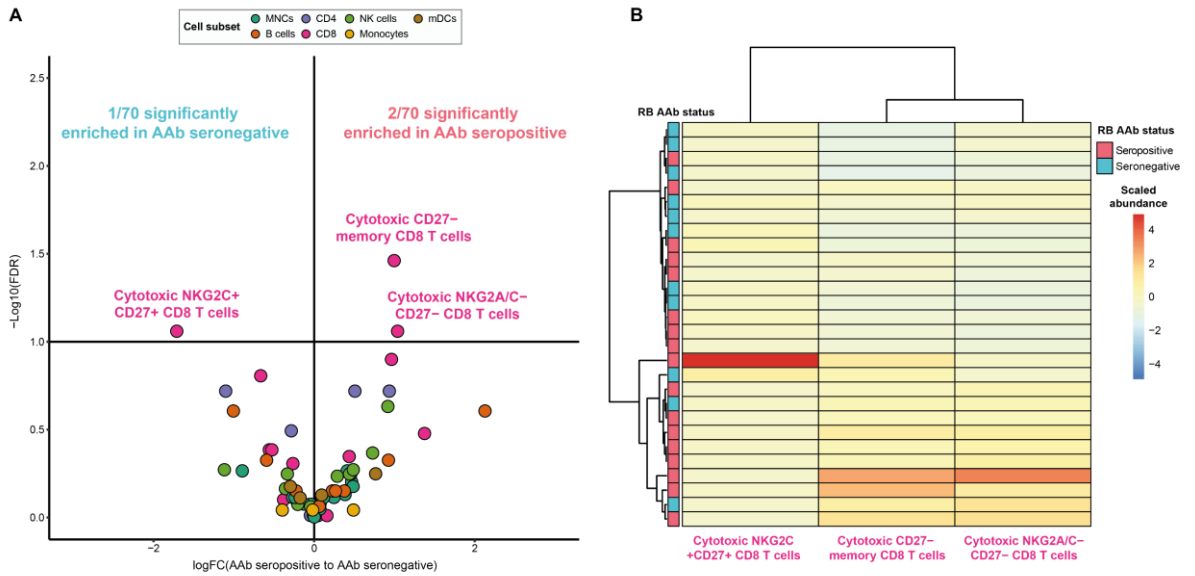


Figure 4.31. CD8 T cell subsets showed significantly different abundances between PreTx samples from OAC patients who were autoantibody seronegative and patients who were autoantibody seropositive. (**A**) Volcano-like plot showing the log fold change between autoantibody (AAb) seropositive and AAb seronegative OAC patients (logFC(AAb seropositive to AAb seronegative)) versus the -Log10(FDR) for the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for the three CD8 T cell subsets with an FDR<=0.1, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from AAb seropositive (red) or AAb seronegative (blue) patients.

To determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between OAC PreTx samples from patients who were RayBiotech AAb seropositive and patients who were RayBiotech AAb seronegative. In total, based on an FDR cut off of 0.1, 20 features were found to be significantly different between these samples, with significant features predominantly related to CD4 T cells, particularly the cytotoxic-like TEM subset (**Figure 4.32A-B**). Subsequent hierarchical clustering on the data for both the full set of 20 significant features, and the subset of 14 features related to CD4 T cells, largely clustered the PreTx OAC samples by RayBiotech AAb status (**Figure 4.32C-D**).

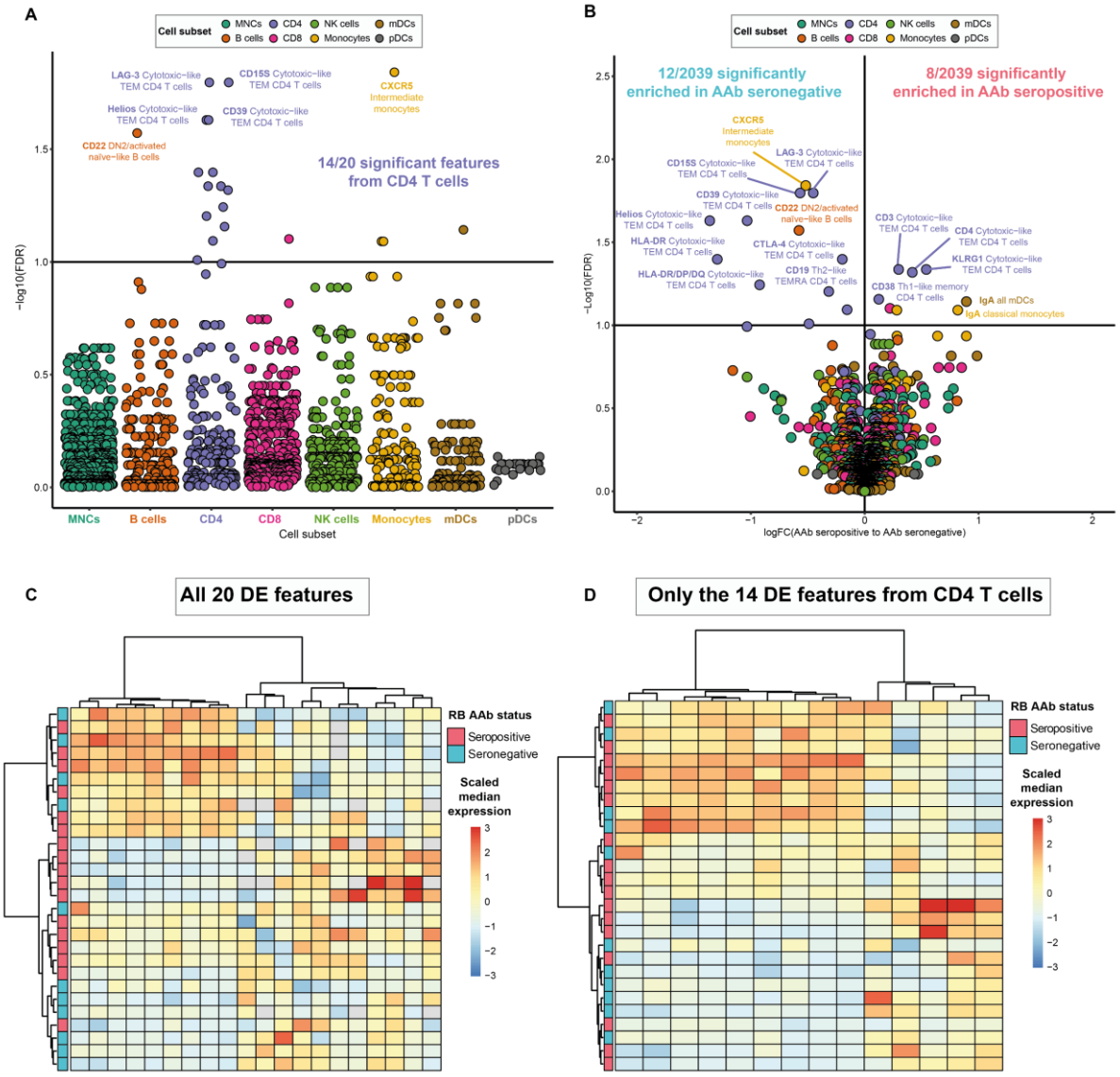


Figure 4.32. CD4+ T cells showed significantly different cell state features from PreTx samples from OAC patients who were autoantibody seronegative and patients who were autoantibody seropositive. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between autoantibody (AAb) seropositive and AAb seronegative OAC patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between autoantibody (AAb) seropositive and AAb seronegative OAC patients (logFC(AAb seropositive to AAb seronegative)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond to. (**C-D**) Heatmaps of the scaled cell state feature data for (**C**) all 20 features with an FDR<0.1 or (**D**) the 14 features from CD4 T cells. Data is scaled by column (feature), with unsupervised hierarchical clustering of both features (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from AAb seropositive (red) or AAb seronegative (blue) patients. Grey cells indicate where data for that feature was not available for that sample. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

In addition to the large number of features from CD4 T cells, expression of the co-inhibitory receptor CD22 by DN2/activated naïve B cells was a notable feature significantly enriched in AAb seronegative

patients. CD22 has been proposed to regulate both T cell independent responses to autoantigens which are driven by BCR and TLR signalling, and GC driven B cell responses[199,451,452].

The non-CD4 T cell related features enriched in AAb seropositive patients of note were related to IgA staining of myeloid cells. Although myeloid cells do not intrinsically express IgA, they are known to express FcRs, such as the myeloid specific FcR CD89 (also known as FcαRI), which are capable of binding to IgA[453,454]. To determine which cell types showed the highest IgA staining, and how this IgA staining differs between AAb seropositive and AAb seronegative patients, IgA staining of the MNCs, monocytes, and mDCs from panel three were further analysed.

The MNCs of patients who were RayBiotech AAb seropositive did not show significantly increased IgA staining relative to the MNCs from patients who were RayBiotech AAb seronegative (**Figure 4.33A-B**). However, when IgA staining of just the monocytes (**Figure 4.33C-D**) and mDCs (**Figure 4.33E-F**) were considered, the myeloid cells from patients who were RayBiotech AAb seropositive showed significantly increased IgA staining compared to the myeloid cell from patients who were RayBiotech AAb seronegative.

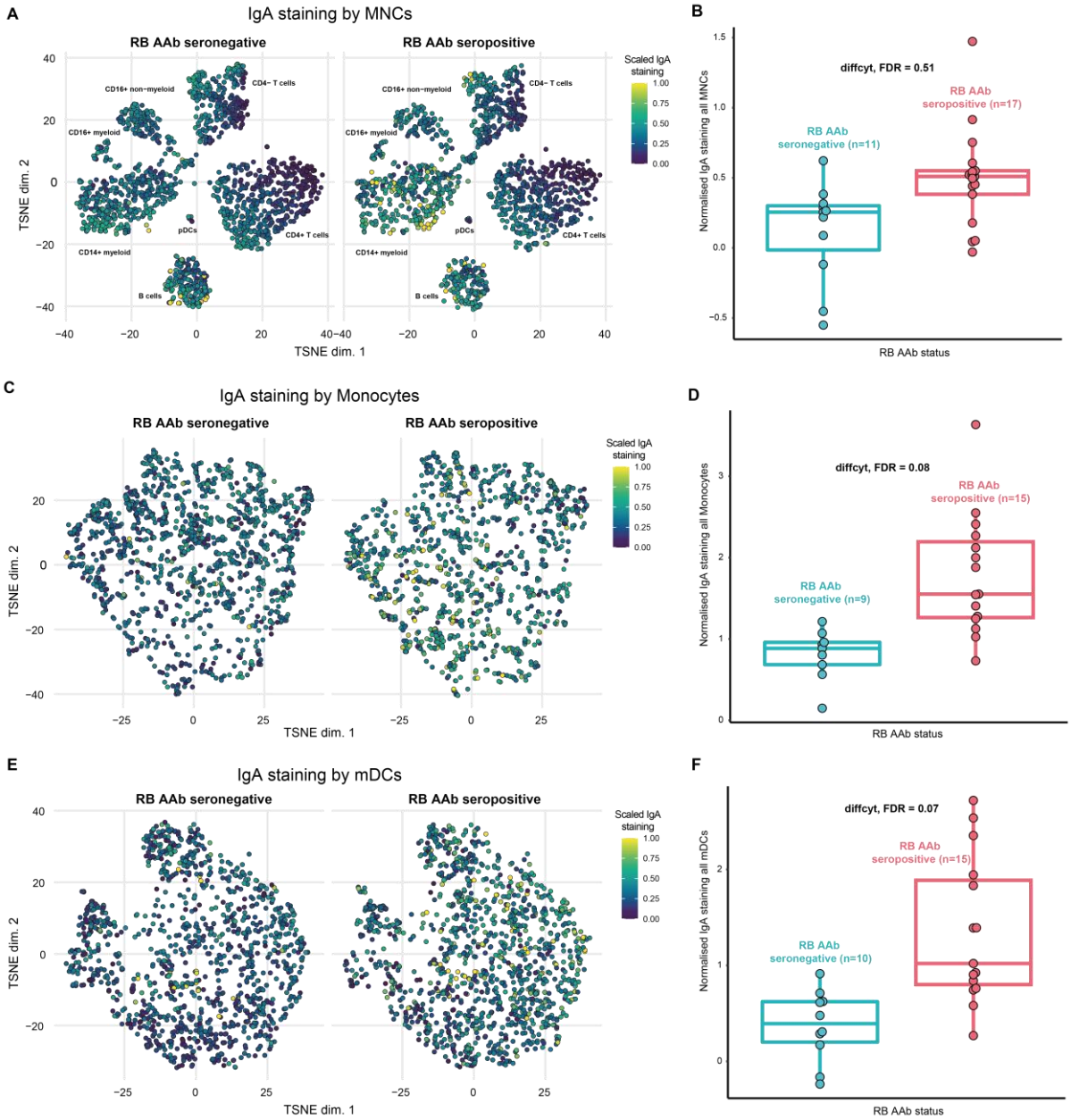


Figure 4.33. IgA staining of monocytes, and myeloid dendritic cells is significantly greater in PreTx samples from patients who are AAb seropositive compared to patients who are AAb seronegative. (**A**) t-SNE plots of mononuclear cells (MNCs) subset clusters from the cells stained with panel three markers. One point per cell. Points coloured by scaled IgA staining. Left plot are cells from PreTx samples of OAC patients who were autoantibody (AAb) seronegative. Right plot are cells from PreTx samples of OAC patients who were AAb seropositive, based on the RayBiotech (RB) array. (**B**) Boxplots of the normalised IgA staining by all MNCs for PreTx samples from either AAb seropositive (blue) or AAb seronegative OAC patients (red). (**C**) As for (A) except monocytes rather than MNCs. (**D**) As for (B) except monocytes rather than MNCs. (**E**) As for (A and C) except myeloid dendritic cells (mDCs) rather than MNCs or monocytes. (**F**) As for (B and D) except mDCs rather than MNCs or monocytes. p values and FDR values were calculated using the diffcyt-DS-limma method from the diffcyt package.

As myeloid cells also express FcRs which can bind IgG antibodies[455], the IgG staining of total MNCs and myeloid cells from these patients were analysed to determine whether the myeloid cells of RayBiotech

AAb patients had increased capability to bind IgG or whether they only had increased capability to bind IgA. This analysis showed no difference in IgG staining of total MNCs (**Figure 4.34A-B**), monocytes (**Figure 4.34C-D**), and mDCs (**Figure 4.34E-F**) between RayBiotech AAb seropositive and seronegative patients.

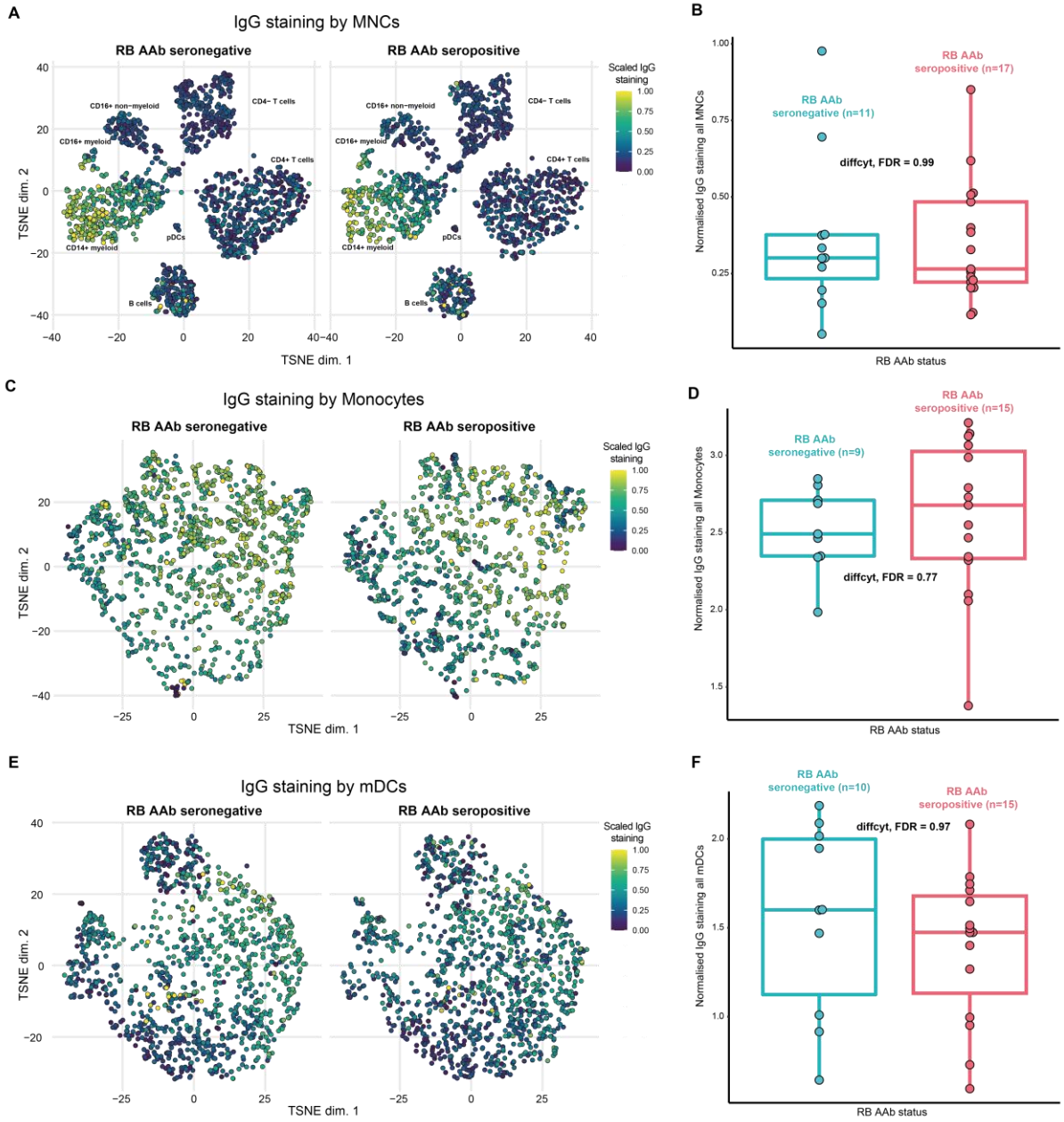


Figure 4.34. IgG staining of mononuclear cells, monocytes, and myeloid dendritic cells is not significantly different between PreTx samples from patients who are AAb seropositive and patients who are AAb seronegative. (**A**) t-SNE plots of mononuclear cells (MNCs) subset clusters from the cells stained with panel three markers. One point per cell. Points coloured by scaled IgG staining. Left plot are cells from PreTx samples of OAC patients who were autoantibody (AAb) seronegative. Right plot are cells from PreTx samples of OAC patients who were AAb seropositive, based on the RayBiotech (RB) array. (**B**) Boxplots of the normalised IgG staining by all MNCs for PreTx samples from either AAb seropositive (blue) or AAb seronegative OAC patients (red). (**C**) As

for (A) except monocytes rather than MNCs. (**D**) As for (B) except monocytes rather than MNCs. (**E**) As for (A and C) except myeloid dendritic cells (mDCs) rather than MNCs or monocytes. (**F**) As for (B and D) except mDCs rather than MNCs or monocytes. p values and FDR values were calculated using the diffcyt-DS-limma method from the diffcyt package.

4.4.5. Barrett's Oesophagus patients show significantly increased IgA expression by B cells than Oesophageal cancer patients

A secondary aim of this chapter was to identify signatures which can distinguish OC patients from BO patients. Subsequently, the abundance of cell subsets and cell state data of PreTx samples from OC patients and BO patients were compared.

Differential analysis of the abundances of the 70 cell subsets showed that the subset of S7- adaptive-like NK cells was significantly increased in OC patients, although this difference seemed to be driven by one patient sample (**Figure 4.35**). In contrast, the CD16+cDC2-like mDC subset, and the non-cytotoxic CCR6+CD27+ memory CD8 T cell subset were significantly enriched in BO patients. However, performing hierarchical clustering on the abundance data for these three cell subsets did not cluster the PreTx OC and BO samples by patient group (**Figure 4.35B**).

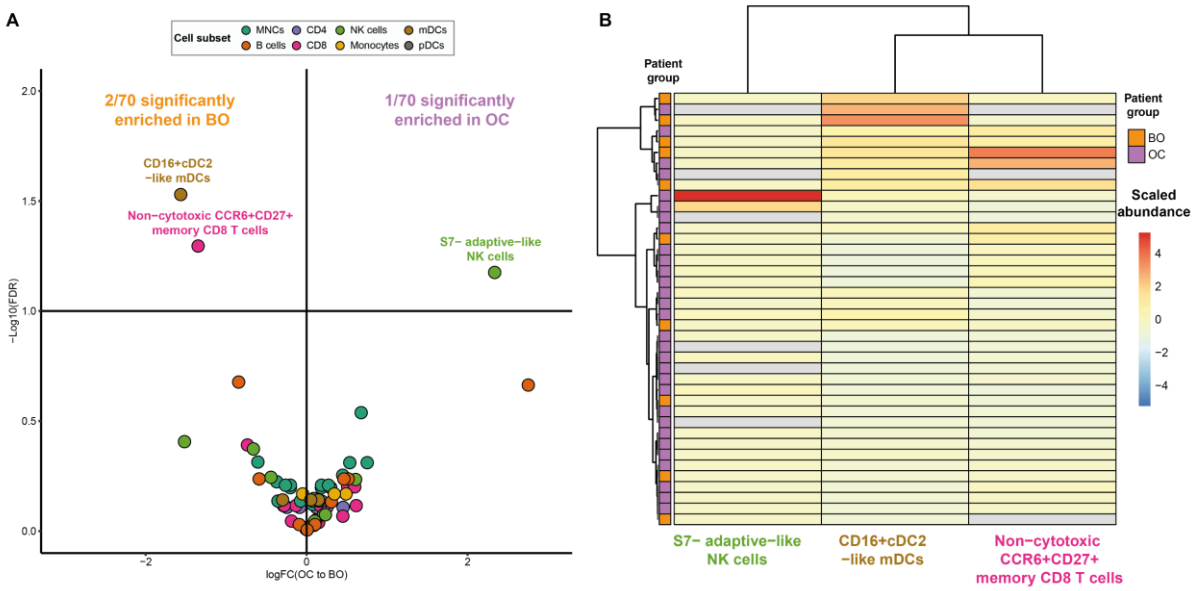


Figure 4.35. NK cell, CD8 T cell and myeloid dendritic cell subsets showed significantly different abundances between samples from OC patients and BO patients. (**A**) Volcano-like plot showing the log fold change between Barrett's oesophagus (BO) and oesophageal cancer (OC) patients (logFC(OC to BO)) versus the -Log10(FDR) for

the abundances of the 70 cell subsets identified by cell type clusterings. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. (**B**) Heatmap of the scaled cell subset abundance data for the three cell subsets with an FDR<0.1, scaled by column (cell subset), with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from OC (purple) or BO (orange) patients. Grey cells indicate where abundance data for that cell subset was not available for that sample.

To determine whether the expression profiles of cells were different, the features of cell state were analysed for differential expression between PreTx samples from OC patients and BO patients. In total, based on an FDR cut off of 0.1, 61 features were found to be significantly different between these samples, with significant features predominantly related to B cells, CD8 T cells, and NK cells (**Figure 4.36A-B**). Subsequent hierarchical clustering on the data for these 61 significant features largely clustered the samples by patient group (**Figure 4.36C**). As all of the samples had expression data for IgA by B cells (from clustering of MNCs), this data was more closely analysed. BO patients had significantly increased expression of IgA by B cells on average compared to untreated OC patients (**Figure 4.36D**).

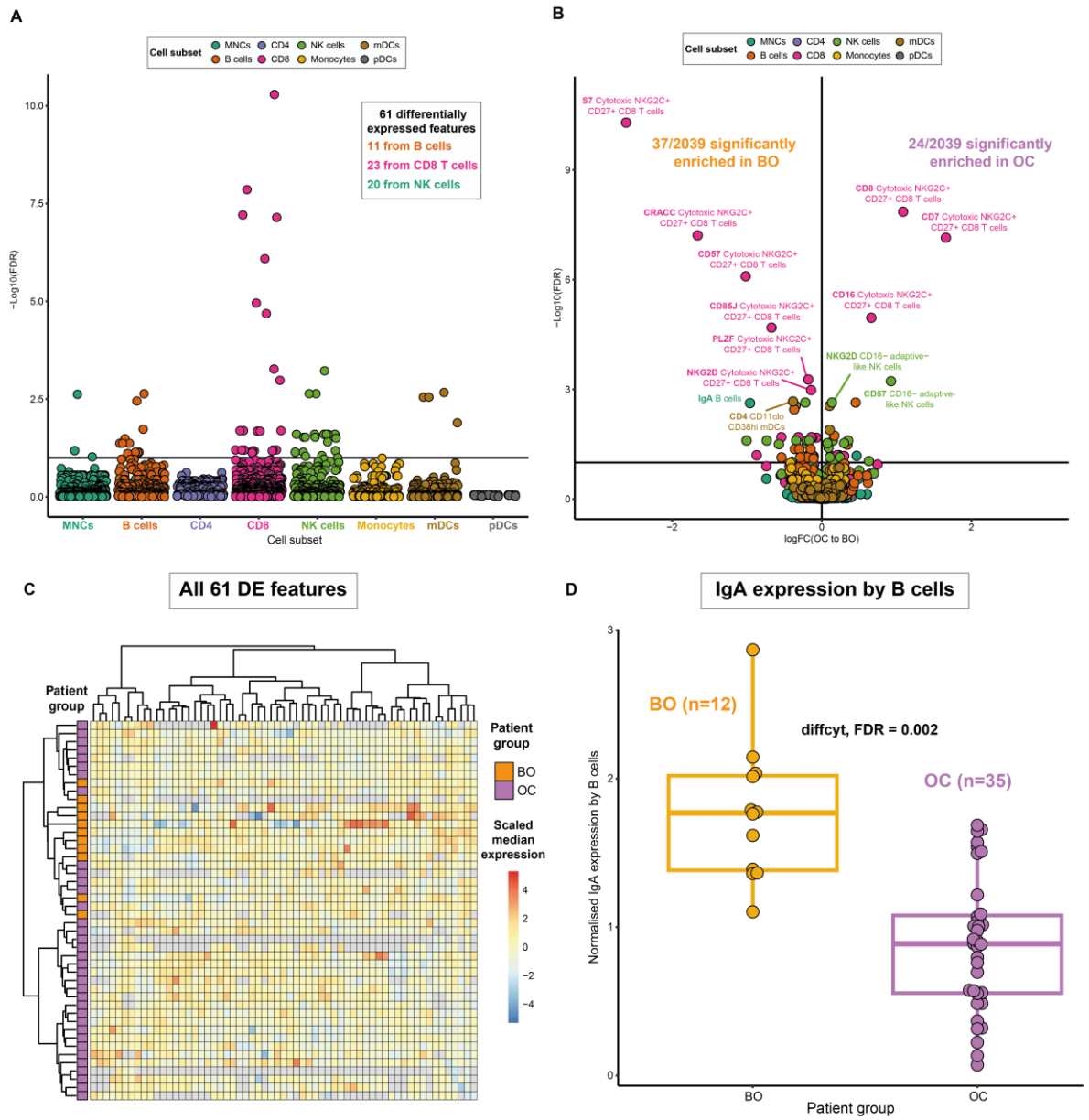


Figure 4.36. CD8+ T cells, B cells, and NK cells showed significantly different cell state features between samples from OC patients and BO patients. (**A**) Dotplot showing -Log10(FDR) for the 2039 cell state features identified by cell type down sampling and clusterings, when comparing between OC and BO patients, versus the cell type the state feature corresponds to**.** (**B**) Volcano-like plot showing the log fold change between OC and BO patients (logFC(OC to BO)) versus the -Log10(FDR) for the same 2039 cell state features. A horizontal line at a -log10(FDR) of 1.0 indicates the cut off for false discovery rate (FDR) which was 0.1. Points coloured by the cell type they correspond to. (**C**) Heatmap of the scaled cell state feature data for all 61 features with an FDR<0.1, with data scaled by column (feature), with unsupervised hierarchical clustering of both features (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from OC (purple) or BO (orange) patients. Grey cells indicate where data for that feature was not available for that sample. (**D**) Boxplots of the normalised IgA expression by the B cell cluster of MNCs comparing OC (purple) and BO (orange) patients. p values, FDRs, and logFC values were calculated using the diffcyt-DS-limma method from the diffcyt package.

Given the increased IgA expression by B cells of BO patients, the IgA and IgG staining of the MNCs from these patients were further analysed. This analysis showed that when considering total MNCs,

BO patients did not show significantly different staining of IgA or IgG compared to OC patients (**Figure 4.37**), despite showing increased IgA expression by their B cells (**Figure 4.36D**).

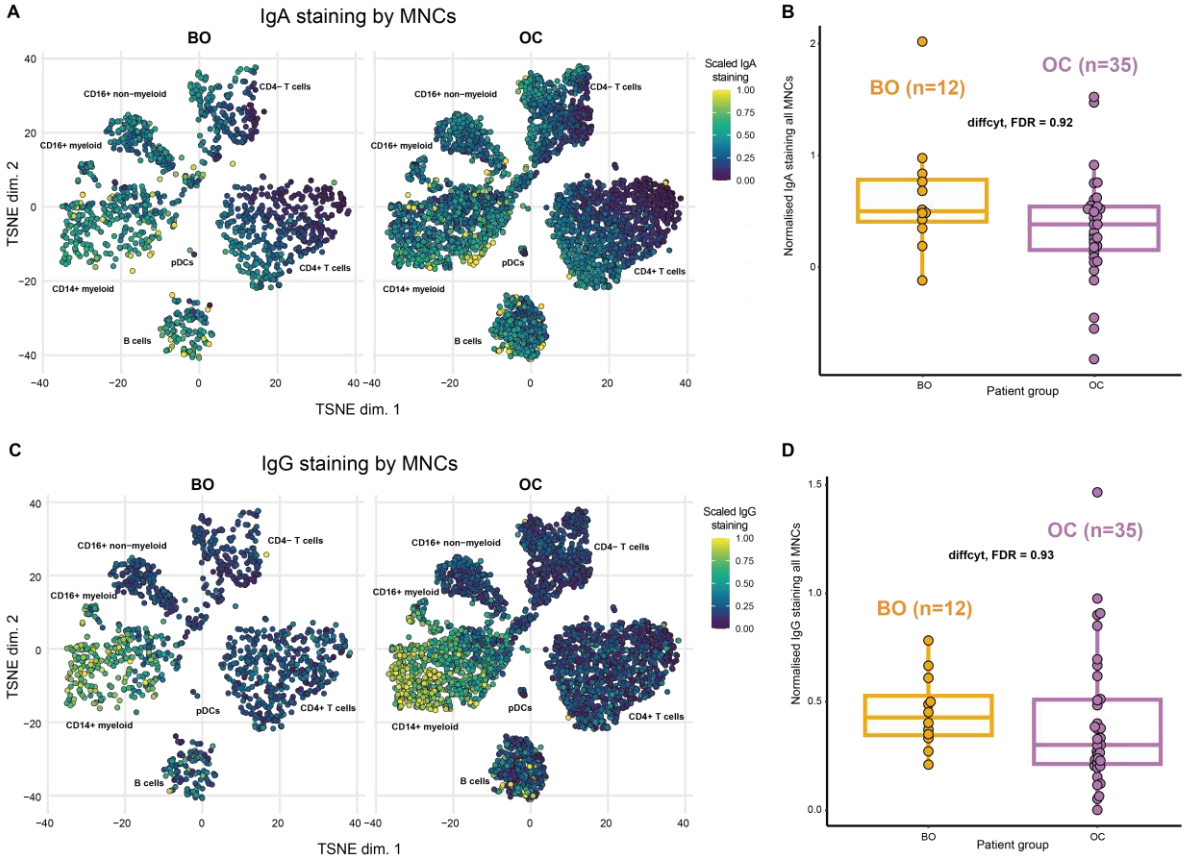


Figure 4.37. IgA and IgG staining of mononuclear cells is not significantly different between BO and OC patients. (**A**) t-SNE plots of mononuclear cells subset clusters from the cells stained with panel three markers. One point per cell. Points coloured by scaled IgA staining. Left plot are cells from patients with Barrett's oesophagus (BO). Right plot are cells from patients with oesophageal cancer (OC). (**B**) Boxplots of the normalised IgA staining by all MNCs for OC (purple) and BO (orange) patients. p values and FDR values were calculated using the diffcyt-DS-limma method from the diffcyt package. (**C**) As for (A) except IgG staining rather than IgA. (**D**) As for (B) except IgG staining rather than IgA.

4.5 Discussion

One of the aims of this thesis was to analyse whether peripheral signatures of increased immune activation were associated with improved clinical outcome and greater AAb numbers in OC patients treated with ICI+CTX. Although there was some overlap in the PreTx signature associated with CB, irAE development, and AAb status, for each of these factors, unique signatures relating to activation of different immune subsets were identified.

4.5.1 Comprehensive immunophenotyping of peripheral immune cells from patients with OC and BO

Comprehensive characterisation of PBMCs from HVs, OC patients and BO patients was performed by *ex vivo* immunophenotyping, using a high dimension cytometry approach which utilised three panels of 30+ markers, which were designed and validated by Dr Isabela Pedroza-Pacheco. Analysis of the corresponding data generated, using a semi-supervised discovery-oriented approach which was designed in collaboration with Dr Isabela Pedroza-Pacheco and based on the previously published CyTOF workflow[334,390], identified 70 subsets of immune cells, including several rare populations. Data for the abundances and marker expression of these cell subsets resulted in a total of up to 2039 features per sample. This is the first instance, to my knowledge, that comprehensive immunophenotyping of *ex vivo* PBMCs from ICI treated OC patients has been performed. This data provided a valuable platform for discovering peripheral immune cell signatures that were associated with various clinical variables of the patient samples profiled.

4.5.2 Signatures predictive of clinical outcome of OC patients to ICI treatment

4.5.3.1. Clinical benefit patients show signatures of increased activation of NKG2C+ CD8 T cells

Differential analysis comparing PreTx samples from OAC patients based on CB status identified significant differences in the abundance and expression profiles of specific subsets of B cells and CD8 cells. Two subsets of DN2-like B cells showed increased abundance in CB patients. B cell subsets with a similar expression profile to these two identified subsets were associated with SLE and were found

to be naïve cells which were more responsive to TLR7 signalling and subsequently prone to differentiation into plasmablasts via the extrafollicular activation pathway[456,457]. Therefore, an increased abundance of these subsets in CB patients could be indicative of an increased capacity for T cell independent responses to antigen and subsequent activation of B cells, which could then promote T cell activation.

The other two differentially abundant subsets were both subsets of cytotoxic CD8 T cells. The NKG2C+CD27- subset showed an increased abundance in CB patients, whilst the CD27+ memory subset showed a decreased abundance in CB patients. These two subsets differed primarily by their expression of the costimulatory molecule CD27, which has been suggested to be differentially expressed between CD8 TEM and central memory (TCM) T cells and inversely associated with the cytotoxic capacity of the cell[415,458,459]. The increased abundance of NKG2C+ expressing cells was particularly interesting as NKG2C is an activating receptor which binds to HLA-E and is normally expressed predominantly by NK cells but is expressed by a small fraction of CD8 T cells, particularly in states of chronic inflammation[424–426]. Like CD27- T cells, NKG2C+ T cells have been previously reported to demonstrate increased cytotoxic capabilities, to both infected and leukemic cells, and not upregulate the immune checkpoint receptor PD-1 after prolonged activation[430]. In addition to being more abundant, cells of the NKG2C+ CD27- CD8 T cell subset from CB patients also showed differences in their expression when compared to cells of the same subset from NCB patients. The expression of multiple NK associated markers differed between these cells, with an apparent skew towards markers of proliferation and activation, such as HLA-DR, Ki-67, and iKIRs, in the cells from CB patients[460,461]. This may suggest that the cytotoxic immune cells from CB patients have an increased baseline activation state prior to ICI treatment and therefore, after migrating to the tumour, can exhibit rapid and strong anti-tumour cytotoxicity following ICI initiation.

This is supported by the analysis of the ICI-4W timepoint at which, in addition to persistence of the activated expression profile of NKG2C+ CD8 T cells, the abundance of NK cells, both total NK cells and specific subsets of adaptive-like NK cells, were increased in CB patients. Together these observations are similar to findings of studies of melanoma and NSCLC in which cytotoxic CD8 T cells and NK cells were associated with improved clinical outcomes to ICI[131,134–136].

However, these finding should be carefully considered in the context that NKG2C+ CD8 T cells and adaptive-like NK cells are thought to expand following CMV infection[429,430], and similar observations have been demonstrated for the CMV seropositive samples used in this chapter. Hence, it would be important to determine whether the signatures of NKG2C+ CD8 T cells and NK cells in CB patients is solely driven by previous or recent CMV infection.

4.5.3.2. Development of immune related adverse events is associated with signatures of reduced LAP expression and a skew towards a more immature and activated B cell compartment

Unsurprisingly, given the association between good clinical outcome and irAE development, a similar signature of active NKG2C+ CD8 T cells was seen observed when differential analysis comparing PreTx samples from patients based on their development of irAEs was performed. However, the dominant signatures related to high risk of development of irAEs during ICI treatment were related to a more immature and activated B cell compartment, and reduced LAP expression.

The increased expression of activation markers, such as CD86, FAS, CD38, and Ki-67, in patients who developed an irAE suggests a skew towards a more activated B cell compartment. Expression of each of these four markers have been shown to be upregulated by B cells located in GCs and TLSs[285,462–464], suggesting that irAE developers may have more active GC and TLS responses, prior to ICI treatment. Moreover, as lower expression of CXCR5 by B cells has been reported in patients with RA and SLE[395],

the decreased CXCR5 expression in B cells from irAE developers suggests that these patients may have a B cell compartment which is more prone to autoreactivity.

Although there is no definitive marker for Bregs, IL-10 producing B cells are thought to be found at an increased prevalence in the transitional and memory B cell compartments[61,465,466]. Therefore, the skew towards an increase in CD24, CD27 expression in the patients who did not develop an irAE could be indicative of an increased prevalence of IL-10 producing Bregs, and hence an increased immunosuppressive capacity. This would agree with previous studies which have shown that Bregs are dysfunctional in lung cancer patients who develop an irAE[467], and in ICI treated melanoma patients, changes in B cells, associated with activation, early during ICI treatment preceded irAE development[468].

However, one limitation of the study described in this chapter is the lack of expression data for IgD, IL-10, and other B cell markers which would be useful for further characterising the balance between signatures of B cell mediated immune regulation and activation.

The other predominant signature associated with irAE development was decreased expression of LAP. LAP is bound to mature TGF-β[469] and so increased LAP expression suggests increased mature TGF-β expression. TGF-β plays an important role in immune suppression and prevention of autoimmunity[470,471]. Therefore, it is plausible that the patients who have increased mature TGF-β expression would have increased systemic immunosuppression and thus a reduced risk of developing an irAE during ICI treatment. Indeed, clinical trials of TGF-β inhibitors have reported a substantial number of irAEs during treatment, highlighting the important role of TGF-β mediated immune suppression[472].

Together, the reduced LAP expression and skew towards a more immature and activated B cell compartment is suggestive of a more active and less suppressive baseline immune state in OC patients who are at high risk of irAE development during ICI treatment.

Overall, these findings suggests that, in this cohort, CB to ICI is primarily mediated through the cytotoxic CD8 T cell compartment, whereas development of irAEs is primarily mediated though the B cell compartment. If correct, assessment of the CD8 T cell and B cell compartments of an OC patient prior to ICI treatment would be highly valuable for assessing their likelihood of both CB and irAE development, and subsequently help to aid treatment decisions.

However, these findings were identified using a broad discovery approach in a relatively small number of patients, and so they need to be more comprehensively validated in a larger independent cohort. For example, in addition to validating the signatures identified in an independent cohort, further *in vitro* assays to functionally validate activation/cytotoxicity of CD8 T cells, LAP expression and immunosuppressive capacity of myeloid cells, and B cell activation or Breg differentiation potential would be highly valuable.

4.5.3 Autoantibody seropositive patients show signatures of a dysregulated antibody response

Differential analysis comparing PreTx samples from patients based on AAb serostatus showed that there was a skew in the expression profiles of CD4 T cells, particularly the cytotoxic-like TEM subset. Although a reasonably rare subset, cytotoxic CD4 T cells have been identified in multiple states of chronic inflammation such as autoimmunity, chronic infection, and cancer[473]. The cytotoxic-like TEM CD4 T cells from AAb seropositive individuals showed a skew towards markers of activation and increased cytokine secretion, based on expression of CD38 and KLRG1[474,475], whilst AAb seronegative patients showed a skew towards a more immune regulatory profile, based on expression of LAG-3,

CTLA-4, and CD15S[476,477]. Interestingly, a recent study of anti-nuclear antibody (ANA) seropositive young adults observed a similar association of AAb seropositivity, in the absence of clinical autoimmunity, with signatures of a dysregulated CD4 T cell compartment[478].

However, arguably the most interesting signature was that AAb seropositive patients exhibited evidence of a dysregulated antibody response, which was characterised by decreased CD22 expression by DN2/activated naïve-like B cells, and an increased IgA staining of myeloid cells. Although, interestingly, there was no difference in IgG staining of myeloid cells seen between AAb seropositive and AAb seronegative patients. This is particularly surprising given that the antigen array to measure AAb seropositivity used in chapter 3 only detected IgG AAbs and not IgA AAbs. As briefly mentioned earlier, CD22 is an inhibitory co-receptor which has been proposed to regulate AAb responses[199,451,452].

One possible explanation for the increased IgA staining of myeloid cells that was observed is that the circulating myeloid cells of AAb seropositive patients may have increased expression of the myeloid specific IgA receptor CD89. Alternatively, it could be an indicator of increased circulating IgA and hence an increased likelihood of binding to myeloid expressed FcRs. Either way, it is unclear whether there is a functional consequence of the observed increased IgA staining of myeloid cells, although it has been suggested that engagement of CD89 by IgA can have either pro- or anti-inflammatory affects depending on whether the receptor is engaged by monomeric or multimeric IgA[479].

To further understand whether these IgG AAb seropositive patients have a more dysregulated antibody response, further experiments should characterise the IgA AAb serostatus against the same panel of 33 autoimmune disease associated antigens, measure the total plasma levels of IgG and IgA, and evaluate the expression of CD89 and other FcRs on the myeloid cells of these patient samples.

4.5.4 IgA expression by B cells is a promising biomarker to distinguish BO and OC patients

Differential analysis comparing untreated samples from OC patients and BO patients showed that there was significantly greater IgA expression by B cells in patients with BO. BO has been reported to be heavily infiltrated by cells involved in the humoral immune response such as CD4 T cells and ASCs,[480] but IgE and IgG were the predominant Ig isotype detected in BO tissue by immunohistochemistry, and not IgA[480]. However, increased expression of IgA, along with IgG and IgM, has been demonstrated in the oesophageal tissue of eosinophilic oesophagitis (EO) patients relative to tissue from HVs[481], and EO has been associated with an increased risk of BO development[482,483]. Hence, it is possible the increased IgA expression by B cells that is observed could be related to increased occurrence or history of EO in BO patients, although this would need to be studied further.

This is the first report, to my knowledge, of increased IgA expression by circulating B cells in BO patients relative to OC patients. However, given the small cohort size, this finding should be validated in an independent cohort with a larger number of patients. If validated, measurement of IgA expression by IgA B cells, in combination with detection of anti-p53 AAbs, could be prove highly valuable for non-invasively distinguishing patients with OC from those with BO, which could aid stratification of patients for endoscopic examination and hopefully increase detection of OC at an early stage. This finding also raises the question of whether detection of anti-p53 AAbs of the IgA isotype, rather than the IgG isotype like in chapter 3, would be more useful for detection of OC.

4.6 Conclusion

Comprehensive *ex vivo* immunophenotyping of ICI treated OC patients identified unique signatures associated with CB, irAE development, and AAb seropositivity. The predominant signature associated with CB was of increased activation of cytotoxic CD8 T cells, whilst the predominant signatures associated with irAE were a skew towards a more activated B cell compartment, and reduced immunosuppression. The identification of these signatures provides evidence for differing underlying mechanisms which may help to explain why some patients develop irAEs but do not show CB to ICI. Similar analysis related to AAb seropositivity identified an increase in IgA staining of myeloid cells, suggesting dysregulated expression of FcRs by myeloid cells in seropositive patients. Finally, analysis comparing samples from OC and BO patients identified IgA expression by B cells as a promising peripheral blood marker for distinguishing BO patients from OC patients.

In the next chapter, the second aim of my thesis will be addressed using data from bulk and scRNA-seq of tumour tissue collected from the LUD2015-005 study which will be analysed to determine whether transcriptional profiles of intratumoural B cells show evidence of GC responses. Paired BCR sequencing data from these samples will also be analysed and the BCR sequences from expanded BCR clones will be synthesised into mAbs and tested *in vitro* to determine whether the antibodies expressed by expanded BCR clones bind to identifiable autoantigens.

5 Intra-tumour B cells and antibodies in immunotherapy treated oesophageal cancer

5.1 Rationale

In the last two chapters, I addressed the hypothesis and aims of my thesis relating to peripheral signals of baseline immune activation and their potential association with the clinical outcome of LUD2015-005 participants to ICI treatment. However, the influence of B cells and TLSs in the tumours on the clinical outcome of these patients has not yet been explored. This is particularly pertinent considering that recent studies of ICI treated melanoma, sarcoma, and renal cell carcinoma (RCC) have shown an association of increased abundance of intra-tumour TLSs and B cells prior to treatment with good clinical outcome[106–109].

Given their high density of immune cells and tendency to form in response to chronic inflammation[111,114], in combination with their positive association with TMB and neoantigen counts[112,113], the presence of TLSs in a tumour could be a marker of a more immunogenic TME which is more likely to contribute to an effective anti-tumour immune response upon ICI treatment. TLSs have not been extensively studied in OC, but mature TLSs have been identified in OSCC and were associated with increased infiltration of B cells and other immune cells[484–486]. TLS have been less well studied in OAC, although one study showed that TLSs were present in OAC tumours and intratumour B cells were located predominantly in these TLSs[487]. However, despite being highly valuable, studies evaluating whether the abundance of intratumour B cells and TLSs associates with good clinical outcome to ICI in OAC have not yet been reported.

In addition to providing an environment for prolonged signalling and interaction between tumour infiltrating immune cells, further studies have supported the growing hypothesis that TLSs are important for maturation of intra-tumour B cells into ASCs which produce antibodies that can bind to tumour cells and contribute to anti-tumour immunity[109,276,286–288,290,488–491]. For example, a recent study identified that antibodies which were produced by circulating plasmablasts of metastatic cancer patients bound to public tumour antigens and caused regression of syngeneic tumours when tested

in mice[488]. Multiple additional studies have shown the presence of ASCs in the tumour and, in some cases, binding of endogenous IgG to tumour cells *in vivo*[109,287,288,489,490]. One of these studies also found that RCC patients who showed increased *in vivo* IgG binding to tumour cells, as assessed by immunofluorescent staining of tumour biopsies, experienced improved clinical outcome to ICI treatment[109]. Further, the authors of this study also reported that there were significantly increased rates of apoptosis in IgG-bound tumour cells; tumours which exhibited high tumour cell IgG binding also had increased macrophage infitration, and closer proximity of macrophages to tumour cells[109]. They went on to suggest that this could be indicative of direct antibody mediated tumour cell killing or increased ADCP as potential mechanisms of anti-tumour immunity provided by antibody binding[109].

Despite these encouraging findings, identification of the antigens that the antibodies expressed by intra-tumour B cells bind to has not been extensively studied. Two recent notable exceptions are a study of ovarian cancer in which patient derived monoclonal antibodies were shown to bind to matrix metalloproteinase 14 (MMP14)[288], and another study of head and neck cancer which identified tumour infiltrating ASCs which produced HPV-binding antibodies[489]. Although challenging, similar studies which identify antigenic targets of antibodies which are endogenously produced in OC tumours would be valuable for further understanding the drivers of antibody production in tumours, which could potentially be exploited by novel therapeutics.

This chapter addresses my hypothesis that the abundance and transcriptional profiles of intratumour B cells influences the clinical outcome of ICI treated OC patients, by perhaps acting as a marker of a highly immunogenic TME. In particular, this chapter aims to determine whether transcriptional profiles of intratumour B cells show evidence of GC responses and whether the antibodies expressed by expanded BCR clones bind to identifiable autoantigens.

5.2 Intratumour transcriptomic signatures of tertiary lymphoid structures and B cells in patient clinical outcome

Although intratumour TLSs and B cells are best explored histologically, various approaches utilising transcriptomics data to quantify signatures of TLSs and B cells have been reported[107,114,116–119]. Further, multiple studies have shown that the signatures quantified by these approaches associate with good clinical outcome in multiple cancer types[107,114,116–119]. Hence, due to the lack of histological material, but abundant transcriptomics data, which was available for the LUD2015-005 study participants, a transcriptomics approach to measuring signatures relating to TLSs and B cells in the tumours of the LUD2015-005 patients was taken.

To this end, in collaboration with Dr Thomas Carroll, two transcriptomic datasets of tumour and non-malignant tissue collected from LUD2015-005 participants were generated (**Figure 5.1**). The first of these was a bulk RNA sequencing data set generated of RNA isolated from 131 endoscopic tissue samples collected from LUD2015-005 participants (**Figure 5.1A**). The second of these datasets was the LUD2015-005 cell atlas which was generated by single cell RNA sequencing (scRNA-seq) of 59 tissue samples from 18 different patients, either OC patients from the LUD2015-005 clinical trial, or BO patients, using 10x Genomics 5' v1.1 platform (**Figure 5.1B and C**). For each sample, a gene expression library, and targeted BCR library were constructed. Full details regarding the experimental procedures used to generate these two datasets and the data processing performed by Dr Thomas Carroll are described in sections 2.11.3 and 2.11.5. Both of these datasets were then analysed for transcriptomic signatures of B cells.

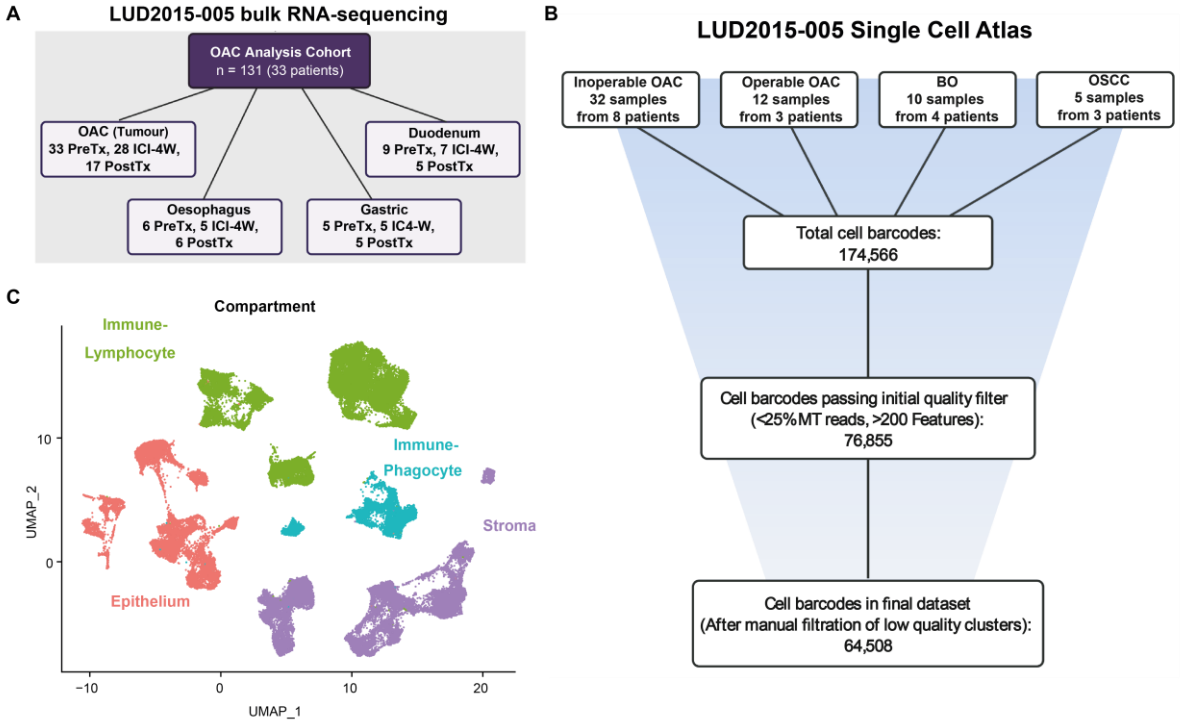


Figure 5.1. Summary of transcriptomic data available. (A) Bulk RNA sequencing (RNA-seq) was performed on RNA extracted from 131 endoscopic tissue biopsies collected from 33 patients of the LUD2015-005 study. 78 samples were OAC tumour, with 33 collected pre-treatment (PreTx), 28 collected after four weeks of immunotherapy treatment (ICI-4W), and 17 collected post-treatment (PostTx). 17 samples were oesophagus tissue, with 6 collected PreTx, 5 collected ICI-4W, and 6 collected PostTx. 15 samples were gastric tissue, with 5 collected PreTx, 5 collected ICI-4W, and 5 collected PostTx. 21 samples were duodenal tissue, with 9 collected PreTx, 7 collected ICI-4W, and 5 collected PostTx. (**B-C**) Single cell RNA-seq (scRNA-seq) was performed on cells isolated from 59 tissue biopsies which were collected from oesophageal cancer and Barrett's oesophagus patients. (**B**) Flow chart showing a summary of the tissue samples that scRNA-seq was performed on and the subsequent processing of the data that was performed by Dr Thomas Carroll. For most patients, multiple tissue samples, corresponding to the different treatment timepoints at which they were collected (e.g. PreTx, ICI-4W, or PostTx) or different tissues of origin (e.g. tumour, gastric, oesophagus, Barrett's, or scar), were processed. From these samples, in total 174,566 cell barcodes were identified, with 76,855 of these passing the initial quality filter of <25% mitochondrial (MT) reads and >200 RNA features, and 64,505 cell barcodes in final dataset after manual filtration of low-quality clusters. (**C**) UMAP plot showing the major cell compartments identified by dimensionality reduction of the final scRNA-seq dataset. One point per cell, with cells coloured by the compartment that the cell belongs to. Adapted from figures generated by Dr Thomas Carroll and used with his permission.

To determine whether transcriptomic signatures of TLSs were associated with clinical outcome to ICI treatment in the LUD2015-005 OAC patients, scores for two previously published TLS gene signatures[107,114] were generated for each of the PreTx OAC LUD2015-005 tumour samples from the bulk RNA sequencing data. Contrary to findings from previous studies of ICI treated melanoma, sarcoma, and renal cell carcinoma patients[106–108]; for the LUD2015-005 PreTx OAC tumour samples, neither of the TLS gene signatures associated with CB (**Figure 5.2**).

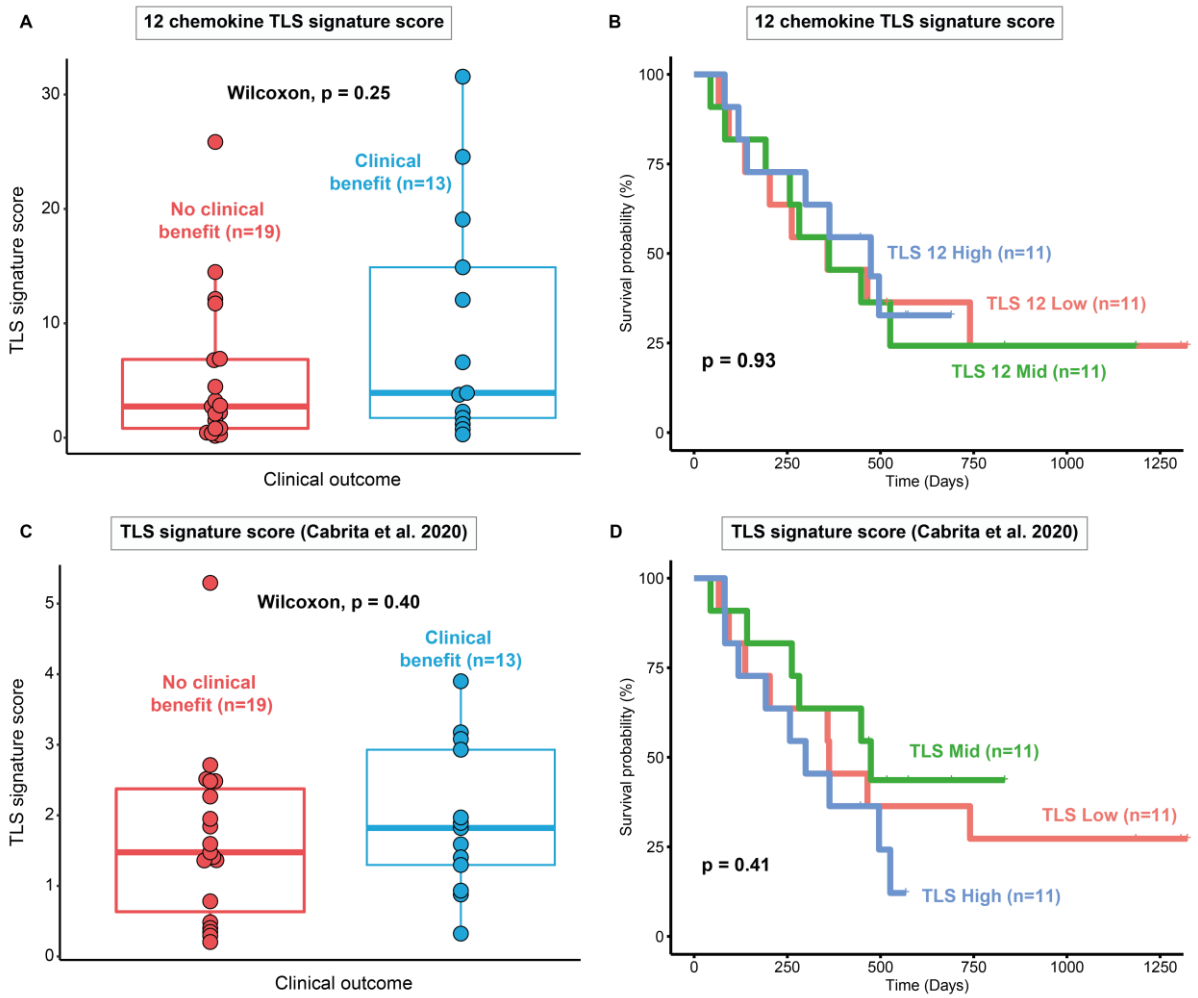


Figure 5.2. Transcriptomic signatures associated with tertiary lymphoid structures are not linked with good clinical outcome to ICI treatment in patients of the LUD2015-005 study. (**A**) Boxplots showing the tertiary lymphoid structure (TLS) transcriptomic signature score, based on gene expression values for 12 different chemokines which are associated with TLSs[114], between no clinical benefit patients (red) and clinical benefit patients (blue) for the PreTx OAC tumour samples from the LUD2015-005 study. (**B**) Kaplan Meier plot of overall survival with patients split into those with a high (top third) (blue), mid (middle third) (green), or low (bottom third) (red) TLS signature score. (**C-D**) Same as for (**A-B**) except the mean gene expression score for the nine genes reported by Cabrita et al., (2020)[107] to be associated with TLSs in melanoma. All gene expression values used were transcripts per million (TPM). For (**A**) and (**C**), statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant differences were observed. For (**B**) and (**D**), samples were trichotomised into high, mid, and low groups as has been done in previous publications assessing these signatures. p value was calculated using the log rank method. No statistically significant differences between groups were observed.

Similarly, analysis of the same PreTx OAC tumour samples, showed no association of either of the TLS gene signatures with irAE development during ICI treatment (**Figure 5.3**).

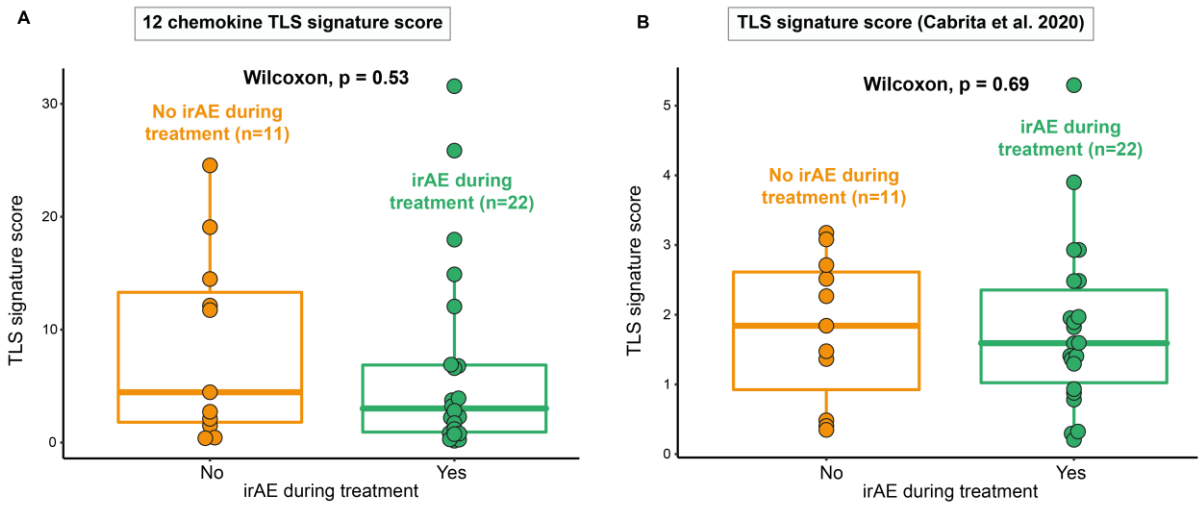


Figure 5.3. Transcriptomic signatures associated with tertiary lymphoid structures are not linked with immune related adverse event development in patients of the LUD2015-005 study. (**A-B**) Boxplots showing the tertiary lymphoid structure (TLS) transcriptomic signature scores, based on gene expression values for genes which are associated with TLSs[114], between irAE developers (green) and irAE non-developers (yellow) for the PreTx OAC tumour samples from the LUD2015-005 study. (**A**) shows the mean gene expression score for the 12 chemokines associated with TLSs, whilst (**B**) shows the mean gene expression score for the nine genes reported by Cabrita et al., (2020)[107] to be associated with TLSs in melanoma. All gene expression values used were transcripts per million (TPM). Statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant differences were observed.

Next, to determine whether the abundance of intratumour B cells was associated with clinical outcome, as has been suggested in studies of ICI treated melanoma, sarcoma and RCC[106–108], the MCP-counter deconvolution tool[340] was used to estimate the proportion of ten different cell subsets, including cells of the "B-cell lineage", in each of the PreTx OAC tumour samples based on the bulk RNA sequencing data. However, neither analysis of the overall profile of abundance of all of the 11 cell types, nor specific analysis of the "B-cell lineage" score generated by MCPCounter from the LUD2015-005 PreTx OAC tumour samples, showed any association with clinical outcome to ICI (**Figure 5.4**).

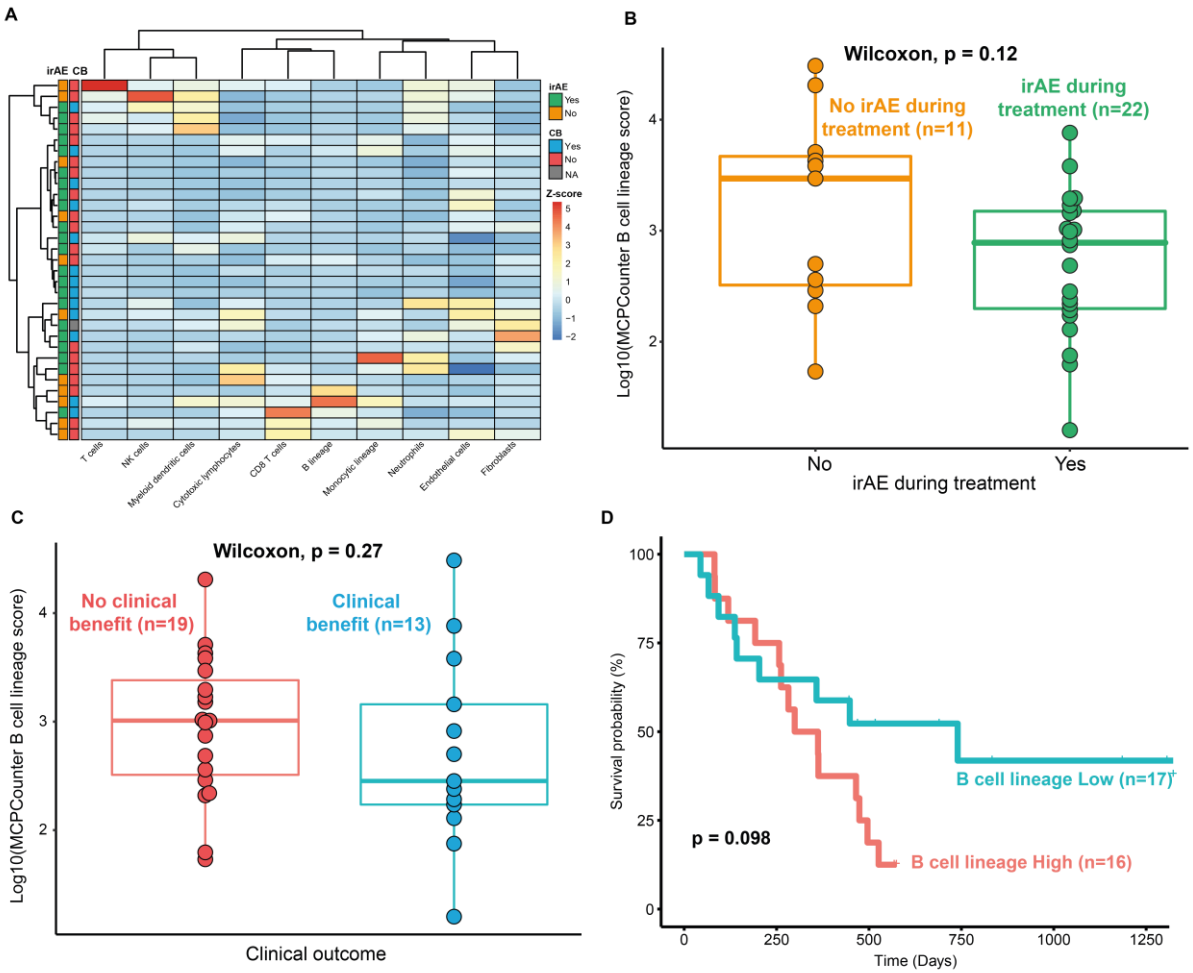


Figure 5.4. Pre-treatment transcriptomic signature of tumour B cell abundance does not associate with clinical outcome of patients of the LUD2015-005 study. (**A**) Heatmap of the z-score transformed MCPCounter deconvolution data for the 10 cell subsets calculated from the bulk RNA sequencing data from the PreTx OAC tumour samples, with unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from CB (blue), NCB (red), or CB NA (grey) patients, and another key indicating whether they were from irAE (green) or no irAE (yellow) patients. (**B-C**) Boxplots showing the log10 transformed MCPCounter B cell lineage score, between either (**B**) irAE developers (green) and irAE non-developers (yellow), or (**C**) NCB (red) and CB (blue) patients, for the PreTx OAC tumour samples from the LUD2015-005 study. (**D**) Kaplan Meier plot of overall survival with patients split into those with a high PreTx tumour B cell lineage score (> median) (red) and those who had a low PreTx tumour B cell lineage score (≤ median) (blue). p value was calculated using the log rank method. No statistically significant difference between the two groups was observed. For (**B**) and (**C**), statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant differences were observed.

To validate the results obtained from analysis of the MCPCounter deconvolution data, a deconvolution approach which uses scRNA-seq derived gene expression profiles from the scRNA-seq dataset(s)[492,493], in this case the LUD2015-005 cell atlas, was used to estimate the proportion of B cells that composed each of the PreTx OAC tumour samples, based on the bulk RNA sequencing data. An additional

advantage of this approach is that, for each sample, the proportion of different subsets of B cells can be inferred, as long as a signature matrix for them is identified from scRNA-seq data.

The LUD2015-005 cell atlas included 10,020 cells of the B cell lineage which had high quality GEX data and were assigned into different cell subsets by Dr Thomas Carroll (**Figure 5.5**). In the B cell compartment, naïve, mature, GC B cells were identified, based on differential expression of key marker genes such as IGHD, IGHM, IGHA1, IGH2, CD83, and CD69 (**Figure 5.5A and C**). In the plasma cell compartment, plasma cells expressing Ig chains of the IgG, IgA, or IgM isotype were identified, along with proliferating plasma cells which were characterised by high Ki-67 expression (**Figure 5.5B and D**).

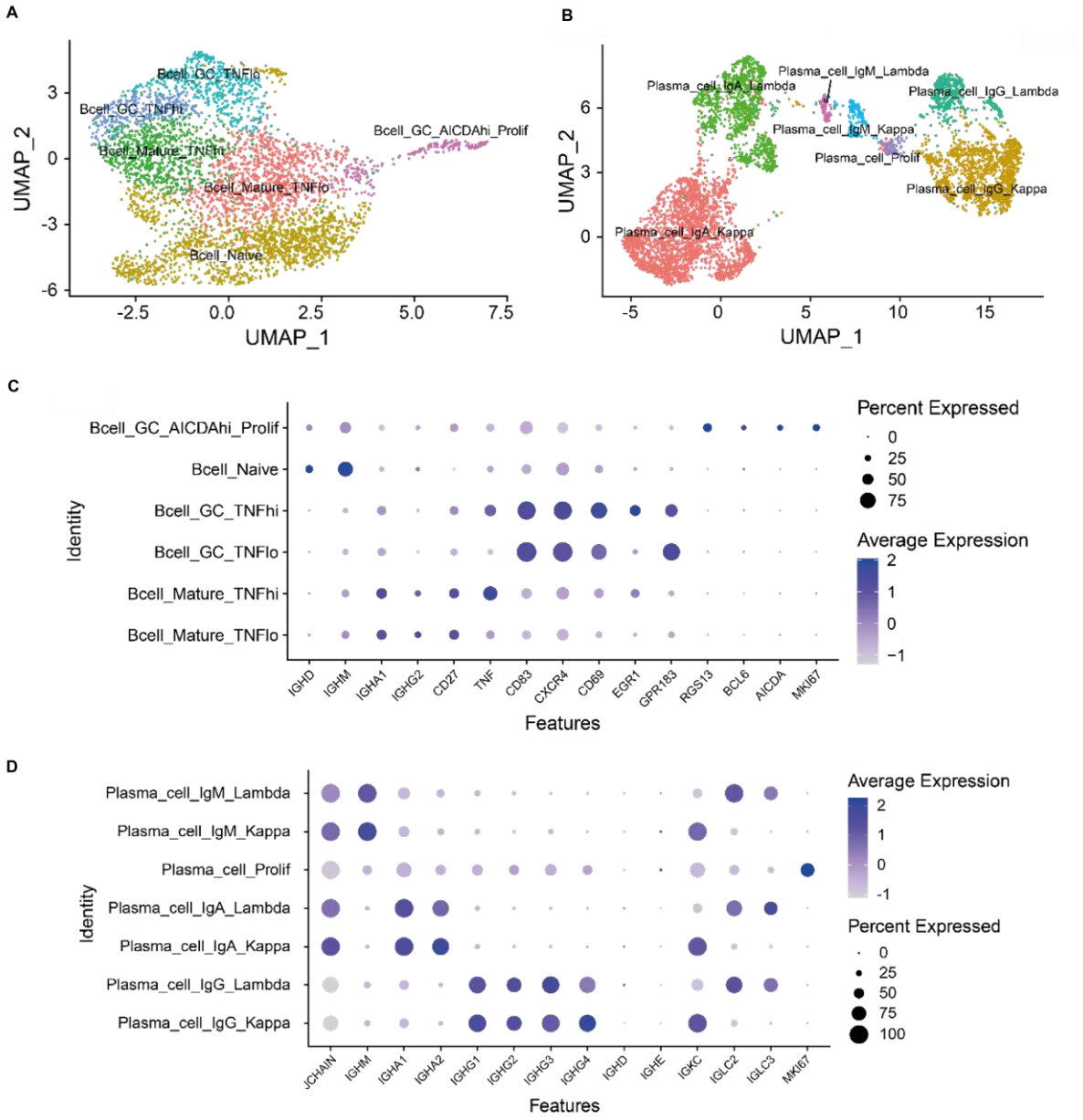


Figure 5.5. Clustering of transcriptomic data identified multiple subsets of B cells. Unsupervised clustering of single cell RNA-sequencing data for B cells from the LUD2015-005 cell atlas, performed by DrThomas Carroll, identified multiple subsets of non-plasma B cells and plasma B cells. (**A-B**) UMAP plots showing the different clusters of B cells which were identified. (**A**) Clustering of non-plasma B cells identified six clusters of cells. (**B**) Clustering of plasma B cells identified seven clusters of cells. (**C-D**) Dot plots showing selected genes that differentiated identified cell subtypes in (**C**) non-plasma B cells or (**D**) plasma B cells. The plots in this figure were generated by Dr Thomas Carroll and used in this thesis with his permission.

For the scRNA-seq based deconvolution, due to similarities in expression profiles and/or low cell numbers, some B cell clusters were merged to form a larger cluster, and this was done by Dr Thomas Carroll. This resulted in a total of seven B cell clusters which were used for deconvolution analysis, and the remaining data analysis in this chapter, and are summarised below (**Table 5.1**).

Compartment	*Original subset cluster*	*Merged subset cluster used for RNA deconvolution*
B cell	Bcell_GC_AICDAhi_Prolif Bcell_GC_TNFlo Bcell_GC_TNFhi	B cell GC
	Bcell_mature_TNFlo Bcell_mature_TNFhi	B cell mature
	Bcell_naive	B cell naïve
Plasma cell	Plasma_cell_IgM_Lambda Plasma_cell_IgM_Kappa	Plasma cell IgM
	Plasma_cell_IgA_Lambda Plasma_cell_IgA_Kappa	Plasma cell IgA
	Plasma_cell_IgG_Lambda Plasma_cell_IgG_Kappa	Plasma cell IgG
	Plasma_cell_Prolif	Plasma cell Prolif

Table 5.1. Summary of merging of similar B cell subset clusters into subsets for use in RNA deconvolution. Due to similarities in expression profiles and/or low cell numbers, some of the original 13 clusters of B cell subsets identified which had similar expression profiles were merged to form a total of seven subsets of B cells which were used for single cell sequencing based RNA deconvolution.

Deconvolution results for each of the seven B cell subsets obtained using BayesPrism[493], were provided to me by Dr Thomas Carroll. For each sample, the sum of each of these seven proportions was also used as a total B cell lineage proportion. Validation of the estimated B cell proportions for each of the PreTx OAC samples was performed by correlation analysis of the MCPCounter "B lineage" score and the BayesPrism total B cell lineage proportion, and this showed a significant positive correlation (**Figure 5.6A**).

However, subsequent analysis of the estimated total B cell lineage proportion for the PreTx OAC samples did not show any association with patient clinical outcome (**Figure 5.6B-D**).

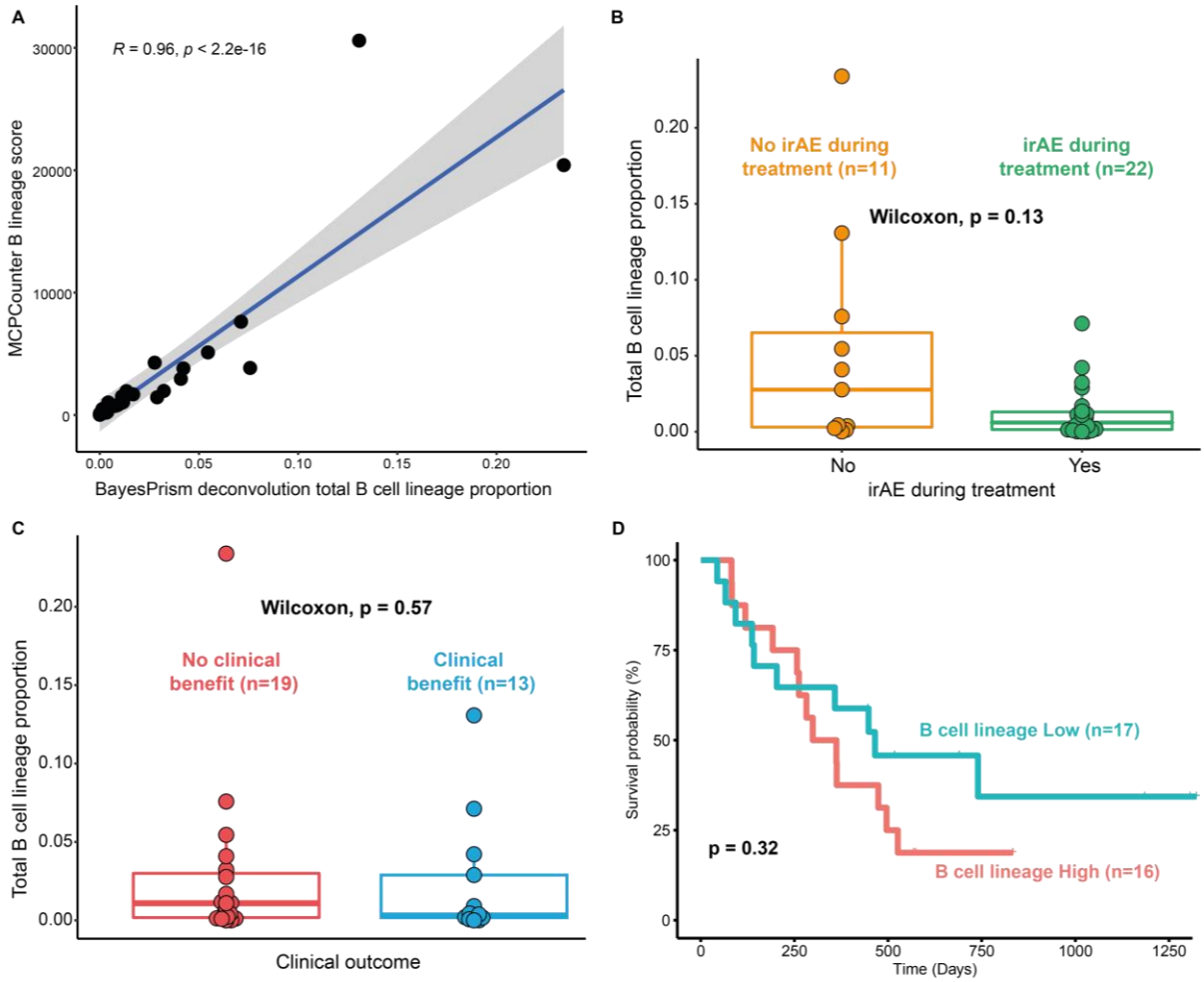


Figure 5.6. Estimated B cell abundance in the tumour pre-treatment does not associate with clinical outcome of patients of the LUD2015-005 study. (**A**) Scatterplot of the MCPCounter B lineage score versus the total B cell lineage proportion, calculated by Dr Thomas Carroll using BayesPrism. One point per sample. The R value and p value were calculated by fitting a linear model using Spearman rank correlation method. Shaded area indicates the 95% confidence interval for the straight line fitted. (**B-C**) Boxplots showing the total proportion of cells estimated to be from the B cell lineage, between either (**B**) irAE developers (green) and irAE non-developers (yellow), or (**C**) NCB (red) and CB (blue) patients, for the PreTx OAC tumour samples from the LUD2015-005 study (**D**) Kaplan Meier plot of overall survival with patients split into those with a high PreTx tumour B cell lineage score (> median) (red) and those who had a low PreTx tumour B cell lineage score (≤ median) (blue). p value was calculated using the log rank method. No statistically significant difference between the two groups was observed. For (**B**) and (**C**), statistical comparisons of group medians were carried out using a Wilcoxon test. No statistically significant differences were observed.

Although, once again, no association of the abundance of cells of the B cell lineage in tumours with clinical outcome was observed, it is possible that the abundance of certain B cell subsets may be more relevant. For example, in a study of ICI treated metastatic melanoma, irRECIST non-responders were found to have a PreTx intra-tumour B cell profile skewed towards naïve B cells, whereas irRECIST responders had a profile skewed towards memory B cells and plasma cells[108]. To explore this, the

scRNA-seq based deconvolution proportions for the seven subsets of intra-tumour B cells identified were analysed for association with clinical outcome.

However, the subsequent analysis did not show any association of particular subsets of intra-tumour B cells with patient clinical outcome (**Figure 5.7A**). When analysing the abundance of subsets relative to total B cells, the B cells of most samples were of predominantly either IgG or IgA plasma cells (**Figure 5.7B**).

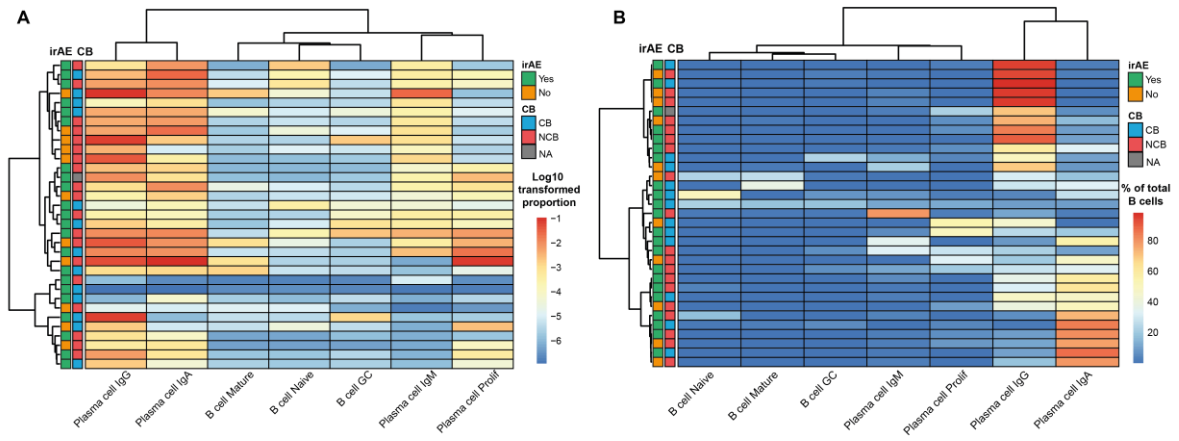


Figure 5.7. Pre-treatment transcriptomic signatures of tumour B cell composition shows no association with clinical outcome of patients of the LUD2015-005 study. (**A-B**) Heatmaps of abundance data of the seven B cell subsets, as calculated by BayesPrism, for the for the bulk RNA sequencing data from the 33 PreTx OAC tumour samples, with data shown as either (**A**) the log10 transformed proportion relative to all cell types found in the tumour, or (**B**) the percentage abundance relative to the total B cells in the tumour. Unsupervised hierarchical clustering of both cell subsets (columns) and samples (rows) was performed using the hclust method with "ward.D" linkage. Samples (rows) are annotated with a key indicating whether they were from CB (blue), NCB (red), or CB NA (grey) patients, and another key indicating whether they were from irAE (green) or no irAE (yellow) patients.

5.3 Single cell RNA sequencing identifies expanded BCR clones within patient tumours

Although there was no association of intra-tumour B cell abundance with patient clinical outcome, scRNA-seq identified the presence of cells from all stages of B cell maturation in the tumour, suggesting that a full B cell response is generated in OAC tumours. Hence, it is possible that the clonal selection, amplification, and maturation of the intra-tumour B cells, rather than their sheer abundance, is more relevant for mounting an effective anti-tumour immune response in OAC.

This would ideally be determined through in-depth BCR repertoire profiling of RNA isolated from tumour samples which has been amplified for Ig transcripts through targeted amplification of VDJ regions, using an approach like that which has been used in a study of immune-mediated diseases[494]. However, although generation of such data is planned, similar analysis is not yet possible for the tumour samples of LUD2015-005 participants. Instead, to explore B cell clonal selection, amplification, and maturation in ICI treated OC, GEX and BCR sequencing data for the 10,020 B cells from the LUD2015-005 cell atlas were analysed for evidence of B cells undergoing GC responses.

5.3.1. Transcriptomic signals of germinal centre responses

Of the 10,020 B cells sequenced, 5,912 cells had high quality GEX data, as determined by Dr Thomas Carroll, and paired IgH and Igκ/λ BCR data. Subsequent analysis of the BCR sequences for these 5,912 cells, using the Change-O[344] package, identified 4,484 different BCR clones. After filtering to only the BCR clones that had a clone size of two cells or greater, 491 expanded BCR clones with paired IGH and IGK/L sequences were identified (**Figure 5.8A**). The abundance of B cells and expanded BCR clones varied substantially across the samples sequenced (**Figure 5.8B-C**).

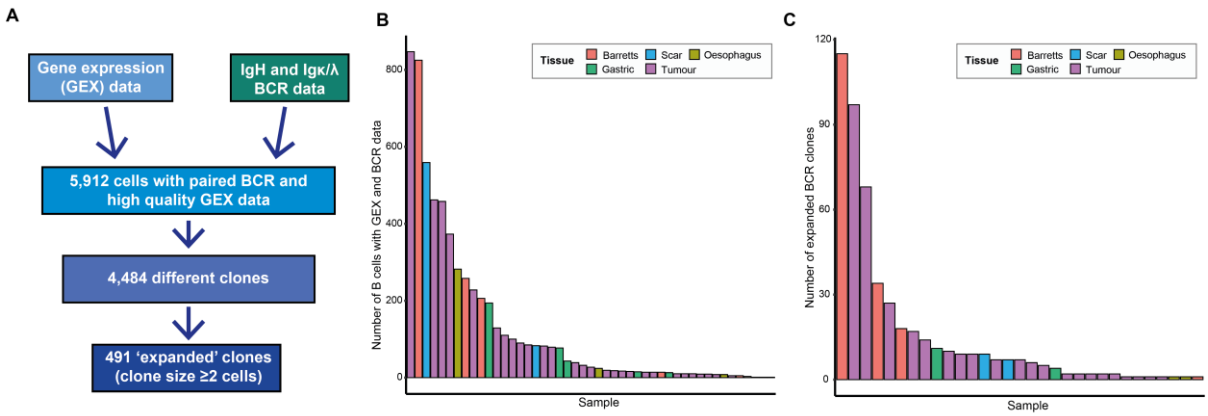


Figure 5.8. Combined single cell RNA and BCR sequencing identifies expanded BCR clones across multiple tissue types. (A) Flow chart showing a summary of the processing of gene expression (GEX) and BCR data to identify expanded BCR clones. Across all 59 samples sequenced, 5,912 B cells had high quality GEX data, as determined by Dr Thomas Carroll, and paired IgH and Igκ/λ BCR data. Change-O analysis of the BCR sequences for these 5,912 cells identified 4,484 different BCR clones, with 491 of these clones expanded, based on a clone size of 2 cells or greater. (**B-C**) Bar plots showing the either (**B**) the number of B cells with high quality GEX and paired BCR data or (**C**) the number of expanded BCR clones for each of the samples in which B cells with high quality GEX and paired BCR data were sequenced. Bars are coloured by the histological tissue type of the sample.

To determine whether intra-tumour B cells show evidence of SHM, and hence have likely experienced GC reactions, the frequency of mutations in the variable region of the *IGH* chain, henceforth referred to as the Vh segment, and CDR3 region, which are both highly variable and contribute to determining antigen binding[163,495–497], relative to the relevant germline sequence were calculated from the BCR sequencing data using the R package SHazaM[344]. This data was then used to compare the mutation frequencies across different B cell subsets (**Figure 5.9**).

On average, naïve B cells showed the lowest frequency of total mutations in both the Vh segment and CDR3 region, with the other B cell subsets all showing a similar average frequency of mutations, although IgA and IgG plasma cells showed the highest average frequency of total mutations (**Figure 5.9A and C**). As mutations can be silent and not result in a change of the amino acid encoded, and thus in theory not alter the antigen specificity of the BCR, the analysis was repeated only considering the frequency of mutations which result in change of the amino acids encoded, also referred to as replacement or non-silent mutations[498]. This analysis showed almost identical results for the Vh segment (**Figure 5.9B**), but showed that cells of the IgG, IgA, and proliferating plasma cell subsets had much higher average frequency of replacement mutations in the CDR3 region, compared to the other four B cell subsets (**Figure 5.9D**).

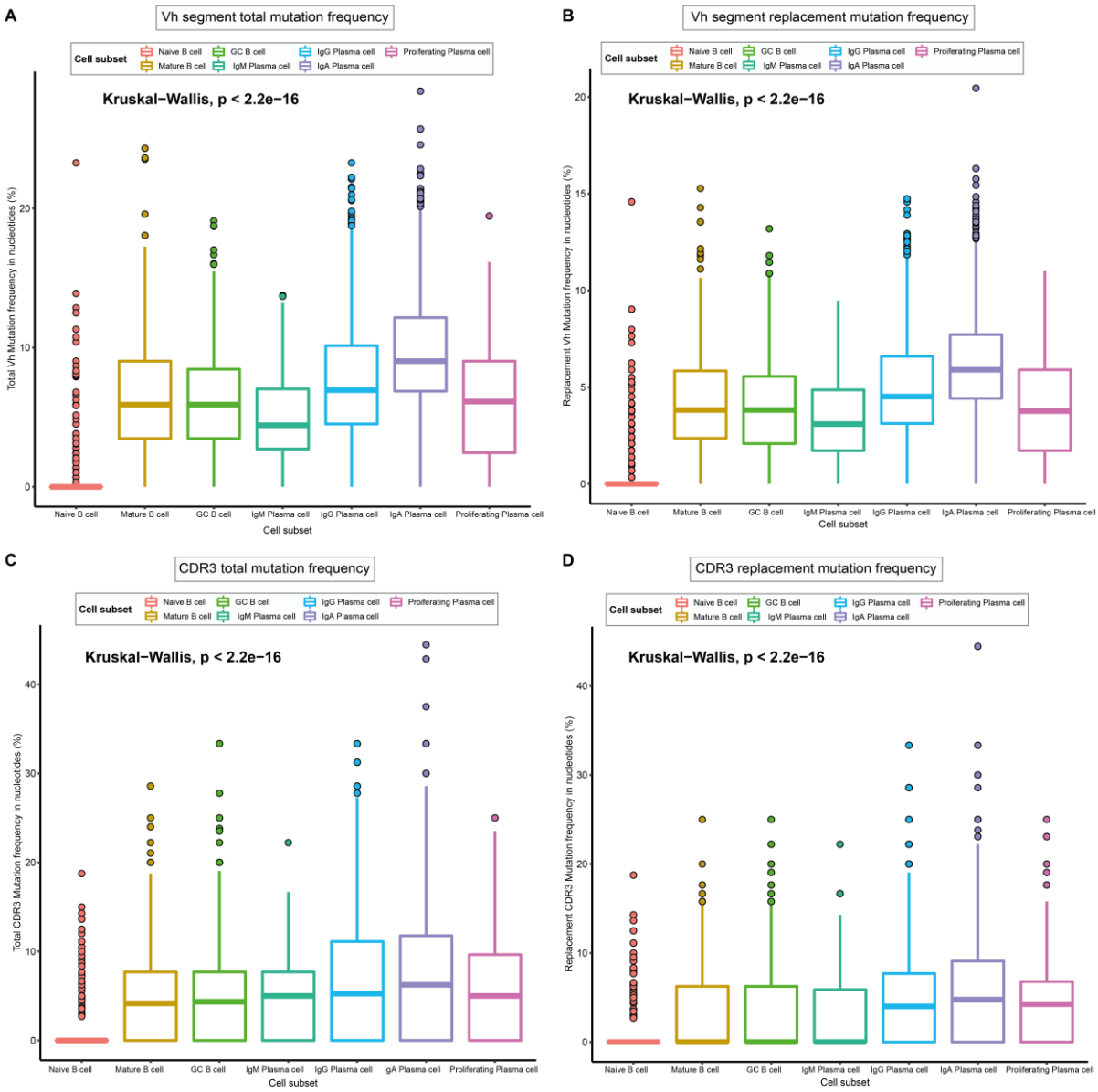


Figure 5.9. IgA and IgG plasma cells show the highest frequency of replacement mutations in their Vh gene segment and CDR3 region. Boxplots showing the frequency of either (**A**) total mutations in the Vh segment, (**B**) replacement/non-silent mutations in the Vh segment, (**C**) total mutations in the CDR3 region, or (**D**) replacement/non-silent mutations in the CDR3 region, each for the B cells across all samples and each separated by the B cell subset expressing the BCR. Mutation frequencies were calculated using the SHazaM package in R and are relative to the germline sequence. One point per cell/BCR, coloured by B cell subset. Statistical comparisons of group medians were carried out using a Kruskal-Wallis test.

Next, to determine which B cell subsets show the most frequent clonal expansion, the total number of B cells, and number of expanded BCR clones per 100 cells were calculated and compared for each of the different B cell subsets. This analysis showed that expanded B cell clones had the highest likelihood of differentiating to plasma cells (**Figure 5.10**).

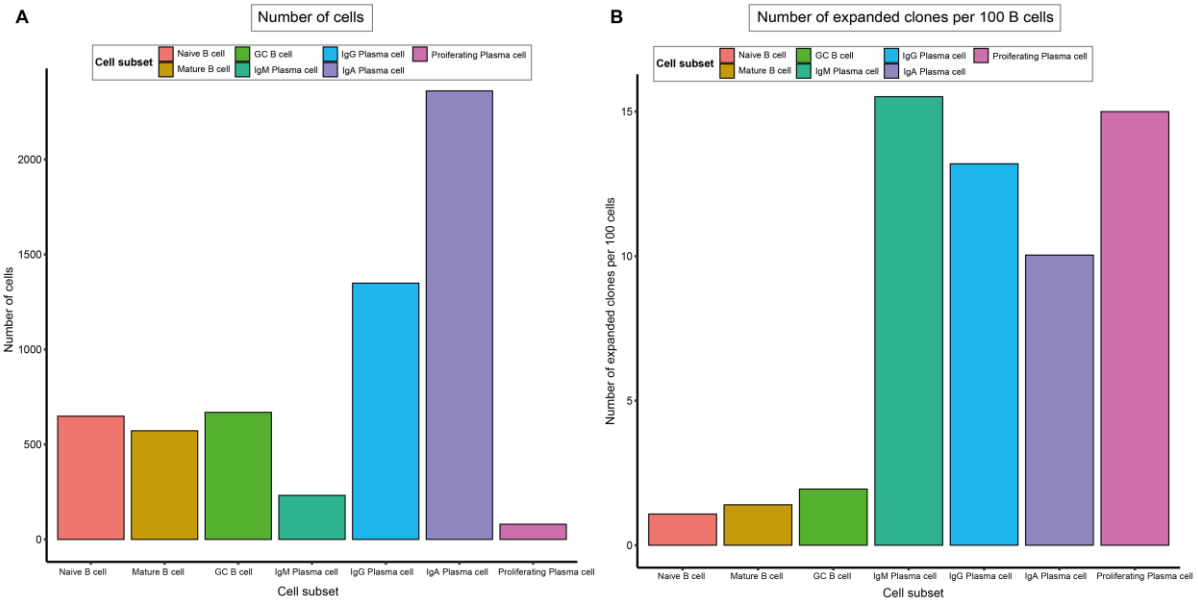


Figure 5.10. Plasma cells accounted for a high percentage of B cell clones. Bar plots showing either (**A**) the number of B cells or (**B**) the number of expanded BCR clones per 100 cells, across all of the samples, separated by B cell subset. Bars are coloured by B cell subset.

Given the skew towards an increased number of replacement mutations in the IgA and IgG plasma cell subsets, the next step was to determine whether certain isotype subclasses of BCRs showed differing degrees of SHM. Hence, similar analysis of mutations and number of expanded clones was performed in the context of BCR isotype subclass, based on the IgH constant region that was called by Change-O. This analysis showed that BCRs of the IgD or IgM isotype, which are predominantly expressed by naïve B cells, had the lowest average mutation frequency in both the Vh segment and CDR3 region (**Figure 5.11)**. The BCRs of the remaining isotypes and subclasses showed similar average frequencies in both total and replacement mutations in the Vh segment (**Figure 5.11A and B**). However, BCRs of the IgG3 subclass showed reduced average frequency of both total and replacement mutations in the CDR3 region compared to the other IgG and IgA subclasses (**Figure 5.11C and D**).

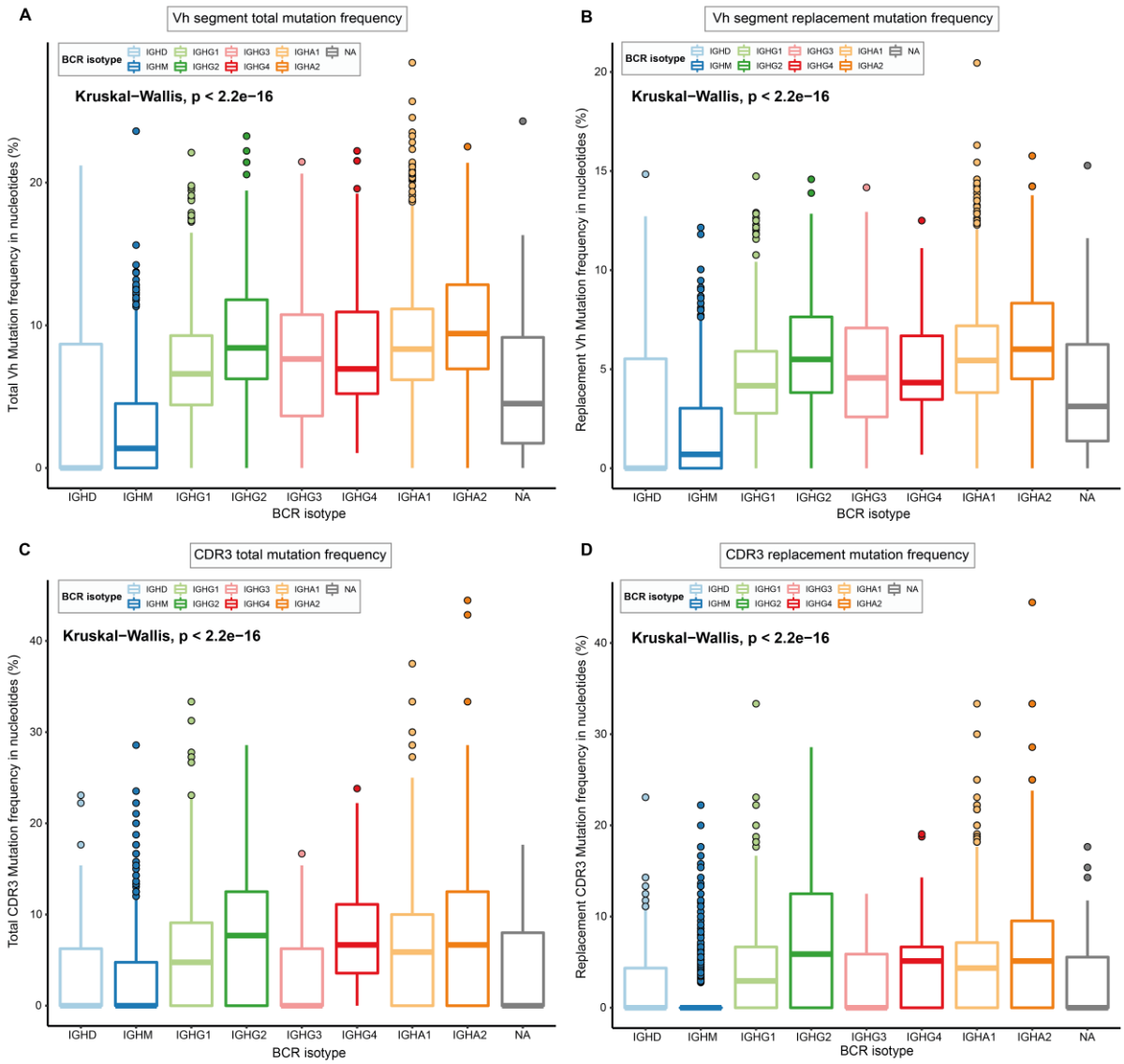


Figure 5.11. BCRs of the IgA or IgG isotype are on average more frequently mutated than BCRs of the IgD or IgM isotype. Boxplots showing the frequency of either (**A**) total mutations in the Vh segment, (**B**) replacement/non-silent mutations in the Vh segment, (**C**) total mutations in the CDR3 region, or (**D**) replacement/non-silent mutations in the CDR3 region, each separated by the isotype of the BCR, based on the IgH constant region expressed. Mutation frequencies were calculated using the SHazaM package in R and are relative to the germline sequence. One point per cell/BCR, coloured by BCR isotype. NA refers to cells/BCR sequences for which the IgH constant region could not be called. Statistical comparisons of group medians were carried out using a Kruskal-Wallis test.

Next, to determine whether the difference in mutation frequencies were reflective of the likelihood of B cells expressing certain BCR isotypes or subclasses to undergo clonal expansion, the total number of B cells, and number of expanded BCR clones per 100 cells were calculated and compared for each of the different BCR isotype and subclasses. This analysis showed that IgG1, IgG2, IgG3, IgG4, IgA1,

and IgA2, showed increased frequencies of expanded BCR clones, relative to IgD, and IgM (**Figure 5.12**).

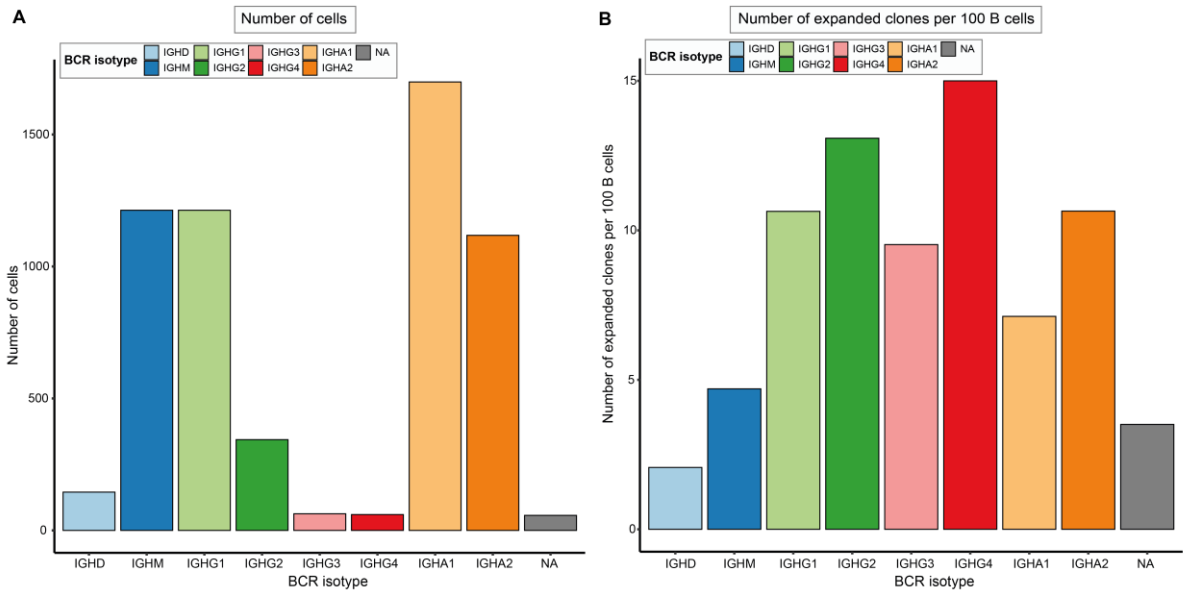


Figure 5.12. Intratumour B cells and expanded BCR clones are most frequently of the IgA or IgG1 isotypes. Bar plots showing the either (**A**) the number of B cells or (**B**) the number of expanded BCR clones per 100 cells, across all of the samples, separated by BCR isotype, based on the IgH constant region called. Bars are coloured by BCR isotype. NA refers to cells/BCR sequences for which the IgH constant region could not be called.

In addition to exploring signatures of GC responses for intra-tumour B cells in OAC, this dataset also provides a valuable opportunity to further investigate the association of increased IgA expression by circulating B cells of BO patients compared to OC patients, which was observed in the last chapter. To do this, the number of tissue infiltrating B cells and expanded BCR clones in each of the five major histological tissue types processed, based on the tissue of origin that the corresponding samples were collected from, were analysed.

Based on percentage of total B cells, this analysis showed that OSCC tissue was predominantly infiltrated by naïve B cells or plasma cells which expressed either IgM and IgG1, whilst OAC tissue was infiltrated predominantly by plasma cells expressing either IgG1, IgG2, IgA1, or IgA2 (**Figure 5.13A and B**). Analysis of the non-malignant tissue showed that oesophageal tissue was primarily infiltrated by non-plasma B cells which expressed IgM, IgG, or IgA, and BO and gastric tissue were almost exclusively

infiltrated by plasma cells expressing IgA1 or IgA2 (**Figure 5.13A and B**). Similar analysis based on the number of expanded BCR clones, showed that expanded BCR clones were frequent in OSCC tissue and were predominantly of the IgM or IgG1 isotype, whilst expanded BCR clones in OAC tissue were reasonably equally distributed amongst the IgM, IgG1, IgG2, IgA1, and IgA2 isotypes, with IgG1 and IgA1 the most frequent (**Figure 5.13C and D**). Analysis of the non-malignant tissue showed that both oesophageal and gastric tissue showed very low frequencies of expanded BCR clones, perhaps due to a lack of chronic inflammation and associated GC responses, whilst expanded BCR clones were frequent in BO tissue, and they were almost exclusively of the IgA1 or IgA2 isotype (**Figure 5.13A and B**).

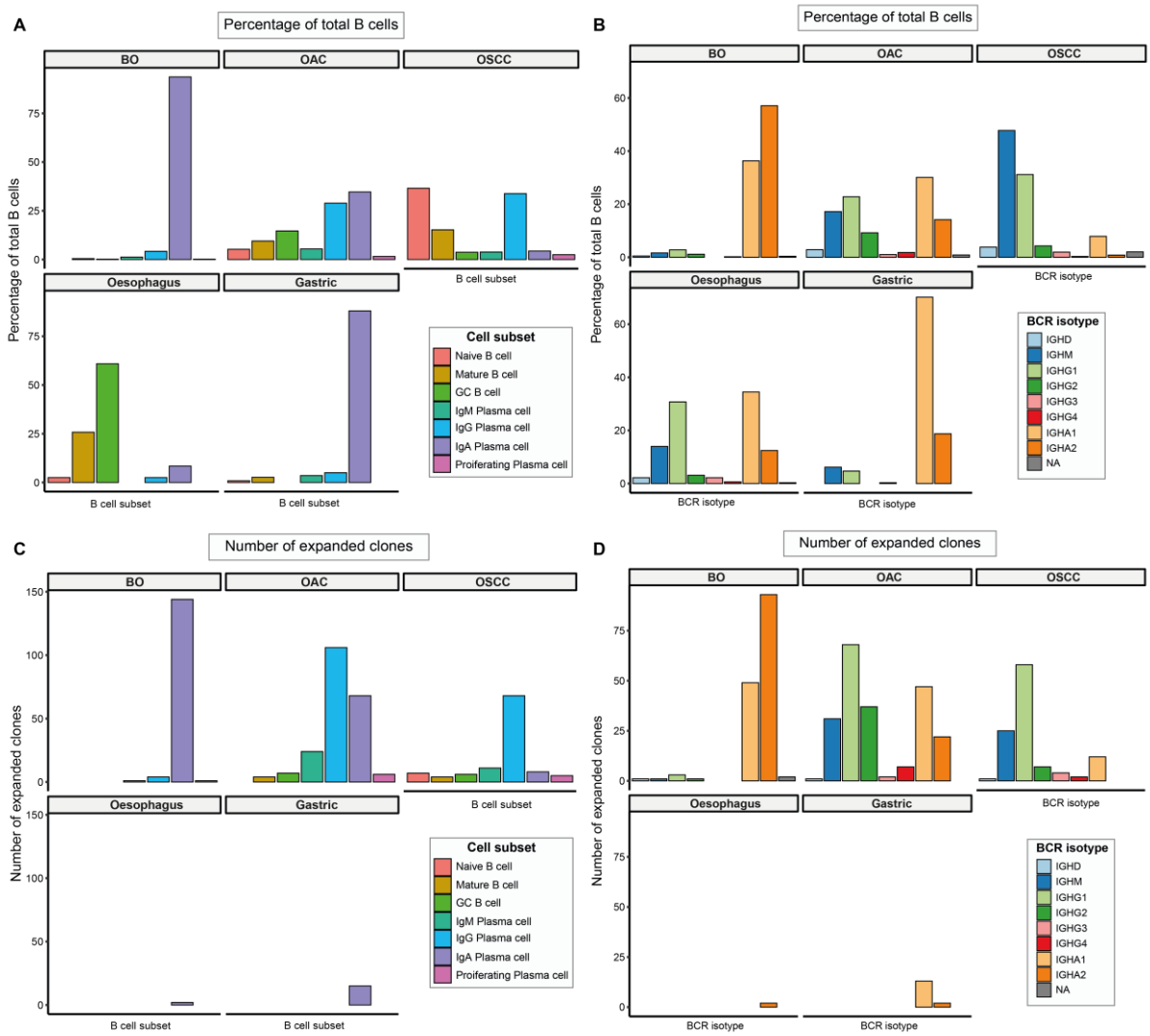


Figure 5.13. Different tissue types show different skew of B cell subsets and BCR isotypes. (**A-B**) Bar plots faceted by tissue type showing the percentage of B cells in the respective tissue which are either (**A**) of a certain B cell subset or (**B**) expressing a BCR of the respective isotype, based on the IgH constant region expressed. For each plot, bars are coloured by either B cell subset or BCR isotype expressed, respectfully. (**C-D**) Same as for (A-B), except showing the number of expanded BCR clones per tissue rather than percentage of B cells. For (**B** and **D**) NA refers to cells/BCR sequences for which the IgH constant region could not be called. BO = Barrett's oesophagus tissue, OAC = oesophageal adenocarcinoma tumour tissue, OSCC = oesophageal squamous cell carcinoma tumour tissue, Oesophagus = normal/non-malignant oesophageal tissue, Gastric = normal/non-malignant gastric tissue.

5.3.2. Selection of expanded BCR clones for in vitro screening

One of the aims of this chapter to determine whether BCRs expressed by expanded B cell clones bind to identifiable autoantigens. To address this, a subset of 34 of these expanded BCR clones, which were identified in the samples from the first four patients processed, were selected for expression into recombinant antibodies and experiment testing *in vitro* (**Figure 5.14**).

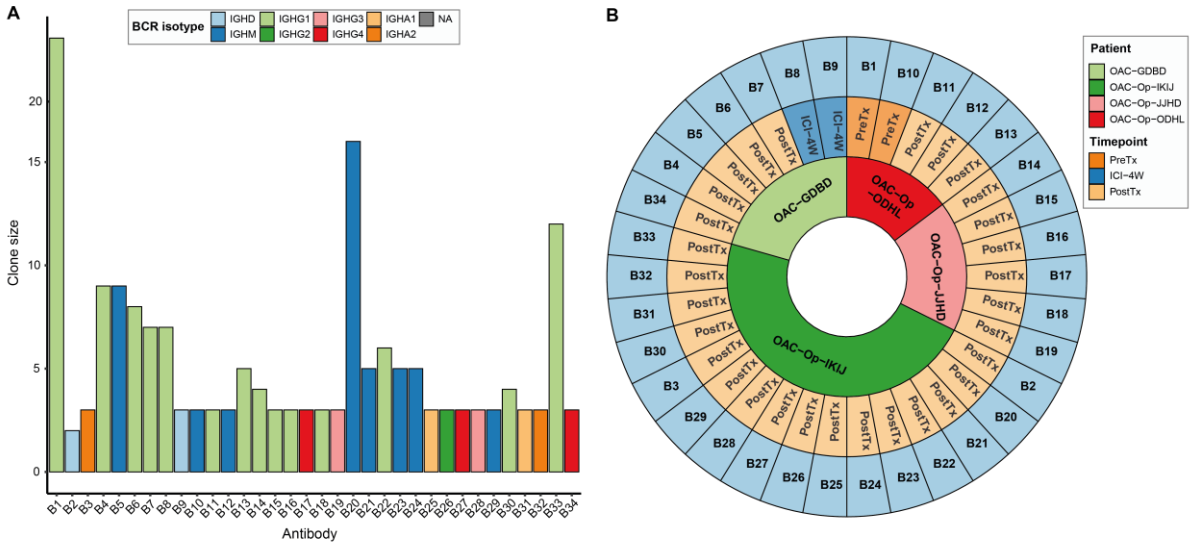


Figure 5.14. 34 expanded BCR clones were selected to be expressed as antibodies *in vitro.* From the samples from the first four patients processed, ~2000 high quality B cells with paired BCR data were sequenced and 34 expanded clones were identified (based on a clone size of 2 cells or greater). (**A**) Bar plot showing the clone size of each of the 34 expanded BCR clones identified. Bars are coloured by the BCR isotype, based on the VH constant region call. (**B**) Sunburst chart showing the patient and treatment timepoint from which each of the 34 expanded BCRs sequenced originated from. Segments are coloured either by the patient, or timepoint that the BCR sequence/antibody originated from.

5.4 Recombinant antibodies bind to human oesophageal adenocarcinoma tissue and cell lines

30 of the 34 expanded BCR clones that were selected were successfully cloned into expression vectors for IgH and Igκ or Igλ protein and subsequently expressed as recombinant IgG1 antibodies *in vitro* with the help of Dr Hantao Lou. Full details regarding cloning of the BCR sequences into these vectors and their subsequent expression by 293T cells are described in section 2.7.2. An example DNA gel demonstrating the successful cloning of a BCR insert corresponding to one of the expanded BCR clones into each of the IgH, Igκ, and Igλ expression vectors, as assessed by vector digestion, is shown below (**Figure 5.15**).

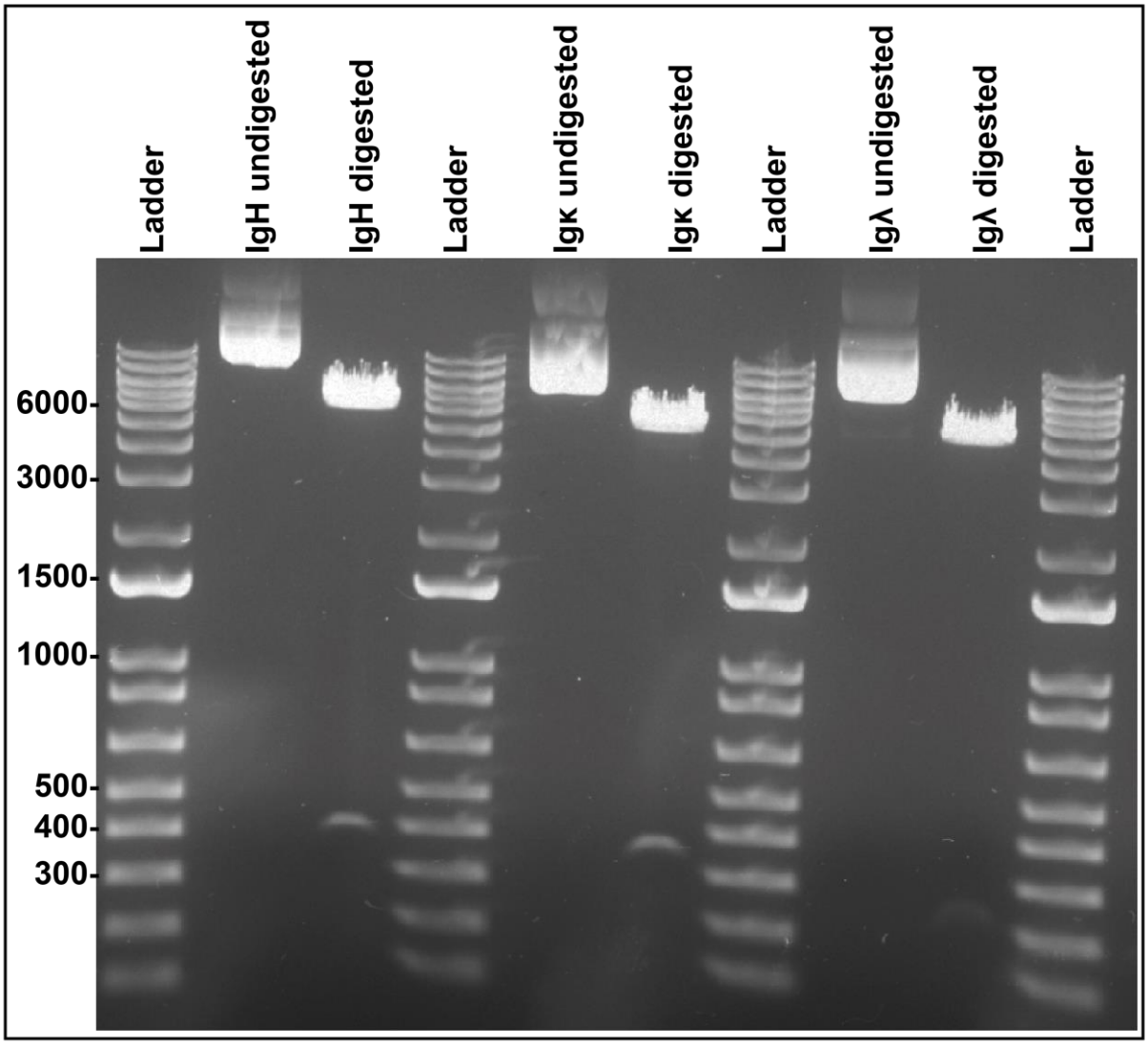


Figure 5.15. Example DNA gel of successfully cloned IgH, Igκ, and Igλ plasmids assessed by vector digestion. DNA gel showing the undigested plasmids for IgH, Igκ and Igλ contain the 250-400bp insert which corresponds to the appropriate V(D)J sequence and are visible on the gel when the plasmids are digested with the appropriate restriction enzymes. Ladder = 1kb plus DNA ladder (Invitrogen). The IgH and Igκ plasmid used for this had inserts corresponding to antibody B26, whilst the Igλ plasmid had an insert corresponding to antibody B31.

5.4.1 Recombinant antibodies bind to cancer cell lines

To assess tumour binding capability, the 30 recombinant antibodies which were successfully expressed were then screened by microscopy for binding to five different oesophageal or gastric cancer cell lines. Based on percentage of IgG-positive cells, nine of the 30 antibodies showed at least moderate binding (≥10% IgG-positive cells) to at least one of the five cell lines screened, and three of these antibodies, B11, B4, and B7, showed strong binding to more than one cell line (**Figure 5.16**).

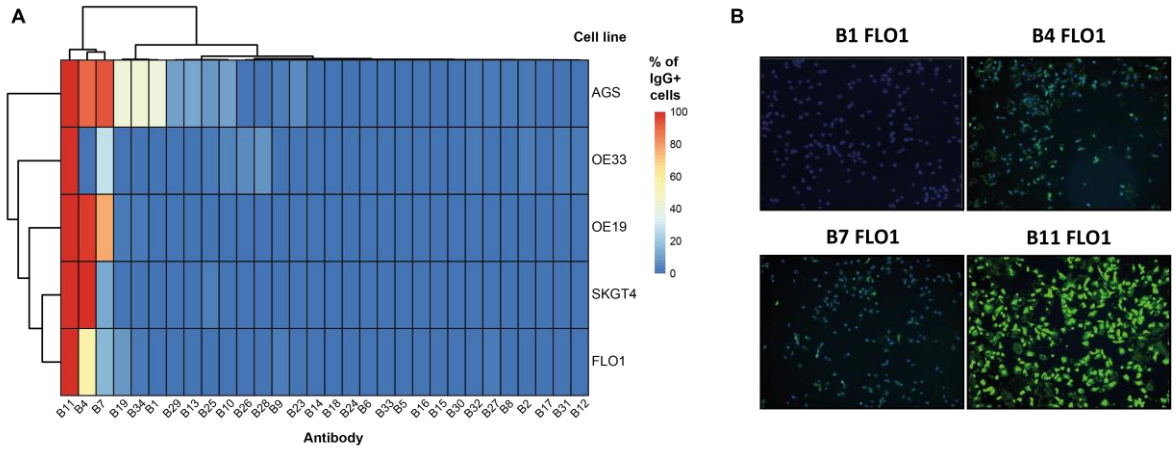


Figure 5.16. Some antibodies cloned from expanded intratumour B cells bind to gastro-oesophageal cancer cell lines. Five oesophageal or gastric cancer cell lines were fixed and incubated individually with each of the 30 antibodies cloned from expanded intratumour B cells and were imaged with an Operetta to assess antibody binding. (**A**) Heatmap of the percentage of IgG-positive cells for each of the five cell lines when stained individually with each of the 30 antibodies, with unsupervised hierarchical clustering of both antibodies (columns) and cell lines (rows) using the hclust method with "ward.D" linkage. (**B**) Examples of the microscopy images from which antibody binding was quantified which show FLO1 cells stained with DAPI (blue) and an anti-human IgG antibody (green), for one antibody which showed no binding (B1) and the three antibodies which showed the greatest binding (B4, B7, and B11).

As the cells were permeabilised prior to the microscopic screening above, the percentage of IgG staining quantified would encompass antibody binding to both intracellularly expressed and surface bound antigens. Hence, to determine whether any of the 30 antibodies were capable of binding to the surface of cells, flow cytometry was also used to assess extracellular binding of these same 30

antibodies to the FLO1 and OE33 cell lines. Six of the 30 antibodies tested showed some binding above background levels to at least one of FLO1 and OE33 (**Figure 5.17**). As with the microscopy-based screening, antibody B11 showed the highest levels of binding (**Figure 5.17C and D**).

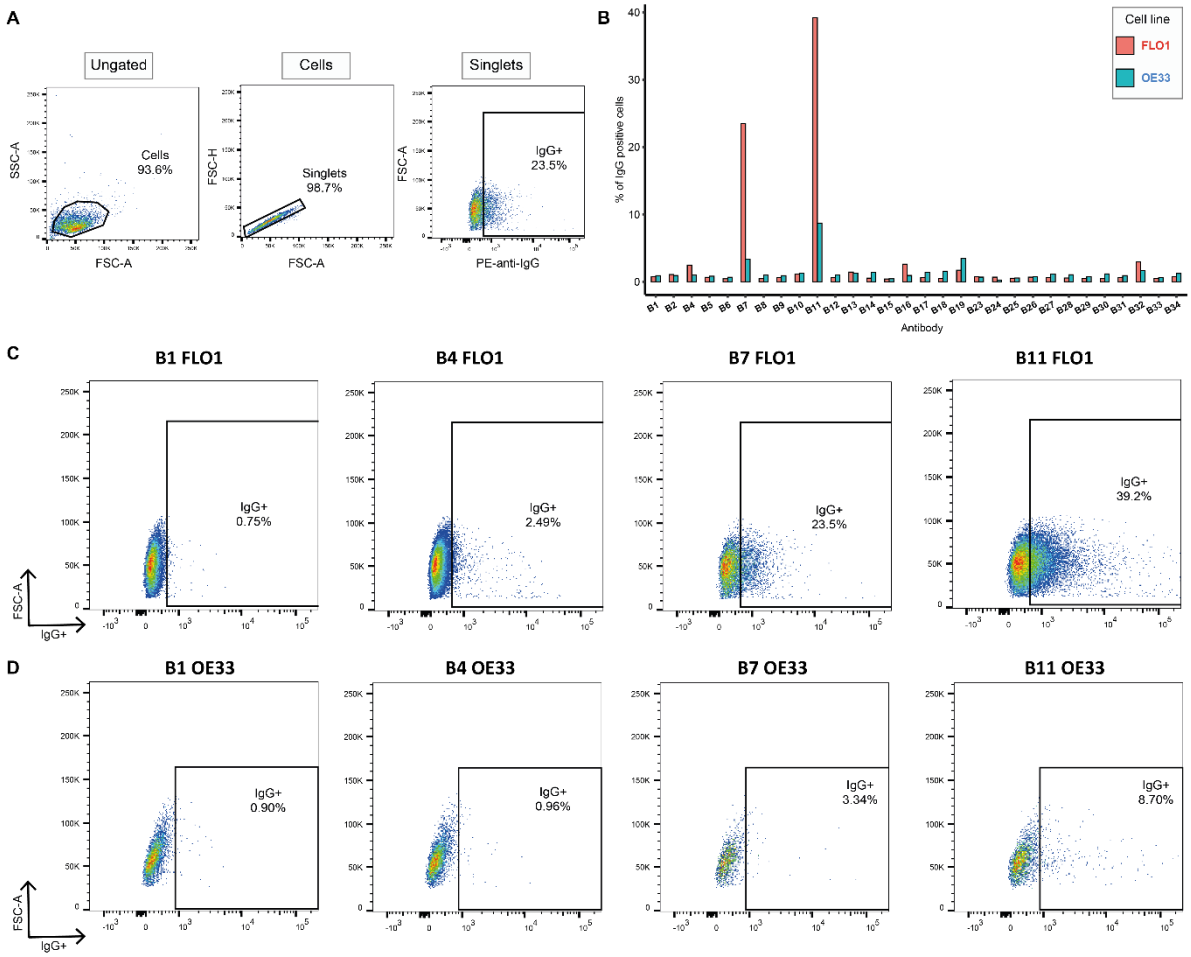


Figure 5.17. Some antibodies cloned from expanded intratumour B cells bind to the surface of oesophageal cancer cell lines. Two oesophageal cancer cell lines, FLO1 and OE33, were incubated individually with each of the 30 antibodies cloned from expanded intratumour B cells and antibody binding to the surface of the cell was assessed by flow cytometry. (**A**) An example gating strategy which was used for the flow cytometry analysis performed using FlowJo™ Software. Cells were gated based on FSC-A vs SSC-A profile, then singlets were gated based on FSC-H vs FSC-A, and finally IgG+ singlets were gated based on IgG+ staining. This gate was set using samples stained with only the secondary anti-human IgG antibody. (**B**) Bar plot of the percentage of IgG-positive cells for FLO1 (red) and OE33 (blue) cells when stained individually with each of the 30 antibodies. (**C-D**) Example IgG staining plots of singlets for either (**C**) FLO1, or (**D**) OE33 cells incubated with the antibodies B1, B4, B7, and B11.

5.4.2 Recombinant antibodies preferentially bind to annexin V-positive cancer cell lines and promote their phagocytosis by myeloid cells

Given the observed stronger binding to permeabilised cells, and previous reports of preferential antibody binding to apoptotic cells[499–501], including from a recent study of RCC which showed binding of locally produced IgG to apoptotic tumour cells *in vivo*[109], the possibility that some of these 30 antibodies bind to intracellular autoantigens which are expressed on the cell membrane during apoptosis, or other forms of cell death, was explored. To do this, nine of the 11 antibodies which showed binding to at least one cancer cell line by either microscopy or flow cytometry were taken forward to see whether they bound preferentially to dying cells, assessed by positive annexin V staining. The two antibodies which were not taken forward, B10 and B19, were excluded because there was insufficient antibody available. However, additional amounts of these antibodies should be synthesised in the future, and they should be tested for binding to annexin V-positive cells.

Cell death was induced in FLO1 cells through treatment for 48 hours with the chemotherapeutic drug cisplatin. Following treatment, binding of each of the nine different antibodies to annexin V-positive and annexin V-negative FLO1 cells was assessed by flow cytometry. Each of the nine antibodies tested showed increased binding to annexin V-positive cells relative to annexin V-negative cells (**Figure 5.18**).

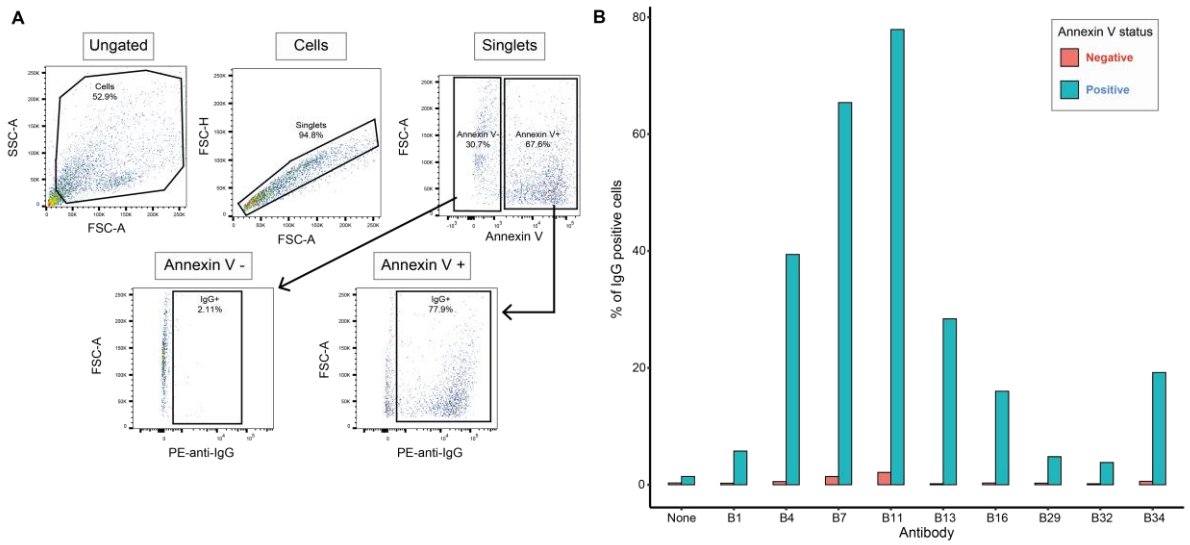


Figure 5.18. Some antibodies cloned from expanded intratumour B cells bind to the surface of FLO1 cells. FLO1 were treated with cisplatin for 48 hours to induce cell death, and then incubated individually with each of the 30 antibodies cloned from expanded intratumour B cells and antibody binding to the surface of the cell was assessed by flow cytometry. (**A**) An example gating strategy which was used for the flow cytometry analysis performed using FlowJo™ Software. Cells were gated based on FSC-A vs SSC-A profile, then singlets were gated based on FSC-H vs FSC-A, annexin V-negative (-) and annexin V-positive (+) cells were gated based on annexin V staining, and finally, for both populations, IgG+ cells were gated based on IgG+ staining. This gate was set using samples stained with only the secondary anti-human IgG antibody. (**B**) Bar plot of the percentage of IgG-positive cells for annexin V-negative (red) and annexin V-positive (blue) cells when stained individually with each of the 9 antibodies, or in the absence of any antibody.

In order to further study the functional consequences of these antibodies binding to annexin V-positive cells without the confounding factors of molecules that may have been present in the 293T culture media that these antibodies were expressed, column purification into purified IgG protein for each of these nine antibodies which showed binding to annexin V-positive cells was attempted. Six of these nine antibodies were successfully purified by column purification (data not shown). However, unfortunately minimal antibody was purified for B16, B29 and B32 (data not shown). An additional antibody, B2, which did not show binding to any of the cell lines during the flow cytometry and microscopy screening, was also successfully purified to serve as a negative control. Full details regarding column purification of antibodies are described in section 2.7.3.

Given the demonstrated role of antibody binding in phagocytic clearance of apoptotic cells[502–504], following successful purification of six of the antibodies, the next step was to determine whether

antibody binding to annexin V-positive cells or their debris would influence their clearance by myeloid cells.

For the six antibodies which showed binding to annexin V-positive cells that were successfully purified, and the negative control antibody B2, a flow cytometry based *in vitro* ADCP assay involving co-culture of cisplatin treated FLO1 cells with primary CD14+ monocytes was performed. This assay showed that the presence of either the B1 or B11 antibody enhanced phagocytosis of cisplatin treated FLO1 cells by CD14+ monocytes compared to the presence of any of the other five antibodies, or monocytes and cisplatin treated FLO1 cells co-cultured in the absence of any antibody (**Figure 5.19**).

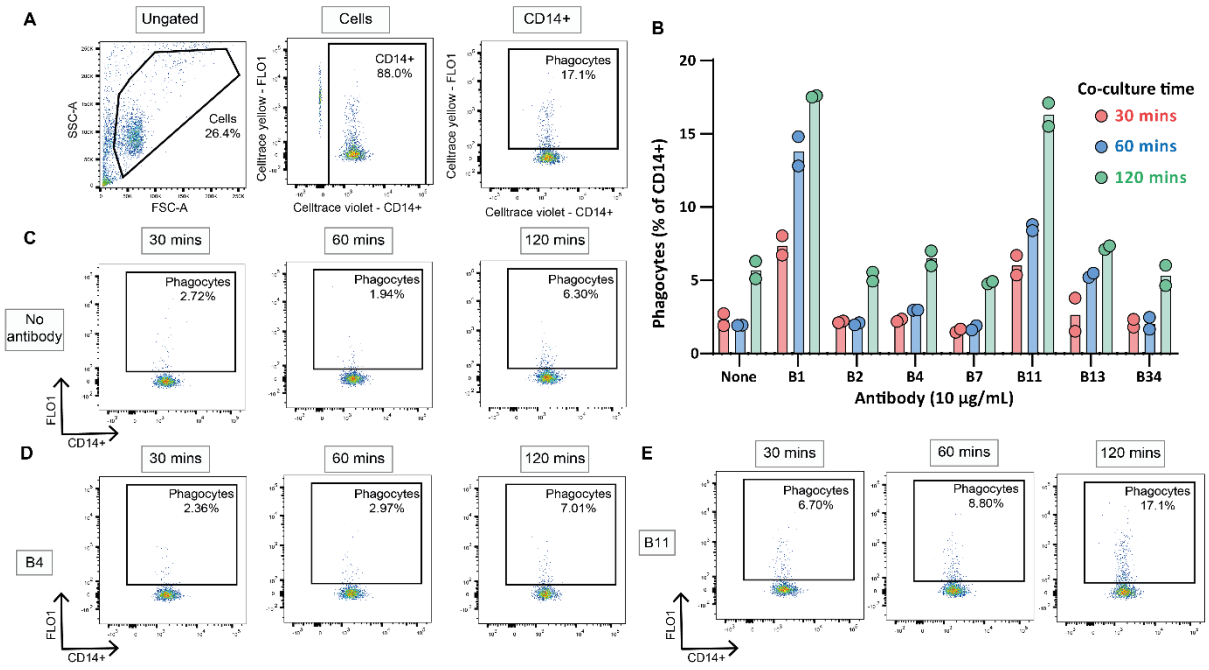


Figure 5.19. The antibodies B1 and B11 enhanced phagocytosis of cisplatin treated FLO1 cells by primary CD14+ monocytes. FLO1 were treated for 48 hours with cisplatin and then labelled with the dye celltrace yellow. These celltrace yellow-positive cells were then co-cultured with celltrace violet labelled CD14+ monocytes magnetically isolated from PBMCs, in the presence or absence of 10µg/ml of purified IgG antibody. Cells were then harvested and fixed after 30 minutes, 60 minutes, or 120 minutes of co-culture and analysed by flow cytometry for celltrace violet and celltrace yellow positive cells, as a readout for phagocytosis. (**A**) An example gating strategy which was used for the flow cytometry analysis performed using FlowJo™ Software. Cells which were likely to be monocytes were gated based on FSC-A vs SSC-A profile, then CD14+ cells were gated based on celltrace violet staining, and finally phagocytes were gated based on celltrace yellow staining. This gate was set using samples of celltrace violet labelled monocytes cultured alone. (**B**) Bar plot of the percentage of IgG-positive cells for FLO1 (red) and OE33 (blue) cells when stained individually with each of the 30 antibodies. (**C-E**) Example plots showing the percentage of phagocytes after 30 minutes, 60 minutes, or 120 minutes of co-culture for either (**C**) monocytes and FLO1 cells co-cultured in the absence of antibody (No antibody), (**D**) monocytes and FLO1

cells co-cultured in the presence of the antibody B4, or (**E**) monocytes and FLO1 cells co-cultured in the presence of the antibody B11.

5.5 Antigen characterisation of recombinant antibodies

One of the aims of this chapter was to explore whether expanded BCR clones from tumours bind to identifiable autoantigens. As the results described in section 5.4 suggested that the target antigens for the cancer cell line reactive antibodies were likely to be expressed intracellularly, all 30 antibodies that were successfully expressed were screened for reactivity to the 33 autoimmune associated antigens present on the RayBiotech AAb array. From antigen screening using this array, three antibodies showed reactivity to one or more antigens (**Figure 5.20**). Antibody B1 showed clear binding to the keratin 20 antigen (**Figure 5.20B**), whilst antibody B11, when diluted 10-fold, showed clear binding to the Scl-70 antigen, which has been widely demonstrated to be topoisomerase I (TOP1)[505,506], and the ribonuclear protein U1-snRNP70 kDa, in addition to weak binding to the CCP peptide antigen (**Figure 5.20F**). Unfortunately, antibody B13 showed a lot of background binding when undiluted so it was unclear which antigens in particular it might bind to, and this will need to be reassessed on this antigen array at a greater dilution (**Figure 5.20D**).

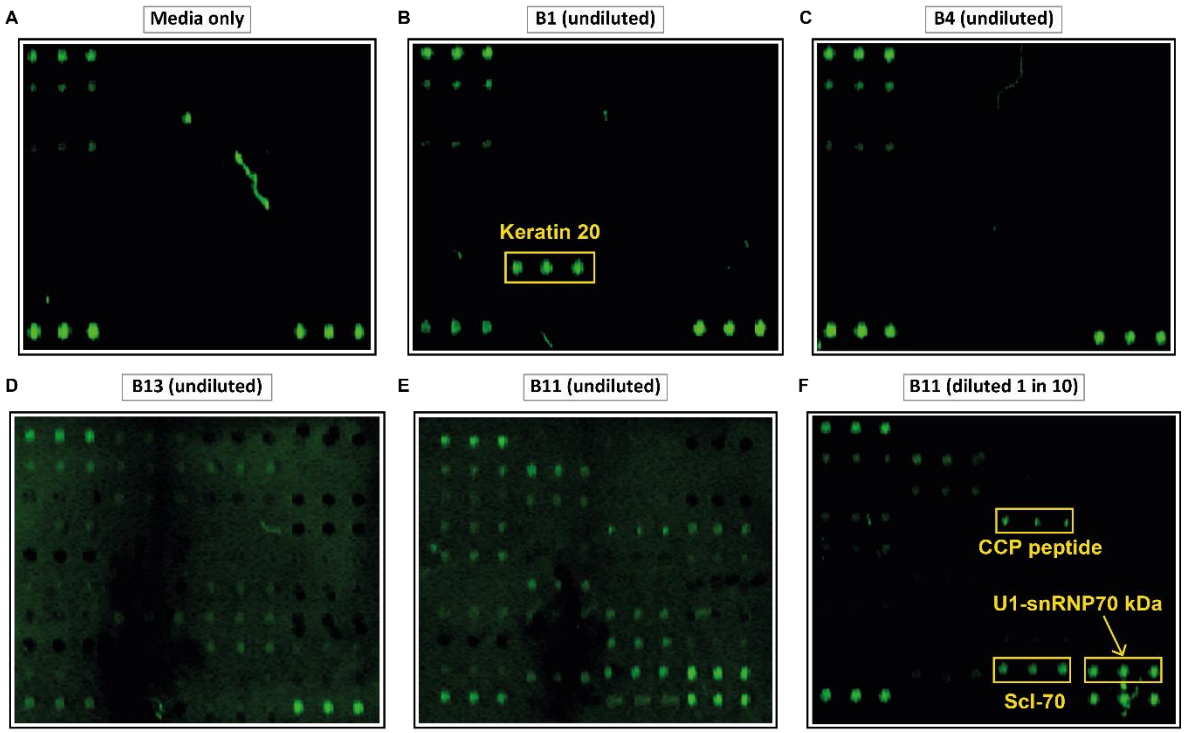


Figure 5.20. Three antibodies showed binding to autoimmune disease associated antigens. Each of the 30 antibodies synthesised, and the culture media that they were synthesised in, were screened for reactivity to the 33 autoantigens present on the RayBiotech AAb array. (**A-F**) Array images for (**A**) media only, (**B**) antibody B1 undiluted, (**C**) antibody B4 undiluted, (**D**) antibody B13 undiluted, (**E**) antibody B11 undiluted, and (**F**) antibody B1 diluted 1 in 10 with media. Clear signals corresponding to a specific autoantigen are highlighted by yellow boxes.

To further elucidate the antigens to which these antibodies bind, the seven antibodies that were successfully purified into IgG were included in the Luminex based protein fragment panel described in section 3.4. Analysis of the subsequent data showed that three of the antibodies were reactive to one or more protein fragments (**Figure 5.21**). Interestingly, one of the antigens for which antibody B11 showed binary reactivity to was a protein fragment corresponding to TOP1 (**Figure 5.21A**). In the Luminex based antigen panel there were three protein fragments spanning the TOP1 sequence, but only one fragment, corresponding to ammino acids 28-78, showed a bimodal like distribution when FMM was performed, as described in section 3.2 of chapter 3, and so only binary serostatus data was calculated for this fragment. However, analysis of the net MFI data for all three of the protein fragments corresponding to TOP1, for each of the seven antibodies screened, showed that antibody B11 bound with similar strength to the two fragments which corresponded to either amino acids 3 to

77, or 28 to 78 of TOP1, but showed minimal binding to the fragment corresponding to amino acids 158 to 243 of TOP1 (**Figure 5.21B**).

In addition to TOP1, antibody B11 also showed binary reactivity to a protein fragment corresponding to KCTD21 and TAMM41. However, subsequent alignment of the reactive protein fragments corresponding to TOP1, KCTD21, and TAMM41 did not reveal any shared epitope based on amino acid sequence, although a shared secondary structure between the three antigens cannot be ruled out. It should also be noted that antibody B7 and antibody B13 also showed binary reactivity to TAMM41, although the other four antibodies did not, so binding to this antigen could be the consequence of common polyreactivity, although this would need to be specifically tested.

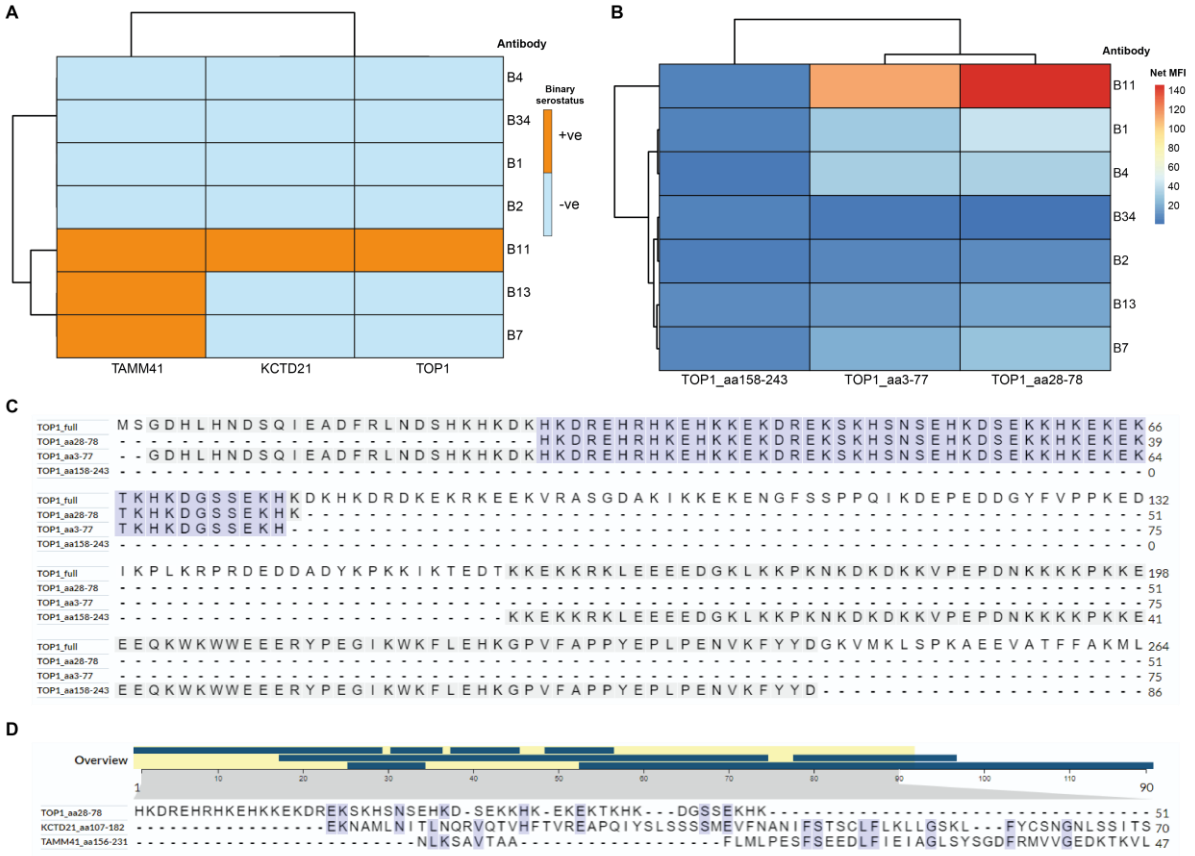


Figure 5.21. Antibody B11 binds to three antigens including topoisomerase I. (**A**) Heatmap of the binary serostatus data of the TAMM41, KCTD21, and TOP1 protein fragments from the seven antibodies which were purified and included in the samples run on antigen panel 2. (**B**) Heatmap of the net MFI data for the seven antibodies, which were purified and included in the samples run on antigen panel 2, against the three protein fragments which correspond to the TOP1 gene. aa refers to the amino acids of TOP1 to which the corresponding protein fragment relates to. For both heatmaps, unsupervised hierarchical clustering of both antigens (columns)

and samples (rows) using the hclust method with "ward.D" linkage was performed.(**C**-**D**). UniProt alignment result comparing the amino acid sequences of either (**C**) the full sequence of topoisomerase I (TOP1_full), the protein fragment corresponding to TOP1 amino acids 28 to 78 (TOP1_aa28-78), the protein fragment corresponding to TOP1 amino acids 3 to 77 (TOP1_aa3-77), and the protein fragment corresponding to TOP1 amino acids 158 to 243 (TOP1_aa158-243), or (**D**) the three protein fragments which antibody B11 had a positive binary serostatus for, which were the protein fragment corresponding to TOP1 amino acids 28 to 78 (TOP1_aa28-78), the protein fragment corresponding to KCTD21 amino acids 107 to 182 (KCTD21_aa107-182), and the protein fragment corresponding to TAMM41 amino acids 156 to 231 (TAMM41_aa156-231). For (**C** and **D**), exact matches across all but one of the sequences are shown in lilac.

Interestingly, B11 did not show a positive binary serostatus to either of the two protein fragments corresponding to snRNP70 (data not shown), despite showing binding to snRNP70 on the RayBiotech AAb array.

5.6 Discussion

One of the hypotheses of this thesis was that the abundance and transcriptional profiles of intratumour B cells influences the clinical outcome of ICI treated OC patients, by perhaps acting as a marker of a highly immunogenic TME.

Whilst transcriptomic measurements of intra-tumour TLSs and B cells did not show significant associations with clinical outcome of ICI treated OAC patients, analysis of a scRNA-seq dataset, which was generated from samples collected from the same patients, showed clear evidence of mature intra-tumour B cells which had undergone GC responses, suggesting that antigen driven GC- or TLS-mediated maturation of B cells is a common occurrence in OAC tumours. Further, expression of a subset of the expanded BCR clones, which were identified from tumours, as recombinant antibodies and subsequent *in vitro* screening showed that tumour reactive ASCs are present in OC tumours *in vivo* and the antibodies they produce can bind to identifiable autoantigens and enhance ADCP of apoptotic cancer cell lines.

5.6.1 Transcriptomic measurements of intra-tumour tertiary lymphoid structures and B cells

Transcriptomics based measurement of intra-tumour TLSs abundance for each of the PreTx OAC LUD2015-005 tumour samples was performed using two previously published TLS gene signatures[107,114], but, in contrast to previous studies of other ICI treated cancer types[106–108,464], no significant association with patient clinical outcome was observed for either signature.

Similarly, once again in contrast to previous studies of other ICI treated cancer types[106–108,464], estimates of intra-tumour B cell abundance in the same PreTx OAC samples using two different RNA deconvolution approaches, showed no significant association with patient clinical outcome. Further analysis regarding the intra-tumour abundance of specific subsets of B cells also showed no significant association with patient clinical outcome.

This difference in observations, regarding the prognostic significance of intra-tumour abundance of TLSs and B cells, between OAC and the other ICI treated cancer patients could possibly be due to differences in cancer type or the treatment received, given that the other studies were of melanoma, sarcoma, and RCC patients treated with ICI without concurrent CTX.

Another important consideration is that, as mentioned previously, intratumour TLSs and B cells are best explored histologically. Hence, the observed of lack of association of transcriptional based measurements of intra-tumour TLS and B cell abundance with good clinical outcome should by validated by immunohistochemistry or immunofluorescence in a larger independent cohort of ICI treated OAC.

To further explore the role of intra-tumour B cells in ICI treated OAC, clonality and diversity analysis of bulk BCR sequencing data of tumour derived RNA is planned and will be used to determine whether CB OAC patients from the LUD2015-005 trial have increased BCR clonality and diversity relative to NCB patients, as was reported for ICI treated melanoma[108]. An alternative approach would be to use the MiXCR[507] package or modified TRUST[508] algorithm, which can both identify BCR IgL and IgH transcripts from bulk RNA-seq data that has not been specifically amplified for Ig transcripts and have previously been used to analyse clonality and diversity of intratumour B cells[108,109], on the existing bulk tumour RNA-seq data that was analysed in this chapter. However, unlike the targeted approach planned, using either MiXCR or the modified TRUST algorithm would not provide isotype data for the Ig transcripts, and would identify substantially fewer Ig transcripts per sample.

5.6.2 Evaluating clonal expansion and subtype and isotype distribution of tissue infiltrating B cells

All stages of B cell maturation from naïve B cells to plasma cells were identified in OC tumour samples by scRNA-seq, and subsequent analysis of a combination of GEX and BCR sequencing data for tissue

infiltrating B cells showed clear evidence of GC responses. Indeed, many GC associated processes, including clonal selection, clonal amplification, SHM, isotype switching, and plasma cell differentiation were supported by measurements of large BCR clone sizes, high SHM load in the Vh segment and CDR3 regions, BCR isotype usage, and cell transcriptomic profile. Together, these findings suggest that full B cell responses were generated, likely in response to chronic antigen exposure, in OC tumours.

Further analysis of BCR sequences revealed differences in the mutation frequencies of both the Vh segment and CDR3 region across both different B cell subset and BCR isotypes. Unsurprisingly, IgA and IgG plasma cells, which are the most terminally differentiated of the B cell subsets identified, showed the highest average mutation frequencies and contribution to expanded BCR clones, whilst naïve B cells showed the lowest average mutation frequency and contribution to expanded BCR clones. Interestingly, IgG3 showed substantially lower frequencies of mutations in the CDR3 region compared to the other IgG and IgA subclasses, although this largely agrees with what has been previously reported in a study of PBMCs from healthy individuals[509].

Strikingly, scRNA-seq based deconvolution of bulk RNA-seq data from PreTx OAC samples, revealed that, for most samples, intra-tumour B cells were predominantly IgA or IgG plasma cells. This is, to my knowledge, the first report of OAC tumours showing a plasma cell rich infiltrate in the majority of OAC tumours, although in OSCC, OAC, and GAC, immunohistochemistry staining of plasma cell markers showed examples of high plasma cell infiltration in a number of samples[510,511]. However, these studies did not quantify plasma cell infiltration relative to the infiltration of other B cell subsets.

Hence, it is possible that dominance of the tumour infiltrating B cell compartment by plasma cells could be a feature of OAC tumour tissue. Indeed, this is supported by the reports of tissues from other inflammatory conditions of the oesophagus, BO and EO, are heavily infiltrated by immune cells involved in the humoral immune response or show increased abundance of Ig protein[480,481] . It would

also be interesting to see if binding of endogenously produced IgG or IgA to tumour cells is a common feature of OAC tumours, or whether the degree of binding could influence clinical outcome to ICI treatment, as has been suggested for RCC[109].

Another interesting observation from the scRNA-seq analysis was a clear dominance of IgA plasma cells within the tissue infiltrating B cell compartment of BO tissue. This observation is particularly interesting given the increased expression of IgA by peripheral B cells of BO patients relative to OC patients that was shown in the previous chapter and raises the question of whether the increased expression of IgA by peripheral B cells in BO patients is a consequence of increased abundance of tissue infiltrating IgA expressing B cells and their subsequent migration into the peripheral blood. Although, it should be noted that this finding contrasts with the previously reported dominance IgE and IgG, and not IgA, in BO tissue examined by immunohistochemistry[480,481]. Despite this, interestingly, studies of breast cancer and CRC have reported a similar finding[512,513]. Both of these studies showed that the isotype distribution of plasma cells was skewed towards IgG and away from IgA in tumour tissue, and in the opposite direction in adjacent normal tissue[512,513].

However, given the small numbers of samples analysed by scRNA-seq in this chapter, these findings should be validated in an independent cohort with a greater number of samples, including those from EO, BO, and OC patients, and preferably using multiple approaches such as scRNA-seq, deconvolution of bulk RNA-seq data, or quantification by flow cytometry, or immunohistochemistry.

5.6.3 Tumour binding capability, function, and antigen reactivity of antibodies expressed by expanded BCR clones identified from OC tumours

In vitro assays, utilising gastric and oesophageal cancer cell lines, were performed to analyse the tumour cell binding capabilities, function, and antigen reactivity of a subset of 30 BCR clones which were expanded in OC tumour samples. From these assays, 11 antibodies showed intracellular or

extracellular binding to at least one of the five cell lines tested. Further testing of nine of these antibodies showed that they all bound preferentially to annexin V-positive cells rather than annexin V-negative cells, and two of the antibodies were found to promote phagocytosis of annexin V-positive cells by primary human monocytes. Due to limited time, most assays were only performed once and did not include all five cancer cell lines and/or all 30 recombinant antibodies. Hence, comprehensive characterisation of binding to additional cancer cell lines and primary tumour tissue, ideally by multiple experiments, should be performed to validate these preliminary findings.

Nonetheless, these findings are of great interest given that IgG binding to apoptotic tumour cells was observed in a recent study of RCC[109]. Interestingly, the authors of this study suggested that this could be reflective of direct anti-tumour effects[109]. Although *in vitro* studies have shown that various mAbs can induce apoptosis of cell lines[514], this seems unlikely for B11 and the other recombinant antibodies screened in this chapter given the lack of binding to annexin V-negative cells. Instead, it seems more probable that the antigens for these antibodies are intracellular and are only exposed by dying cells, like has been shown for laminin-binding protein[500]. Nevertheless, the results of experiments to determine whether B11 or any of the other tumour binding recombinant antibodies could promote apoptosis of cancer cell lines would be interesting.

The study of RCC mentioned above also showed increased tumour infiltration of macrophages and their closer proximity to tumour cells in tumours which exhibited high tumour cell IgG binding[109]. This is particularly interesting given the enhanced ADCP of annexin V-positive cells FLO1 cells by primary monocytes in the presence of antibodies B1 and B11. Given that ADCP has been shown to promote further immune activation either by presenting peptides from the engulfed cell as antigen to T cells via MHC class II, or by producing pro-inflammatory mediators[250,274], it is possible that these antibodies which promote ADCP could be enhancing anti-tumour immunity. However, it is important to note that it is widely believed that phagocytosis of apoptotic cells by myeloid cells can lead to induction of anti-

inflammatory signalling[515–519]. This should be addressed by further experiments to determine whether the increased ADCP of annexin V-positive cells by antibodies B1 and B11 has any influence on the expression profile, differentiation, or functional capacity of the phagocytic monocytes.

Given the recent finding that tumour reactive antibodies can originate from tumour reactive or non-reactive precursors[288], antibody B11, and other cell line binding antibodies, should be expressed as their corresponding germline sequence to determine whether they acquired binding capabilities due to SHM or this binding was germline encoded. Similarly, for ease of high throughput expression and screening, all 30 antibodies were initially cloned as IgG1, but for functional testing they should be cloned as their native isotype as this could influence their ability to promote ADCP[520,521].

Three antibodies, including the two which enhanced ADCP of annexin V-positive cells, were reactive to autoantigens included on the RayBiotech AAb array and/or the custom Luminex antigen array which was described in section 3.2 of chapter 3. Antibody B11 showed binding to multiple antigens from these arrays, including the extractable nuclear antigens TOP1 and U1-snRNP70, which have both been suggested to undergo oxidative modification during apoptosis that could generate immunogenic neoepitopes[522]. If true, this could help to explain why this antibody was clonally selected in the tumour, with a possible hypothesis being that excessive apoptosis in the inflammatory TME led to chronic exposure to modified TOP1 and U1snRNP70 antigens. It should also be noted that a significant association between the presence of anti-U1snRNP antibodies and anti-CMV antibodies has been reported[523] and so the possibility of antibody B11 binding to U1-snRNP70 due to cross-reactivity with CMV cannot be ruled out.

Nonetheless, it is interesting that TOP1 was the only antigen which was included on both antigen arrays and for which antibody B11 showed reactivity, especially given that anti-TOP1 antibodies are clinically used to diagnose patients with the autoimmune condition systemic sclerosis (SSc)[524], and the

various suggestions that SSc patients are at increased risk of developing cancer[525,526]. Patients with diagnosed autoimmune diseases, including SSc, were excluded from enrolment in the LUD2015-005 study, so it is unlikely, although not impossible, that the patient from which antibody B11 was expressed had generated anti-TOP1 antibodies independently of the tumour. In addition to the previously discussed reversion of the B11 BCR to the corresponding germline sequence, the aetiology of antibody B11 should be explored further by additional BCR repertoire analysis of the intra-tumour B cells of this patient, with a focus on identifying clones that belong to the same clonal family. Additionally, to confirm binding to TOP1 and identify other potential reactive antigens, immunoprecipitation, using antibody B11 and any related clones, and subsequent mass spectrometry should be performed.

Finally, as only a subset of 34 of the 491 expanded BCR clones which were identified in this chapter were expressed as recombinant antibodies and screened *in vitro*, further expanded BCR clones should be expressed and screened in a similar manner to confirm whether tumour reactive antibodies are commonly expressed by tumour infiltrating B cells.

5.7 Conclusion

Analysis of the abundance of intra-tumour TLSs and B cells using transcriptomic approaches did not reveal significant associations with clinical outcome of ICI+CTX treated OC patients. However, analysis of scRNA-seq data showed clear evidence of GC responses, suggesting that full B cell responses were generated in the OC tumours analysed. Expression and *in vitro* testing of expanded BCR clones identified from OC tumours showed that tumour reactive ASCs are present in OC tumours *in vivo*. Further functional assays showed that the antibodies produced by some of these ASCs bind to known autoantigens and they may function to modulate anti-tumour through antibody-mediated phagocytosis of dying tumour cells.

6 Conclusion and future directions

6.1 Overall conclusions

ICI therapy has revolutionised treatment of many cancer types and was recently approved by the FDA to treat advanced gastro-oesophageal cancer[27–30,32,39]. However, there are still many patients who do not show good clinical outcome to ICI therapy[30–33,36,43]. Further, although the development of irAEs during treatment has been linked to good clinical outcome, there are often several patients who develop irAEs and do not show good clinical outcome, or patients who must cease therapy due to development of serious irAEs[139,140,143–145].

Understanding the heterogeneity of patient clinical outcome and risk of irAE development is limited, particularly in OC. Studies of other cancer types have suggested that clinical outcome and irAE development during ICI therapy are influenced by the patient's baseline immune activation and the presence of intra-tumour B cells[42,106–108,120,125–127,129–136,146,147,275–280]. However, similar studies to determine whether this is the case in ICI treated OC have not yet been reported. Hence, the aim of this thesis was to investigate the role of baseline immune activation and tumour infiltrating B cells in the clinical outcome of OC patients to ICI+CTX treatment.

6.1.1 Signatures of peripheral blood CD8 T cells and B cells, but not autoantibodies, reflect likelihood of clinical benefit and irAE development

In chapter 3, I performed comprehensive AAb profiling using two custom antigen arrays to investigate whether the number of reactive antigens could be used as a surrogate measurement for baseline immune activation. Although no significant association of IgG AAb reactivity with either patient clinical outcome or irAE development was found, similar analysis of LAC patients treated with 1st line ICI showed a positive association with number of reactive antigens and good tumour irRECIST responses, suggesting that the number of reactive antigens may be a more pertinent predictor of good clinical outcome in patients treated with ICI in the absence of concurrent CTX.

To further characterise the baseline immune activation of OAC patients of the LUD2015-005 inoperable cohort, in chapter 4, comprehensive *ex vivo* immunophenotyping of PBMCs was performed to identify peripheral immune signatures which were associated with increased likelihood of good clinical outcome and irAE development. The predominant signature associated with good clinical outcome was the increased activation and proliferation of cytotoxic CD8 T cells. In contrast, the predominant signatures associated with irAE development were a skew towards a more activated expression profile in the B cell compartment, and reduced TGF-β-mediated immunosuppression (**Figure 6.1**).

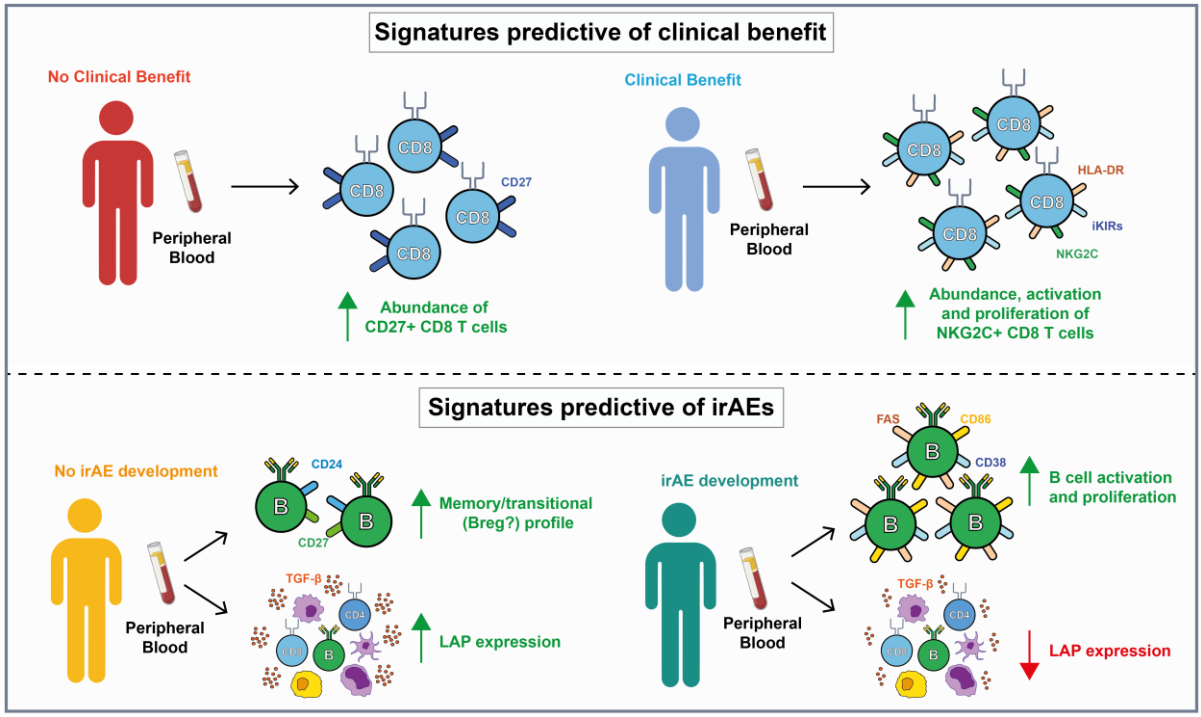


Figure 6.1. Peripheral signatures predictive of clinical benefit and immune related adverse events in the LUD2015-005 inoperable cohort. Comprehensive *ex vivo* immunophenotyping identified unique signatures detectable prior to ICI treatment which were associated with clinical benefit (CB) and immune related adverse events (irAEs). (Top) Differences in the abundance and expression profiles of the cytotoxic CD8 compartment were the predominant predictive signature for likelihood of CB. (Bottom) Differences in the activation and expression profiles of the B cell compartment, and differences in latency-associated peptide (LAP) expression by PBMCs were the predominant predictive signatures for likelihood of irAE development.

In conclusion, although AAbs did not show significant prognostic potential in the LUD2015-005 inoperable cohort, signatures derived from PBMCs expand our understanding of mechanisms that

predispose to good clinical outcome or irAE development during ICI treatment. The validation of these results in larger independent cohorts could inform patient treatment decisions based on biomarkers predictive of clinical outcome.

6.1.2 Tumour infiltrating B cells undergo germinal centre responses and, in some cases, express tumour-reactive antibodies

In chapter 5, I used transcriptomic approaches to investigate the role of intra-tumour TLSs and B cells in the clinical outcome of OC patients to ICI+CTX treatment. Neither the abundance of intra-tumour TLSs nor B cells showed significant associations with clinical outcome. However, analysis of scRNA-seq data showed clear evidence of GC responses, suggesting that full B cell responses were generated in the OC tumours analysed. Expression of expanded BCR clones identified from OC tumours as recombinant antibodies and subsequent *in vitro* testing showed that tumour reactive ASCs are present in OC tumours *in vivo* (**Figure 6.2**).

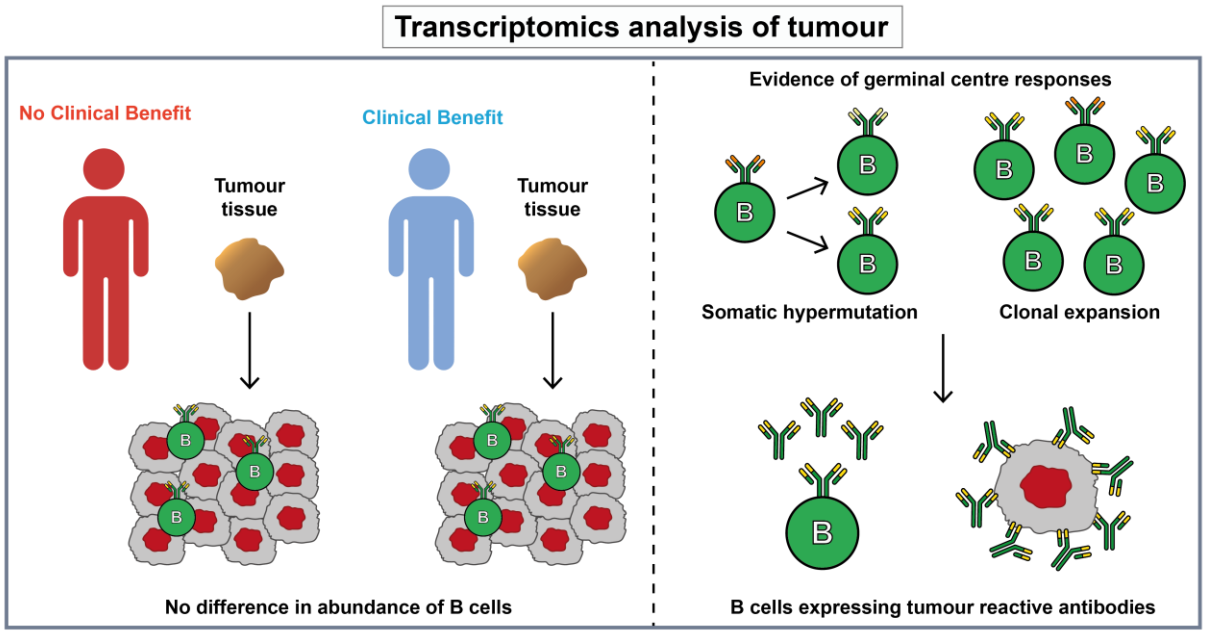


Figure 6.2. Abundance of intra-tumour B cells does not reflect clinical outcome to ICI in the LUD2015-005 inoperable cohort, but some intra-tumour B cells express tumour reactive antibodies. (Left) Transcriptomic analysis of oesophageal tumour tissue from patients of the LUD2015-005 inoperable cohort did not show any association of intra-tumour B cell abundance with clinical outcome. (Right) Intra-tumour B cells showed evidence of germinal centre responses and in some cases expressed tumour reactive antibodies.

Further functional *in vitro* assays showed that the antibodies produced by some of these ASCs bind to known autoantigens and they may function to modulate anti-tumour through antibody mediated phagocytosis of dying tumour cells.

6.1.3 Circulating anti-p53 autoantibodies and IgA-expressing B cells distinguish individuals with oesophageal cancer from those with Barrett's Oesophagus

In chapter 3, analysis comparing IgG AAb reactivity in BO and OC patients led to the identification of anti-p53 AAbs as a promising peripheral blood biomarker for distinguishing BO patients from OC patients with high specificity, but reasonably low sensitivity. Further work to identify additional peripheral blood biomarker was performed in chapter 4 by comparing PBMC samples from OC and BO patients. This identified IgA expression by B cells as a promising peripheral blood marker which was elevated in BO patients relative to OC patients (**Figure 6.3**).

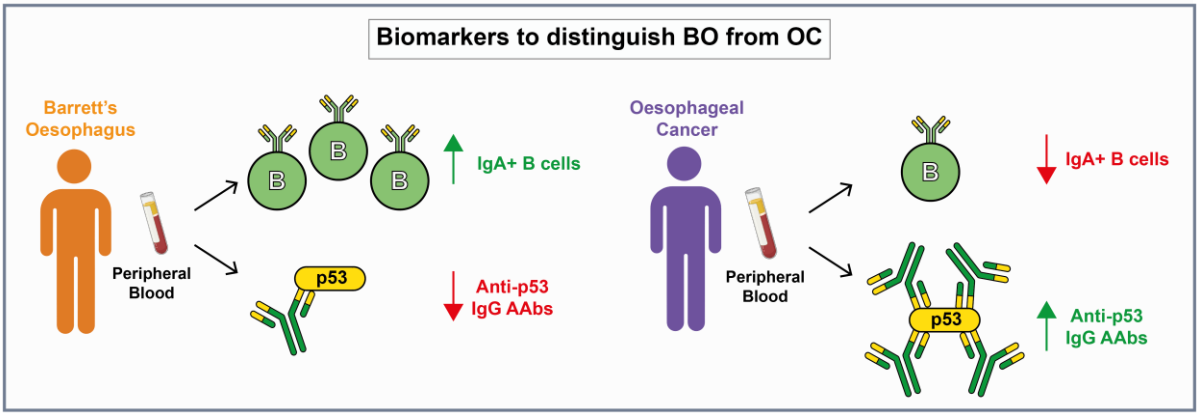


Figure 6.3 Circulating anti-p53 autoantibodies and IgA-expressing B cells non-invasively distinguish Barrett's Oesophagus and oesophageal cancer patients. Autoantibody (AAb) profiling and comprehensive *ex vivo* immunophenotyping of peripheral blood form Barrett's Oesophagus (BO) and oesophageal cancer (OC) patients identified increased prevalence of anti-p53 IgG AAbs and decreased IgA-expression by circulating B cells relative to BO patients.

In conclusion, I identified anti-p53 AAbs and B cell IgA expression as promising peripheral blood biomarkers to non-invasively distinguish OC and BO patients. Together, these findings provide a platform to develop a high-throughput clinical biomarker test to help monitor and stratify BO patients for endoscopic examination and hopefully detection of OC at an early stage.

6.2 Limitations

Despite the clinical potential of the findings summarised above, there are several limitations of the work presented in this thesis. The most notable of these is the relatively small cohort size of the LUD2015-005 inoperable cohort on which this work focuses on. Other notable limitations are the relatively limited antigen coverage of the antigen arrays used in Chapter 3, compared to newer approaches such as the REAP antigen platform[355]; the lack of *in vitro* functional assays to validate the signatures of activation/cytotoxicity of CD8 T cells, immunosuppressive capacity of myeloid cells, and B cell activation which were observed in Chapter 4; the lack of histological measurements of intra-tumour TLSs and B cells; and the limited replicates of experiments used to profile the subset of 34 of 491 expanded BCR clones which were expressed as recombinant antibodies in Chapter 5.

6.3 Future directions

To address the notable limitation of small cohort size, validation of key findings in the operable arm of LUD2015-005 trial and samples from other clinical trials of ICI treated OC should be performed. In particular, the findings in this thesis that merit the greatest consideration for further study are anti-p53 AAbs as a peripheral blood biomarker of OC, increased IgA expression by peripheral B cells in BO patients, the distinct peripheral blood-based signatures of CB and irAE which were identified, the lack of association of intra-tumour TLSs and B cells with clinical outcome to ICI, and the expression of tumour reactive BCRs within OC tumours.

To validate anti-p53 AAbs, an alternative method of AAb detection which utilises full length and post translationally modified p53, such as yeast display, should be used to detect anti-p53 AAbs in a large independent cohort of OC and BO patients. The same or a similar cohort could be used to validate the increased IgA expression by peripheral B cells in BO patients through a targeted flow cytometry-based assay. IgA titres from the blood plasma of the same patients should also be measured by ELISA to determine whether serum IgA is elevated in BO patients.

To further investigate and validate the distinct peripheral blood-based signatures of CB and irAE which were identified in Chapter 4, similar *ex vivo* immunophenotyping experiments should be performed in a large independent cohort of ICI treated OC patients. PBMCs samples from the same cohort should also be functional tested using *in vitro* assays of T cells, such as T cell killing of tumour cells, and B cells, such as activation, cytokine expression, and differentiation into Bregs. If high quality tumour biopsies are collected, the same cohort of ICI treated OC patients could also be used to validate the lack of association of intra-tumour TLSs and B cells with clinical outcome to ICI by immunohistochemistry or immunofluorescence-based measurement of TLS and B cell abundance in tumour sections.

To determine whether CB OAC patients from the LUD2015-005 trial inoperable cohort have increased BCR clonality and diversity relative to NCB patients, as was reported for ICI treated melanoma[108], targeted bulk BCR RNA-sequencing is planned in collaboration with Dr Rachael Bashford-Rogers (University of Oxford).

Finally, to further investigate and validate the expression of tumour reactive BCRs within OC tumours, additional batches of the identified 491 expanded BCR clones should be expressed and screened in a similar manner to that described in Chapter 5, in order to confirm whether tumour reactive antibodies are commonly expressed by tumour infiltrating B cells. Then, comprehensive characterisation of antibody binding to additional cancer cell lines and primary tumour tissue, ideally by multiple experiments, should be performed to validate the preliminary findings obtained from the screening of the 34 antibodies described in Chapter 5. Any cell line binding antibodies should then be expressed as their corresponding germline sequence to determine whether they acquired binding capabilities due to SHM or this binding was germline encoded. Further, to identify reactive antigens for antibodies of interest, immunoprecipitation using the recombinant antibodies, and subsequent mass spectrometry, should be performed.

7 References

1. Sung, H. *et al.* Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* **71**, 209–249 (2021).

2. Arnold, M., Laversanne, M., Brown, L. M., Devesa, S. S. & Bray, F. Predicting the Future Burden of Esophageal Cancer by Histological Subtype: International Trends in Incidence up to 2030. *American Journal of Gastroenterology* **112**, 1247–1255 (2017).

3. Jones, C. M., Hawkins, M., Mukherjee, S., Radhakrishna, G. & Crosby, T. Considerations for the Treatment of Oesophageal Cancer With Radiotherapy During the COVID-19 Pandemic. *Clin Oncol (R Coll Radiol)* **32**, 354 (2020).

4. National Institute for Health and Care Excellence. Recommendations | Oesophago-gastric cancer: assessment and management in adults | Guidance | NICE. https://www.nice.org.uk/guidance/ng83/chapter/Recommendations#radical-treatment-2.

5. Abnet, C. C., Arnold, M. & Wei, W. Q. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* **154**, 360 (2018).

6. Pennathur, A., Gibson, M. K., Jobe, B. A. & Luketich, J. D. Oesophageal carcinoma. *The Lancet* **381**, 400–412 (2013).

7. Fitzgerald, R. C. Review article: Barrett's oesophagus and associated adenocarcinoma – a UK perspective. *Aliment Pharmacol Ther* **20**, 45–49 (2004).

8. McColl, K. E. L. What is causing the rising incidence of esophageal adenocarcinoma in the West and will it also happen in the East? *J Gastroenterol* **54**, 669–673 (2019).

9. Solaymani-Dodaran, M., Logan, R. F. A., West, J., Card, T. & Coupland, C. Risk of oesophageal cancer in Barrett's oesophagus and gastro-oesophageal reflux. *Gut* **53**, 1070–1074 (2004).

10. Peters, Y. *et al.* Barrett oesophagus. *Nature Reviews Disease Primers 2019 5:1* **5**, 1–22 (2019).

11. Shekhar, R., Biyyani, S. & Chak, A. Barrett's esophagus: review of diagnosis and treatment. *Gastroenterol Rep (Oxf)* **1**, 9–18 (2013).

12. Mulholland, M. W., Reid, B. J., Levine, D. S. & Rubin, C. E. Elevated gastric acid secretion in patients with Barrett's metaplastic epithelium. *Digestive Diseases and Sciences 1989 34:9* **34**, 1329–1334 (1989).

13. Jagadesham, V. P. & Kelty, C. J. Low grade dysplasia in Barrett's esophagus: Should we worry? *World J Gastrointest Pathophysiol* **5**, 91 (2014).

14. Reid, B. J., Li, X., Galipeau, P. C. & Vaughan, T. L. Barrett's oesophagus and oesophageal adenocarcinoma: time for a new synthesis. *Nat Rev Cancer* **10**, 87 (2010).

15. Gerson, L. B., Shetler, K. & Triadafilopoulos, G. Prevalence of Barrett's esophagus in asymptomatic individuals. *Gastroenterology* **123**, 461–467 (2002).

16. Taylor, J. B. & Rubenstein, J. H. Meta-analyses of the effect of symptoms of gastroesophageal reflux on the risk of barrett's esophagus. *American Journal of Gastroenterology* **105**, 1730–1737 (2010).

17. Masclee, G. M. C., Coloma, P. M., de Wilde, M., Kuipers, E. J. & Sturkenboom, M. C. J. M. The incidence of Barrett's oesophagus and oesophageal adenocarcinoma in the United Kingdom and the Netherlands is levelling off. *Aliment Pharmacol Ther* **39**, 1321–1330 (2014).

18. Schnell, T. G. *et al.* Long-term nonsurgical management of Barrett's esophagus with high-grade dysplasia. *Gastroenterology* **120**, 1607–1619 (2001).

19. Pennathur, A., Landreneau, R. J. & Luketich, J. D. Surgical Aspects of the Patient with High-Grade Dysplasia. *Semin Thorac Cardiovasc Surg* **17**, 326–332 (2005).

20. Lekakos, L. *et al.* Barrett's esophagus with high-grade dysplasia: Focus on current treatment options. *World J Gastroenterol* **17**, 4174–4183 (2011).

21. Berry, M. F. Esophageal cancer: staging system and guidelines for staging and treatment. *J Thorac Dis* **6**, S289 (2014).

22. TNM staging | Oesophageal cancer | Cancer Research UK. https://www.cancerresearchuk.org/about-cancer/oesophageal-cancer/stages-types-and-grades/tnm-staging.

23. American Cancer Society. Treating Esophageal Cancer by Stage | Treating All Stages of Esophagus Cancer. https://www.cancer.org/cancer/esophagus-cancer/treating/by-stage.html.

24. NIH NCI SEER. SEER*Explorer Application. *Stage Distribution of SEER Incidence Cases, 2010-2019* https://seer.cancer.gov/statistics-network/explorer/application.html?site=17&data_type=1&graph_type=4&compareBy=sex&chk_sex_1=1&chk_sex_3=3&chk_sex_2=2&race=1&age_range=1&advopt_precision=1&hdn_view=0&advopt_display=2#graphArea.

25. Guyer, D. L., Almhanna, K. & McKee, K. Y. Palliative care for patients with esophageal cancer: a narrative review. *Ann Transl Med* **8**, 1103–1103 (2020).

26. Safran, H. P. *et al.* Trastuzumab with trimodality treatment for oesophageal adenocarcinoma with HER2 overexpression (NRG Oncology/RTOG 1010): a multicentre, randomised, phase 3 trial. *Lancet Oncol* **23**, 259–269 (2022).

27. FDA approves pembrolizumab for esophageal or GEJ carcinoma | FDA. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-pembrolizumab-esophageal-or-gej-carcinoma.

28. FDA approves nivolumab in combination with chemotherapy for metastatic gastric cancer and esophageal adenocarcinoma | FDA. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-nivolumab-combination-chemotherapy-metastatic-gastric-cancer-and-esophageal.

29. US Food and Drug Administration. FDA approves Opdivo in combination with chemotherapy and Opdivo in combination with Yervoy for first-line esophageal squamous cell carcinoma indications | FDA. https://www.fda.gov/drugs/resources-information-approved-drugs/fda-approves-opdivo-combination-chemotherapy-and-opdivo-combination-yervoy-first-line-esophageal.

30. Shitara, K. *et al.* Nivolumab plus chemotherapy or ipilimumab in gastro-oesophageal cancer. *Nature 2022 603:7903* **603**, 942–948 (2022).

31. Boku, N. *et al.* LBA7_PR Nivolumab plus chemotherapy versus chemotherapy alone in patients with previously untreated advanced or recurrent gastric/gastroesophageal junction (G/GEJ) cancer: ATTRACTION-4 (ONO-4538-37) study. *Annals of Oncology* **31**, S1192 (2020).

32. Janjigian, Y. Y. *et al.* First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *The Lancet* **398**, 27–40 (2021).

33. Huang, J. *et al.* Camrelizumab versus investigator's choice of chemotherapy as second-line therapy for advanced or metastatic oesophageal squamous cell carcinoma (ESCORT): a multicentre, randomised, open-label, phase 3 study. *Lancet Oncol* **21**, 832–842 (2020).

34. Shen, L. *et al.* Tislelizumab Versus Chemotherapy as Second-Line Treatment for Advanced or Metastatic Esophageal Squamous Cell Carcinoma (RATIONALE-302): A Randomized Phase III Study. *Journal of Clinical Oncology* **71**, (2022).

35. Lu, Z. *et al.* Sintilimab versus placebo in combination with chemotherapy as first line treatment for locally advanced or metastatic oesophageal squamous cell carcinoma (ORIENT-15): multicentre, randomised, double blind, phase 3 trial. *BMJ* **377**, (2022).

36. Sun, J. M. *et al.* Pembrolizumab plus chemotherapy versus chemotherapy alone for first-line treatment of advanced oesophageal cancer (KEYNOTE-590): a randomised, placebo-controlled, phase 3 study. *The Lancet* **398**, 759–771 (2021).

37. Fang, P. *et al.* Immunotherapy resistance in esophageal cancer: Possible mechanisms and clinical implications. *Front Immunol* **13**, 4980 (2022).

38. Zhu, Y. *et al.* Predicting response to immunotherapy plus chemotherapy in patients with esophageal squamous cell carcinoma using non-invasive Radiomic biomarkers. *BMC Cancer* **21**, 1–11 (2021).

39. Hodi, F. S. *et al.* Improved Survival with Ipilimumab in Patients with Metastatic Melanoma. *New England Journal of Medicine* **363**, 711–723 (2010).

40. Wang, D. R., Wu, X. L. & Sun, Y. L. Therapeutic targets and biomarkers of tumor immunotherapy: response versus non-response. *Signal Transduction and Targeted Therapy 2022 7:1* **7**, 1–27 (2022).

41. Hogan, S. A., Levesque, M. P. & Cheng, P. F. Melanoma immunotherapy: Next-generation biomarkers. *Front Oncol* **8**, 178 (2018).

42. Das, S. & Johnson, D. B. Immune-related adverse events and anti-tumor efficacy of immune checkpoint inhibitors. *J Immunother Cancer* **7**, 306 (2019).

43. Sambi, M., Bagheri, L. & Szewczuk, M. R. Current Challenges in Cancer Immunotherapy: Multimodal Approaches to Improve Efficacy and Patient Response Rates. *J Oncol* **2019**, (2019).

44. Takeuchi, O. & Akira, S. Pattern Recognition Receptors and Inflammation. *Cell* **140**, 805–820 (2010).

45. Iwasaki, A. & Medzhitov, R. Control of adaptive immunity by the innate immune system. *Nature Immunology 2015 16:4* **16**, 343–353 (2015).

46. Vénéreau, E., Ceriotti, C. & Bianchi, M. E. DAMPs from cell death to new life. *Front Immunol* **6**, 422 (2015).

47. Janeway, C. A. & Medzhitov, R. Innate Immune Recognition. *Annu Rev Immunol* **20**, 197–216 (2002).

48. Medzhitov, R. Origin and physiological roles of inflammation. *Nature 2008 454:7203* **454**, 428–435 (2008).

49. Fearon, D. T. & Locksley, R. M. The Instructive Role of Innate Immunity in the Acquired Immune Response. *Science* **272**, 50–54 (1996).

50. Banchereau, J. & Steinman, R. M. Dendritic cells and the control of immunity. *Nature 1998 392:6673* **392**, 245–252 (1998).

51. Steinman, R. M. Dendritic cells: versatile controllers of the immune system. *Nature Medicine 2007 13:10* **13**, 1155–1159 (2007).

52. Collin, M. & Bigley, V. Human dendritic cell subsets: an update. *Immunology* **154**, 3 (2018).

53. Kadowaki, N. Dendritic Cells—A Conductor of T Cell Differentiation—. *Allergology International* **56**, 193–199 (2007).

54. Wieczorek, M. *et al.* Major histocompatibility complex (MHC) class I and MHC class II proteins: Conformational plasticity in antigen presentation. *Front Immunol* **8**, 292 (2017).

55. Sigal, L. J. Activation of CD8 T Lymphocytes during Viral Infections. *Encyclopedia of Immunobiology* **4**, 286 (2016).

56. Philip, M. & Schietinger, A. CD8+ T cell differentiation and dysfunction in cancer. *Nature Reviews Immunology 2021 22:4* **22**, 209–223 (2021).

57. Luckheeram, R. V., Zhou, R., Verma, A. D. & Xia, B. CD4+T Cells: Differentiation and Functions. *Clin Dev Immunol* **2012**, 12 (2012).

58. Heesters, B. A., van der Poel, C. E., Das, A. & Carroll, M. C. Antigen Presentation to B Cells. *Trends Immunol* **37**, 844–854 (2016).

59. Kwak, K., Akkaya, M. & Pierce, S. K. B cell signaling in context. *Nature Immunology 2019 20:8* **20**, 963–969 (2019).

60. Forthal, D. N. Functions of Antibodies. *Microbiol Spectr* **2**, 1 (2015).

61. Iwata, Y. *et al.* Characterization of a rare IL-10–competent B-cell subset in humans that parallels mouse regulatory B10 cells. *Blood* **117**, 530–541 (2011).

62. Cashman, K. S. *et al.* Understanding and Measuring Human B Cell Tolerance and its Breakdown in Autoimmune Disease. *Immunol Rev* **292**, 76 (2019).

63. Xing, Y. & Hogquist, K. A. T-Cell Tolerance: Central and Peripheral. *Cold Spring Harb Perspect Biol* **4**, 1–15 (2012).

64. da Gama Duarte, J., Woods, K., Andrews, M. C. & Behren, A. The good, the (not so) bad and the ugly of immune homeostasis in melanoma. *Immunol Cell Biol* **96**, 497–506 (2018).

65. Greten, F. R. & Grivennikov, S. I. Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity* **51**, 27–41 (2019).

66. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov* **12**, 31–46 (2022).

67. Yi, M., Niu, M., Xu, L., Luo, S. & Wu, K. Regulation of PD-L1 expression in the tumor microenvironment. *Journal of Hematology & Oncology 2021 14:1* **14**, 1–13 (2021).

68. Okazaki, T., Maeda, A., Nishimura, H., Kurosaki, T. & Honjo, T. PD-1 immunoreceptor inhibits B cell receptor-mediated signaling by recruiting src homology 2-domain-containing tyrosine phosphatase 2 to phosphotyrosine. *Proceedings of the National Academy of Sciences* **98**, 13866–13871 (2001).

69. 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 (1992).

70. Agata, Y. *et al.* Expression of the PD-1 antigen on the surface of stimulated mouse T and B lymphocytes. *Int Immunol* **8**, 765–772 (1996).

71. Leach, D. R., Krummel, M. F. & Allison, J. P. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* **271**, 1734–1736 (1996).

72. Qin, S. *et al.* Novel immune checkpoint targets: moving beyond PD-1 and CTLA-4. *Molecular Cancer 2019 18:1* **18**, 1–14 (2019).

73. Conroy, M. & Naidoo, J. Immune-related adverse events and the balancing act of immunotherapy. *Nature Communications 2022 13:1* **13**, 1–4 (2022).

74. Robert, C. A decade of immune-checkpoint inhibitors in cancer therapy. *Nature Communications 2020 11:1* **11**, 1–3 (2020).

75. Han, Y., Liu, D. & Li, L. PD-1/PD-L1 pathway: current researches in cancer. *Am J Cancer Res* **10**, 727 (2020).

76. Thompson, R. H., Allison, J. P. & Kwon, E. D. Anti-cytotoxic T lymphocyte antigen-4 (CTLA-4) immunotherapy for the treatment of prostate cancer. *Urologic Oncology: Seminars and Original Investigations* **24**, 442–447 (2006).

77. Chambers, C. A., Kuhns, M. S., Egen, J. G. & Allison, J. P. CTLA-4-mediated inhibition in regulation of T cell responses: mechanisms and manipulation in tumor immunotherapy. *Annu Rev Immunol* **19**, 565–594 (2001).

78. Schirrmacher, V. From chemotherapy to biological therapy: A review of novel concepts to reduce the side effects of systemic cancer treatment (Review). *Int J Oncol* **54**, 407 (2019).

79. Zhou, X., Ren, T., Zan, H., Hua, C. & Guo, X. Novel Immune Checkpoints in Esophageal Cancer: From Biomarkers to Therapeutic Targets. *Front Immunol* **13**, (2022).

80. Shah, M. A. *et al.* Efficacy and Safety of Pembrolizumab for Heavily Pretreated Patients With Advanced, Metastatic Adenocarcinoma or Squamous Cell Carcinoma of the Esophagus: The Phase 2 KEYNOTE-180 Study. *JAMA Oncol* **5**, 546–550 (2019).

81. Kojima, T. *et al.* Randomized phase III KEYNOTE-181 study of pembrolizumab versus chemotherapy in advanced esophageal cancer. *Journal of Clinical Oncology* **38**, 4138–4148 (2020).

82. Xu, J. *et al.* Clinical and biomarker analyses of sintilimab versus chemotherapy as second-line therapy for advanced or metastatic esophageal squamous cell carcinoma: a randomized, open-label phase 2 study (ORIENT-2). *Nature Communications 2022 13:1* **13**, 1–12 (2022).

83. Wang, F. *et al.* Association of frequent amplification of chromosome 11q13 in esophageal squamous cell cancer with clinical benefit to immune check point blockade. *https://doi.org/10.1200/JCO.2019.37.15_suppl.4036* **37**, 4036–4036 (2019).

84. Bai, R., Lv, Z., Xu, D. & Cui, J. Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors. *Biomarker Research 2020 8:1* **8**, 1–17 (2020).

85. Rozeman, E. A. *et al.* Survival and biomarker analyses from the OpACIN-neo and OpACIN neoadjuvant immunotherapy trials in stage III melanoma. *Nature Medicine 2021 27:2* **27**, 256–263 (2021).

86. Chowell, D. *et al.* Patient HLA class I genotype influences cancer response to checkpoint blockade immunotherapy. *Science* **359**, 582–587 (2018).

87. Li, H., van der Merwe, P. A. & Sivakumar, S. Biomarkers of response to PD-1 pathway blockade. *British Journal of Cancer 2022 126:12* **126**, 1663–1675 (2022).

88. Topalian, S. L. *et al.* Safety, Activity, and Immune Correlates of Anti–PD-1 Antibody in Cancer. *New England Journal of Medicine* **366**, 2443–2454 (2012).

89. Taube, J. M. *et al.* Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clinical Cancer Research* **20**, 5064–5074 (2014).

90. Kulangara, K. *et al.* Development of the combined positive score (CPS) for the evaluation of PD-L1 in solid tumors with the immunohistochemistry assay PD-L1 IHC 22C3 pharmDx. *https://doi.org/10.1200/JCO.2017.35.15_suppl.e14589* **35**, e14589–e14589 (2017).

91. Kulangara, K. *et al.* Clinical Utility of the Combined Positive Score for Programmed Death Ligand-1 Expression and the Approval of Pembrolizumab for Treatment of Gastric Cancer. *Arch Pathol Lab Med* **143**, 330–337 (2019).

92. Topalian, S. L., Taube, J. M., Anders, R. A. & Pardoll, D. M. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nature Reviews Cancer 2016 16:5* **16**, 275–287 (2016).

93. Patel, S. P. & Kurzrock, R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol Cancer Ther* **14**, 847–856 (2015).

94. Chae, Y. K. *et al.* Biomarkers for PD-1/PD-L1 Blockade Therapy in Non–Small-cell Lung Cancer: Is PD-L1 Expression a Good Marker for Patient Selection? *Clin Lung Cancer* **17**, 350–361 (2016).

95. Strickler, J. H., Hanks, B. A. & Khasraw, M. Tumor mutational burden as a predictor of immunotherapy response: Is more always better? *Clinical Cancer Research* **27**, 1236–1241 (2021).

96. Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *New England Journal of Medicine* **377**, 2500–2501 (2017).

97. 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).

98. Rizvi, N. A. *et al.* Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* **348**, 124–128 (2015).

99. Snyder, A. *et al.* Genetic Basis for Clinical Response to CTLA-4 Blockade in Melanoma. *New England Journal of Medicine* **371**, 2189–2199 (2014).

100. Chan, T. A. *et al.* Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Annals of Oncology* **30**, 44–56 (2019).

101. Schumacher, T. N. & Schreiber, R. D. Neoantigens in cancer immunotherapy. *Science* **348**, 69–74 (2015).

102. Chao, J. *et al.* Pembrolizumab (pembro) in microsatellite instability-high (MSI-H) advanced gastric/gastroesophageal junction (G/GEJ) cancer by line of therapy. *https://doi.org/10.1200/JCO.2020.38.4_suppl.430* **38**, 430–430 (2020).

103. André, T. *et al.* Pembrolizumab in Microsatellite-Instability–High Advanced Colorectal Cancer. *New England Journal of Medicine* **383**, 2207–2218 (2020).

104. Overman, M. J. *et al.* Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet Oncol* **18**, 1182–1191 (2017).

105. Snyder, A. *et al.* Contribution of systemic and somatic factors to clinical response and resistance to PD-L1 blockade in urothelial cancer: An exploratory multi-omic analysis. *PLoS Med* **14**, e1002309 (2017).

106. Petitprez, F. *et al.* B cells are associated with survival and immunotherapy response in sarcoma. *Nature 2020 577:7791* **577**, 556–560 (2020).

107. Cabrita, R. *et al.* Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature 2020 577:7791* **577**, 561–565 (2020).

108. Helmink, B. A. *et al.* B cells and tertiary lymphoid structures promote immunotherapy response. *Nature 2020 577:7791* **577**, 549–555 (2020).

109. Meylan, M. *et al.* Tertiary lymphoid structures generate and propagate anti-tumor antibody-producing plasma cells in renal cell cancer. *Immunity* **55**, 527-541.e5 (2022).

110. Munoz-Erazo, L., Rhodes, J. L., Marion, V. C. & Kemp, R. A. Tertiary lymphoid structures in cancer – considerations for patient prognosis. *Cellular & Molecular Immunology 2020 17:6* **17**, 570–575 (2020).

111. Neyt, K., Perros, F., GeurtsvanKessel, C. H., Hammad, H. & Lambrecht, B. N. Tertiary lymphoid organs in infection and autoimmunity. *Trends Immunol* **33**, 297–305 (2012).

112. Gunderson, A. J. *et al.* Germinal center reactions in tertiary lymphoid structures associate with neoantigen burden, humoral immunity and long-term survivorship in pancreatic cancer. *Oncoimmunology* **10**, (2021).

113. Lin, Z. *et al.* Pan-cancer analysis of genomic properties and clinical outcome associated with tumor tertiary lymphoid structure. *Scientific Reports 2020 10:1* **10**, 1–14 (2020).

114. Sautès-Fridman, C., Petitprez, F., Calderaro, J. & Fridman, W. H. Tertiary lymphoid structures in the era of cancer immunotherapy. *Nature Reviews Cancer 2019 19:6* **19**, 307–325 (2019).

115. Dieu-Nosjean, M. C. *et al.* Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. *Immunol Rev* **271**, 260–275 (2016).

116. Prabhakaran, S. *et al.* Evaluation of invasive breast cancer samples using a 12-chemokine gene expression score: Correlation with clinical outcomes. *Breast Cancer Research* **19**, 1–11 (2017).

117. Coppola, D. *et al.* Unique Ectopic Lymph Node-Like Structures Present in Human Primary Colorectal Carcinoma Are Identified by Immune Gene Array Profiling. *Am J Pathol* **179**, 37–45 (2011).

118. Groeneveld, C. S. *et al.* Tertiary lymphoid structures marker CXCL13 is associated with better survival for patients with advanced-stage bladder cancer treated with immunotherapy. *Eur J Cancer* **148**, 181–189 (2021).

119. Messina, J. L. *et al.* 12-Chemokine Gene Signature Identifies Lymph Node-like Structures in Melanoma: Potential for Patient Selection for Immunotherapy? *Scientific Reports 2012 2:1* **2**, 1–6 (2012).

120. Nixon, A. B. *et al.* Peripheral immune-based biomarkers in cancer immunotherapy: can we realize their predictive potential? *J Immunother Cancer* **7**, (2019).

121. Choo, S. Y. The HLA System: Genetics, Immunology, Clinical Testing, and Clinical Implications. *Yonsei Med J* **48**, 11 (2007).

122. Lee, C.-H. *et al.* Genomic biomarkers of response to nivolumab/ipilimumab (nivo/ipi) and nivolumab (nivo) monotherapy in 108 patients with advanced renal cell carcinoma. *https://doi.org/10.1200/JCO.2019.37.7_suppl.641* **37**, 641–641 (2019).

123. Yang, Y., Kim, E. & Kim, S. Insignificant effects of loss of heterozygosity in HLA in the efficacy of immune checkpoint blockade treatment. *Genes Genomics* **44**, 509–515 (2022).

124. Lu, Z. *et al.* Germline HLA-B evolutionary divergence influences the efficacy of immune checkpoint blockade therapy in gastrointestinal cancer. *Genome Med* **13**, (2021).

125. Wistuba-Hamprecht, K. *et al.* Peripheral CD8 effector-memory type 1 T-cells correlate with outcome in ipilimumab-treated stage IV melanoma patients. *Eur J Cancer* **73**, 61–70 (2017).

126. Ottonello, S. *et al.* Association Between Response to Nivolumab Treatment and Peripheral Blood Lymphocyte Subsets in Patients With Non-small Cell Lung Cancer. *Front Immunol* **11**, 125 (2020).

127. Wu, T. D. *et al.* Peripheral T cell expansion predicts tumour infiltration and clinical response. *Nature 2020 579:7798* **579**, 274–278 (2020).

128. Han, J. *et al.* TCR repertoire diversity of peripheral PD-1þCD8þ T cells predicts clinical outcomes after immunotherapy in patients with non–small cell lung cancer. *Cancer Immunol Res* **8**, 146–154 (2020).

129. Kagamu, H. *et al.* CD4+ T-cell immunity in the peripheral blood correlates with response to Anti-PD-1 therapy. *Cancer Immunol Res* **8**, 334–344 (2020).

130. Valpione, S. *et al.* Immune awakening revealed by peripheral T cell dynamics after one cycle of immunotherapy. *Nature Cancer 2020 1:2* **1**, 210–221 (2020).

131. Watson, R. A. *et al.* Immune checkpoint blockade sensitivity and progression-free survival associates with baseline CD8+ T cell clone size and cytotoxicity. *Sci Immunol* **6**, 8825 (2021).

132. 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 2020 26:2* **26**, 193–199 (2020).

133. Krieg, C. *et al.* High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nature Medicine 2018 24:2* **24**, 144–153 (2018).

134. Cho, Y. H. *et al.* Natural Killer Cells as a Potential Biomarker for Predicting Immunotherapy Efficacy in Patients with Non-Small Cell Lung Cancer. *Target Oncol* **15**, 241–247 (2020).

135. Mazzaschi, G. *et al.* The circulating pool of functionally competent NK and CD8+ cells predicts the outcome of anti-PD1 treatment in advanced NSCLC. *Lung Cancer* **127**, 153–163 (2019).

136. Wistuba-Hamprecht, K. *et al.* Establishing High Dimensional Immune Signatures from Peripheral Blood via Mass Cytometry in a Discovery Cohort of Stage IV Melanoma Patients. *The Journal of Immunology* **198**, 927–936 (2017).

137. Weber, J. S. *et al.* Safety profile of nivolumab monotherapy: A pooled analysis of patients with advanced melanoma. *Journal of Clinical Oncology* **35**, 785–792 (2017).

138. Postow, M. A., Sidlow, R. & Hellmann, M. D. Immune-Related Adverse Events Associated with Immune Checkpoint Blockade. *New England Journal of Medicine* **378**, 158–168 (2018).

139. Ye, W. *et al.* Checkpoint-blocker-induced autoimmunity is associated with favourable outcome in metastatic melanoma and distinct T-cell expression profiles. *British Journal of Cancer 2021 124:10* **124**, 1661–1669 (2021).

140. Freeman-Keller, M. *et al.* Nivolumab in Resected and Unresectable Metastatic Melanoma: Characteristics of Immune-Related Adverse Events and Association with Outcomes. *Clin Cancer Res* **22**, 886–894 (2016).

141. Fukihara, J. *et al.* Prognostic Impact and Risk Factors of Immune-Related Pneumonitis in Patients With Non–Small-Cell Lung Cancer Who Received Programmed Death 1 Inhibitors. *Clin Lung Cancer* **20**, 442-450.e4 (2019).

142. Suresh, K. *et al.* Impact of Checkpoint Inhibitor Pneumonitis on Survival in NSCLC Patients Receiving Immune Checkpoint Immunotherapy. *Journal of Thoracic Oncology* **14**, 494–502 (2019).

143. Brahmer, J. R. *et al.* Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy: American society of clinical oncology clinical practice guideline. *Journal of Clinical Oncology* **36**, 1714–1768 (2018).

144. Kalinich, M. *et al.* Prediction of severe immune-related adverse events requiring hospital admission in patients on immune checkpoint inhibitors: study of a population level insurance claims database from the USA. *J Immunother Cancer* **9**, e001935 (2021).

145. Tison, A., Garaud, S., Chiche, L., Cornec, D. & Kostine, M. Immune-checkpoint inhibitor use in patients with cancer and pre-existing autoimmune diseases. *Nature Reviews Rheumatology 2022* 1–16 (2022) doi:10.1038/s41584-022-00841-0.

146. Lauss, M., Donia, M., Svane, I. M. & Jönsson, G. B Cells and Tertiary Lymphoid Structures: Friends or Foes in Cancer Immunotherapy? *Clinical Cancer Research* **28**, 1751–1758 (2022).

147. Fridman, W. H. *et al.* B cells and cancer: To B or not to B? *Journal of Experimental Medicine* **218**, (2021).

148. Rothstein, T. L., Griffin, D. O., Holodick, N. E., Quach, T. D. & Kaku, H. Human B-1 cells take the stage. *Ann N Y Acad Sci* **1285**, 97–114 (2013).

149. Griffin, D. O., Holodick, N. E. & Rothstein, T. L. Human B1 cells in umbilical cord and adult peripheral blood express the novel phenotype CD20+CD27+CD43+CD70−. *Journal of Experimental Medicine* **208**, 67–80 (2011).

150. Galili, U. Human Natural Antibodies to Mammalian Carbohydrate Antigens as Unsung Heroes Protecting against Past, Present, and Future Viral Infections. doi:10.3390/antib9020025.

151. Gunti, S. & Notkins, A. L. Polyreactive Antibodies: Function and Quantification. doi:10.1093/infdis/jiu512.

152. Binder, C. J. Natural IgM Antibodies Against Oxidation-Specific Epitopes. *Journal of Clinical Immunology 2010 30:1* **30**, 56–60 (2010).

153. Grönwall, C., Vas, J. & Silverman, G. J. Protective Roles of Natural IgM Antibodies. *Front Immunol* **3**, (2012).

154. Ochsenbein, A. F. *et al.* Control of early viral and bacterial distribution and disease by natural antibodies. *Science* **286**, 2156–2159 (1999).

155. Eichinger, S. *et al.* Natural antibodies to oxidation-specific epitopes: innate immune response and venous thromboembolic disease. *Journal of Thrombosis and Haemostasis* **16**, 31–35 (2018).

156. Appelgren, D., Eriksson, P., Ernerudh, J. & Segelmark, M. Marginal-Zone B-Cells are main producers of IgM in humans, and are reduced in patients with autoimmune vasculitis. *Front Immunol* **9**, 2242 (2018).

157. Galy, A., Travis, M., Cen, D. & Chen, B. Human T, B, natural killer, and dendritic cells arise from a common bone marrow progenitor cell subset. *Immunity* **3**, 459–473 (1995).

158. Nemazee, D. Mechanisms of central tolerance for B cells. *Nature Reviews Immunology 2017 17:5* **17**, 281–294 (2017).

159. Schroeder, K., Herrmann, M. & Winkler, T. H. The role of somatic hypermutation in the generation of pathogenic antibodies in SLE. *http://dx.doi.org/10.3109/08916934.2012.748751* **46**, 121–127 (2013).

160. Musette, P. & Bouaziz, J. D. B cell modulation strategies in autoimmune diseases: New concepts. *Front Immunol* **9**, 622 (2018).

161. Elkon, K. & Casali, P. Nature and functions of autoantibodies. *Nat Clin Pract Rheumatol* **4**, 491 (2008).

162. Pieper, K., Grimbacher, B. & Eibel, H. B-cell biology and development. *Journal of Allergy and Clinical Immunology* **131**, 959–971 (2013).

163. Tonegawa, S. Somatic generation of antibody diversity. *Nature 1983 302:5909* **302**, 575–581 (1983).

164. Schroeder, H. W. & Cavacini, L. Structure and Function of Immunoglobulins. *J Allergy Clin Immunol* **125**, S41 (2010).

165. Stanhope-Baker, P., Hudson, K. M., Shaffer, A. L., Constantinescu, A. & Schlissel, M. S. Cell Type–Specific Chromatin Structure Determines the Targeting of V(D)J Recombinase Activity In Vitro. *Cell* **85**, 887–897 (1996).

166. Karasuyama, H., Kudo, A. & Melchers, F. The proteins encoded by the VpreB and lambda 5 pre-B cell-specific genes can associate with each other and with mu heavy chain. *Journal of Experimental Medicine* **172**, 969–972 (1990).

167. Melchers, F. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nature Reviews Immunology 2005 5:7* **5**, 578–584 (2005).

168. LeBien, T. W. Fates of human B-cell precursors. *Blood* **96**, 9–23 (2000).

169. ten Boekel, E., Melchers, F. & Rolink, A. G. Precursor B Cells Showing H Chain Allelic Inclusion Display Allelic Exclusion at the Level of Pre-B Cell Receptor Surface Expression. *Immunity* **8**, 199–207 (1998).

170. Guo, B., Kato, R. M., Garcia-Lloret, M., Wahl, M. I. & Rawlings, D. J. Engagement of the Human Pre-B Cell Receptor Generates a Lipid Raft–Dependent Calcium Signaling Complex. *Immunity* **13**, 243–253 (2000).

171. Herzog, S., Reth, M. & Jumaa, H. Regulation of B-cell proliferation and differentiation by pre-B-cell receptor signalling. *Nature Reviews Immunology 2009 9:3* **9**, 195–205 (2009).

172. Turner, M. *et al.* Perinatal lethality and blocked B-cell development in mice lacking the tyrosine kinase Syk. *Nature 1995 378:6554* **378**, 298–302 (1995).

173. Cheng, A. M. *et al.* Syk tyrosine kinase required for mouse viability and B-cell development. *Nature 1995 378:6554* **378**, 303–306 (1995).

174. Reth, M., Petrac, E., Wiese, P., Lobel, L. & Alt, F. W. Activation of V kappa gene rearrangement in pre-B cells follows the expression of membrane-bound immunoglobulin heavy chains. *EMBO J* **6**, 3299–3305 (1987).

175. Campana, D., Farrant, J., Inamdar, N., Webster, A. D. & Janossy, G. Phenotypic features and proliferative activity of B cell progenitors in X-linked agammaglobulinemia. *The Journal of Immunology* **145**, (1990).

176. Ruddle, N. H. & Akirav, E. M. Secondary Lymphoid Organs: Responding to Genetic and Environmental Cues in Ontogeny and the Immune Response. *J Immunol* **183**, 2205 (2009).

177. Victora, G. D. & Nussenzweig, M. C. Germinal Centers. (2012) doi:10.1146/annurev-immunol-020711-075032.

178. Bemark, M. & Bemark, M. Translating transitions - how to decipher peripheral human B cell development. *The Journal of Biomedical Research, 2015, Vol. 29, Issue 4, Pages: 264-284* **29**, 264–284 (2015).

179. Martin, V. G. *et al.* Transitional B cells in early human B cell development - Time to revisit the paradigm? *Front Immunol* **7**, 546 (2016).

180. Vossenkämper, A. *et al.* A role for gut-associated lymphoid tissue in shaping the human B cell repertoire. *Journal of Experimental Medicine* **210**, 1665–1674 (2013).

181. Agrawal, S., Smith, S. A. B. C., Tangye, S. G. & Sewell, W. A. Transitional B cell subsets in human bone marrow. *Clin Exp Immunol* **174**, 53–59 (2013).

182. Sims, G. P. *et al.* Identification and characterization of circulating human transitional B cells. *Blood* **105**, 4390–4398 (2005).

183. Palanichamy, A. *et al.* Novel Human Transitional B Cell Populations Revealed by B Cell Depletion Therapy. *The Journal of Immunology* **182**, 5982–5993 (2009).

184. Suryani, S. *et al.* Differential expression of CD21 identifies developmentally and functionally distinct subsets of human transitional B cells. *Blood* **115**, 519–529 (2010).

185. Klein, U. *et al.* Transcriptional analysis of the B cell germinal center reaction. *Proc Natl Acad Sci U S A* **100**, 2639–2644 (2003).

186. Farmer, J. R. *et al.* Induction of metabolic quiescence defines the transitional to follicular B cell switch. *Sci Signal* **12**, 5573 (2019).

187. Schiemann, B. *et al.* An essential role for BAFF in the normal development of B cells through a BCMA-independent pathway. *Science* **293**, 2111–2114 (2001).

188. Nemazee, D. Natural history of MZ B cells. *Journal of Experimental Medicine* **218**, (2021).

189. Bemark, M. *et al.* A glycosylation-dependent CD45RB epitope defines previously unacknowledged CD27−IgMhigh B cell subpopulations enriched in young children and after hematopoietic stem cell transplantation. *Clinical Immunology* **149**, 421–431 (2013).

190. Descatoire, M. *et al.* Identification of a human splenic marginal zone B cell precursor with NOTCH2-dependent differentiation properties. *Journal of Experimental Medicine* **211**, 987–1000 (2014).

191. Zhao, Y. *et al.* Spatiotemporal segregation of human marginal zone and memory B cell populations in lymphoid tissue. *Nature Communications 2018 9:1* **9**, 1–15 (2018).

192. Tull, T. J. *et al.* Human marginal zone B cell development from early T2 progenitors. *Journal of Experimental Medicine* **218**, (2021).

193. Siu, J. H. Y. *et al.* Two subsets of human marginal zone B cells resolved by global analysis of lymphoid tissues and blood. *Sci Immunol* **7**, eabm9060 (2022).

194. Weller, S. *et al.* Human blood IgM "memory" B cells are circulating splenic marginal zone B cells harboring a prediversified immunoglobulin repertoire. *Blood* **104**, 3647–3654 (2004).

195. Spencer, J., Perry, M. E. & Dunn-Walters, D. K. Human marginal-zone B cells. *Immunol Today* **19**, 421–426 (1998).

196. Kruetzmann, S. *et al.* Human Immunoglobulin M Memory B Cells Controlling Streptococcus pneumoniae Infections Are Generated in the Spleen. *Journal of Experimental Medicine* **197**, 939–945 (2003).

197. Suzuki, K., Grigorova, I., Phan, T. G., Kelly, L. M. & Cyster, J. G. Visualizing B cell capture of cognate antigen from follicular dendritic cells. *Journal of Experimental Medicine* **206**, 1485–1493 (2009).

198. Carrasco, Y. R. & Batista, F. D. B Cells Acquire Particulate Antigen in a Macrophage-Rich Area at the Boundary between the Follicle and the Subcapsular Sinus of the Lymph Node. *Immunity* **27**, 160–171 (2007).

199. Duong, B. H. *et al.* Decoration of T-independent antigen with ligands for CD22 and Siglec-G can suppress immunity and induce B cell tolerance in vivo. *Journal of Experimental Medicine* **207**, 173–187 (2010).

200. Coutinho, A., Gronowicz, E., Bullock, W. W. & Möller, G. MECHANISM OF THYMUS-INDEPENDENT IMMUNOCYTE TRIGGERING MITOGENIC ACTIVATION OF B CELLS RESULTS IN SPECIFIC IMMUNE RESPONSES. *Journal of Experimental Medicine* **139**, 74–92 (1974).

201. Akkaya, M., Kwak, K. & Pierce, S. K. B cell memory: building two walls of protection against pathogens. *Nature Reviews Immunology 2019 20:4* **20**, 229–238 (2019).

202. Jenks, S. A., Cashman, K. S., Woodruff, M. C., Lee, F. E. H. & Sanz, I. Extrafollicular responses in humans and SLE. *Immunol Rev* **288**, 136 (2019).

203. Goodnow, C. C., Vinuesa, C. G., Randall, K. L., MacKay, F. & Brink, R. Control systems and decision making for antibody production. *Nature Immunology 2010 11:8* **11**, 681–688 (2010).

204. MacLennan, I. C. M. *et al.* Extrafollicular antibody responses. *Immunol Rev* **194**, 8–18 (2003).

205. Frölich, D. *et al.* Secondary Immunization Generates Clonally Related Antigen-Specific Plasma Cells and Memory B Cells. *The Journal of Immunology* **185**, 3103–3110 (2010).

206. Victora, G. D. *et al.* Germinal Center Dynamics Revealed by Multiphoton Microscopy with a Photoactivatable Fluorescent Reporter. *Cell* **143**, 592–605 (2010).

207. Klein, U. & Dalla-Favera, R. Germinal centres: role in B-cell physiology and malignancy. *Nature Reviews Immunology 2008 8:1* **8**, 22–33 (2008).

208. Allen, C. D. C., Okada, T., Tang, H. L. & Cyster, J. G. Imaging of germinal center selection events during affinity maturation. *Science* **315**, 528–531 (2007).

209. Schwickert, T. A. *et al.* In vivo imaging of germinal centres reveals a dynamic open structure. *Nature 2006 446:7131* **446**, 83–87 (2007).

210. Allen, C. D. C. *et al.* Germinal center dark and light zone organization is mediated by CXCR4 and CXCR5. *Nature Immunology 2004 5:9* **5**, 943–952 (2004).

211. Roco, J. A. *et al.* Class Switch Recombination Occurs Infrequently in Germinal Centers. *Immunity* **51**, 337 (2019).

212. Chaudhuri, J. & Alt, F. W. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol* **4**, 541–552 (2004).

213. Vries Gascan, de H., Gauchat, J. F., Aversa, G. & van Vlasselaer, P. different signaling pathways. switching in purified human B cells via cell clones and IL-4 induce IgG4 and IgE Anti-CD40 monoclonal antibodies or CD4+ T. (1991).

214. Gascan, H. *et al.* Human B cell clones can be induced to proliferate and to switch to IgE and IgG4 synthesis by interleukin 4 and a signal provided by activated CD4+ T cell clones. *Journal of Experimental Medicine* **173**, 747–750 (1991).

215. Cocks, B. G., de Waal Malefyt, R., Galizzi, J. pierre, de Vries, J. E. & Aversa, G. IL-13 induces proliferation and differentiation of human B cells activated by the CD40 ligand. *Int Immunol* **5**, 657–663 (1993).

216. Stavnezer, J. Regulation of antibody production and class switching by TGF-beta. *The Journal of Immunology* **155**, (1995).

217. Islam, K. B., Nilsson, L., Sideras, P., Hammarstrom, L. & Smith, C. I. E. TGF-p1 induces germ-line transcripts of both IgA subclasses in human B lymphocytes. *Int Immunol* **3**, 1099–1106 (1991).

218. Tangye, S. G., Ferguson, A., Avery, D. T., Ma, C. S. & Hodgkin, P. D. Isotype Switching by Human B Cells Is Division-Associated and Regulated by Cytokines. *The Journal of Immunology* **169**, 4298–4306 (2002).

219. Fujieda, S., Saxon, A. & Zhang, K. E. Direct evidence that γ1 and γ3 switching in human B cells is interleukin-10 dependent. *Mol Immunol* **33**, 1335–1343 (1996).

220. Brière, F., Servet-Delprat, C., Bridon, J. M., Saint-Remy, J. M. & Banchereau, J. Human interleukin 10 induces naive surface immunoglobulin D+ (sIgD+) B cells to secrete IgG1 and IgG3. *Journal of Experimental Medicine* **179**, 757–762 (1994).

221. Pelanda, R. & Torres, R. M. Central B-Cell Tolerance: Where Selection Begins. *Cold Spring Harb Perspect Biol* **4**, (2012).

222. Shlomchik, M. J. Sites and Stages of Autoreactive B Cell Activation and Regulation. *Immunity* **28**, 18–28 (2008).

223. Tiegs, S. L., Russell, D. M. & Nemazee, D. Receptor editing in self-reactive bone marrow B cells. *J Exp Med* **177**, 1009–1020 (1993).

224. Gay, D., Saunders, T., Camper, S. & Weigert, M. Receptor editing: an approach by autoreactive B cells to escape tolerance. *Journal of Experimental Medicine* **177**, 999–1008 (1993).

225. Nemazee, D. A. & Bürki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature 1989 337:6207* **337**, 562–566 (1989).

226. Nossal, G. J. V. & Pike, B. L. Clonal anergy: Persistence in tolerant mice of antigen-binding B lymphocytes incapable of responding to antigen or mitogen. *Proceedings of the National Academy of Sciences* **77**, 1602–1606 (1980).

227. Goodnow, C. C. Transgenic Mice and Analysis of B-Cell Tolerance. *http://dx.doi.org/10.1146/annurev.iy.10.040192.002421* **10**, 489–518 (2003).

228. Melchers, F. Checkpoints that control B cell development. *J Clin Invest* **125**, 2203–2210 (2015).

229. Hartley, S. B. *et al.* Elimination from peripheral lymphoid tissues of self-reactive B lymphocytes recognizing membrane-bound antigens. *Nature 1991 353:6346* **353**, 765–769 (1991).

230. Casellas, R. *et al.* Contribution of receptor editing to the antibody repertoire. *Science* **291**, 1541–1544 (2001).

231. Cooke, M. P. *et al.* Immunoglobulin signal transduction guides the specificity of B cell-T cell interactions and is blocked in tolerant self-reactive B cells. *Journal of Experimental Medicine* **179**, 425–438 (1994).

232. Eris, J. M. *et al.* Anergic self-reactive B cells present self antigen and respond normally to CD40-dependent T-cell signals but are defective in antigen-receptor-mediated functions. *Proceedings of the National Academy of Sciences* **91**, 4392–4396 (1994).

233. Klinman, N. R. The 'Clonal Selection Hypothesis' Minireview and Current Concepts of B Cell Tolerance. *Immunity* **5**, 189–195 (1996).

234. Cyster, J. G. & Goodnow, C. C. Antigen-induced exclusion from follicles and anergy are separate and complementary processes that influence peripheral B cell fate. *Immunity* **3**, 691–701 (1995).

235. Cyster, J. G., Hartley, S. B. & Goodnow, C. C. Competition for follicular niches excludes self-reactive cells from the recirculating B-cell repertoire. *Nature 1994 371:6496* **371**, 389–395 (1994).

236. Ekland, E. H., Forster, R., Lipp, M. & Cyster, J. G. Requirements for follicular exclusion and competitive elimination of autoantigen-binding B cells. *J Immunol* **172**, 4700–4708 (2004).

237. Liu, S. *et al.* Receptor Editing Can Lead to Allelic Inclusion and Development of B Cells That Retain Antibodies Reacting with High Avidity Autoantigens. *The Journal of Immunology* **175**, 5067–5076 (2005).

238. Kenny, J. J. *et al.* Autoreactive B Cells Escape Clonal Deletion by Expressing Multiple Antigen Receptors. *The Journal of Immunology* **164**, 4111–4119 (2000).

239. Rice, J. S., Newman, J., Wang, C., Michael, D. J. & Diamond, B. Receptor editing in peripheral B cell tolerance. *Proc Natl Acad Sci U S A* **102**, 1608–1613 (2005).

240. Brooks, J. F., Murphy, P. R., Barber, J. E. M., Wells, J. W. & Steptoe, R. J. Peripheral Tolerance Checkpoints Imposed by Ubiquitous Antigen Expression Limit Antigen-Specific B Cell Responses under Strongly Immunogenic Conditions. *The Journal of Immunology* **205**, 1239–1247 (2020).

241. Wang, Y. H. & Diamond, B. B cell receptor revision diminishes the autoreactive B cell response after antigen activation in mice. *J Clin Invest* **118**, 2896–2907 (2008).

242. Wang, H. & Shlomchik, M. J. Autoantigen-Specific B Cell Activation in FAS-Deficient Rheumatoid Factor Immunoglobulin Transgenic Mice. *Journal of Experimental Medicine* **190**, 639–650 (1999).

243. Koncz, G. & Hueber, A. O. The Fas/CD95 receptor regulates the death of autoreactive B cells and the selection of antigen-specific B cells. *Front Immunol* **3**, 207 (2012).

244. Espéli, M. *et al.* FcγRIIb differentially regulates pre-immune and germinal center B cell tolerance in mouse and human. *Nature Communications 2019 10:1* **10**, 1–14 (2019).

245. Yau, I. W. *et al.* Censoring of Self-Reactive B Cells by Follicular Dendritic Cell–Displayed Self-Antigen. *The Journal of Immunology* **191**, 1082–1090 (2013).

246. van der Poel, C. E. *et al.* Follicular Dendritic Cells Modulate Germinal Center B Cell Diversity through FcγRIIB. *Cell Rep* **29**, 2745-2755.e4 (2019).

247. Sage, P. T., Paterson, A. M., Lovitch, S. B. & Sharpe, A. H. The Coinhibitory Receptor CTLA-4 Controls B Cell Responses by Modulating T Follicular Helper, T Follicular Regulatory, and T Regulatory Cells. *Immunity* **41**, 1026–1039 (2014).

248. Honjo, T. *IMMUNOGLOBULIN GENES*. *Ann. Rev. ImmunoL* vol. 1 www.annualreviews.org (1983).

249. Charles A Janeway, J., Travers, P., Walport, M. & Shlomchik, M. J. The distribution and functions of immunoglobulin isotypes. (2001).

250. Lu, L. L., Suscovich, T. J., Fortune, S. M. & Alter, G. Beyond binding: antibody effector functions in infectious diseases. *Nature Reviews Immunology 2017 18:1* **18**, 46–61 (2017).

251. Boes, M. Role of natural and immune IgM antibodies in immune responses. *Mol Immunol* **37**, 1141–1149 (2000).

252. Avrameas, S. Natural autoantibodies: from 'horror autotoxicus' to 'gnothi seauton'. *Immunol Today* **12**, 154–159 (1991).

253. Cohen, I. R. & Norins, L. C. Natural Human Antibodies to Gram-Negative Bacteria: Immunoglobulins G, A, and M. *Science* **152**, 1257–1259 (1966).

254. Sharp, T. H. *et al.* Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. *Proc Natl Acad Sci U S A* **116**, 11900–11905 (2019).

255. Pone, E. J. *et al.* Toll-like Receptors and B-cell Receptors Synergize to Induce Immunoglobulin Class Switch DNA Recombination: Relevance to Microbial Antibody Responses. *Crit Rev Immunol* **30**, 1 (2010).

256. Shan, M. *et al.* Secreted IgD Amplifies Humoral T Helper 2 Cell Responses by Binding Basophils via Galectin-9 and CD44. *Immunity* **49**, 709-724.e8 (2018).

257. Vidarsson, G., Dekkers, G., Rispens, T. & Klinman, D. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol* (2014) doi:10.3389/fimmu.2014.00520.

258. Hjelholt, A., Christiansen, G., Sørensen, U. S. & Birkelund, S. IgG subclass profiles in normal human sera of antibodies specific to five kinds of microbial antigens. *Pathog Dis* **67**, 206–213 (2013).

259. Umetsu, D. T., Ambrosino, D. M., Quinti, I., Siber, G. R. & Geha, R. S. Recurrent sinopulmonary infection and impaired antibody response to bacterial capsular polysaccharide antigen in children with selective IgG-subclass deficiency. *N Engl J Med* **313**, 1247–51 (1985).

260. Bruhns, P. *et al.* Specificity and affinity of human Fcγ receptors and their polymorphic variants for human IgG subclasses. *Blood* **113**, 3716–3725 (2009).

261. Nirula, A., Glaser, S. M., Kalled, S. L. & Taylora, F. R. What is IgG4? A review of the biology of a unique immunoglobulin subtype. *Curr Opin Rheumatol* **23**, 119–124 (2011).

262. Cerutti, A. The regulation of IgA class switching. *Nature Reviews Immunology 2008 8:6* **8**, 421–434 (2008).

263. Luker, A. J., Lownik, J. C., Conrad, D. H. & Martin, R. K. A new look at IgE beyond allergies. *F1000Res* **8**, (2019).

264. Karagiannis, S. N. *et al.* Activity of human monocytes in IgE antibody-dependent surveillance and killing of ovarian tumor cells. *Eur J Immunol* **33**, 1030–1040 (2003).

265. Henault, J. *et al.* Self-reactive IgE exacerbates interferon responses associated with autoimmunity. *Nature Immunology 2015 17:2* **17**, 196–203 (2015).

266. Nybakken, G. E. *et al.* Structural basis of West Nile virus neutralization by a therapeutic antibody. *Nature 2005 437:7059* **437**, 764–769 (2005).

267. Kaur, R., Surendran, N., Ochs, M. & Pichichero, M. E. Human antibodies to PhtD, PcpA, and ply reduce adherence to human lung epithelial cells and murine nasopharyngeal colonization by Streptococcus pneumoniae. *Infect Immun* **82**, 5069–5075 (2014).

268. Menon, M. *et al.* Inactive disease in lupus patients is linked to autoantibodies to type-I interferons that normalize blood IFNα and B cell subsets. *medRxiv* 2021.04.07.21255049 (2021) doi:10.1101/2021.04.07.21255049.

269. Rosenberg, J. M. *et al.* Neutralizing anti-cytokine autoantibodies against interferon-a in immunodysregulation polyendocrinopathy enteropathy x-linked. *Front Immunol* **9**, 544 (2018).

270. Wang, E. Y. *et al.* Diverse functional autoantibodies in patients with COVID-19. *Nature 2021 595:7866* **595**, 283–288 (2021).

271. Merle, N. S., Noe, R., Halbwachs-Mecarelli, L., Fremeaux-Bacchi, V. & Roumenina, L. T. Complement system part II: Role in immunity. *Front Immunol* **6**, 257 (2015).

272. Zahavi, D., AlDeghaither, D., O'Connell, A. & Weiner, L. M. Enhancing antibody-dependent cell-mediated cytotoxicity: a strategy for improving antibody-based immunotherapy. *Antib Ther* **1**, 7–12 (2018).

273. Cao, X. *et al.* Promoting antibody-dependent cellular phagocytosis for effective macrophage-based cancer immunotherapy. *Sci Adv* **8**, 9171 (2022).

274. Tay, M. Z., Wiehe, K. & Pollara, J. Antibody dependent cellular phagocytosis in antiviral immune responses. *Front Immunol* **10**, 332 (2019).

275. Rodgers, C. B., Mustard, C. J., McLean, R. T., Hutchison, S. & Pritchard, A. L. A B-cell or a key player? The different roles of B-cells and antibodies in melanoma. *Pigment Cell Melanoma Res* **35**, 303 (2022).

276. Sharonov, G. v., Serebrovskaya, E. O., Yuzhakova, D. v., Britanova, O. v. & Chudakov, D. M. B cells, plasma cells and antibody repertoires in the tumour microenvironment. *Nature Reviews Immunology 2020 20:5* **20**, 294–307 (2020).

277. Tsou, P., Katayama, H., Ostrin, E. J. & Hanash, S. M. The emerging role of b cells in tumor immunity. *Cancer Res* **76**, 5591–5601 (2016).

278. Wouters, M. C. A. & Nelson, B. H. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clinical Cancer Research* **24**, 6125–6135 (2018).

279. Griss, J. *et al.* B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. *Nature Communications 2019 10:1* **10**, 1–14 (2019).

280. Lundberg, A., Li, B. & Li, R. B cell-related gene signature and cancer immunotherapy response. *British Journal of Cancer 2021 126:6* **126**, 899–906 (2021).

281. Sarvaria, A., Madrigal, J. A. & Saudemont, A. B cell regulation in cancer and anti-tumor immunity. *Cellular & Molecular Immunology 2017 14:8* **14**, 662–674 (2017).

282. Chen, J. *et al.* Single-cell transcriptome and antigen-immunoglobin analysis reveals the diversity of B cells in non-small cell lung cancer. *Genome Biol* **21**, 1–21 (2020).

283. Shi, J. Y. *et al.* Margin-infiltrating CD20+ B cells display an atypical memory phenotype and correlate with favorable prognosis in hepatocellular carcinoma. *Clinical Cancer Research* **19**, 5994–6005 (2013).

284. Bruno, T. C. *et al.* Antigen-presenting intratumoral B cells affect CD4+ TIL phenotypes in non–small cell lung cancer patients. *Cancer Immunol Res* **5**, 898–907 (2017).

285. Wennhold, K. *et al.* Cd86 antigen-presenting b cells are increased in cancer, localize in tertiary lymphoid structures, and induce specific t-cell responses. *Cancer Immunol Res* **9**, 1098–1108 (2021).

286. Germain, C. *et al.* Presence of B cells in tertiary lymphoid structures is associated with a protective immunity in patients with lung cancer. *Am J Respir Crit Care Med* **189**, 832–844 (2014).

287. Montfort, A. *et al.* A strong B-cell response is part of the immune landscape in human high-grade serous ovarian metastases. *Clinical Cancer Research* **23**, 250–262 (2017).

288. Mazor, R. D. *et al.* Tumor-reactive antibodies evolve from non-binding and autoreactive precursors. *Cell* **185**, 1208-1222.e21 (2022).

289. Su, S. *et al.* Immune Checkpoint Inhibition Overcomes ADCP-Induced Immunosuppression by Macrophages. *Cell* **175**, 442-457.e23 (2018).

290. Biswas, S. *et al.* IgA transcytosis and antigen recognition govern ovarian cancer immunity. *Nature 2021 591:7850* **591**, 464–470 (2021).

291. Michaud, D., Steward, C. R., Mirlekar, B. & Pylayeva-Gupta, Y. Regulatory B cells in cancer. *Immunol Rev* **299**, 74 (2021).

292. Xiao, X. *et al.* PD-1hi identifies a novel regulatory b-cell population in human hepatoma that promotes disease progression. *Cancer Discov* **6**, 546–559 (2016).

293. Wang, W. *et al.* CD19+CD24hiCD38hiBregs involved in downregulate helper T cells and upregulate regulatory T cells in gastric cancer. *Oncotarget* **6**, 33486 (2015).

294. Ishigami, E. *et al.* Coexistence of regulatory B cells and regulatory T cells in tumor-infiltrating lymphocyte aggregates is a prognostic factor in patients with breast cancer. *Breast Cancer* **26**, 180–189 (2019).

295. Murakami, Y. *et al.* Increased regulatory B cells are involved in immune evasion in patients with gastric cancer. *Scientific Reports 2019 9:1* **9**, 1–9 (2019).

296. Zhou, X., Su, Y. X., Lao, X. M., Liang, Y. J. & Liao, G. Q. CD19+IL-10+ regulatory B cells affect survival of tongue squamous cell carcinoma patients and induce resting CD4+ T cells to CD4+Foxp3+ regulatory T cells. *Oral Oncol* **53**, 27–35 (2016).

297. Wei, X. *et al.* Regulatory B cells contribute to the impaired antitumor immunity in ovarian cancer patients. *Tumor Biology* **37**, 6581–6588 (2016).

298. Toney, N. J. *et al.* Tumor-B-cell interactions promote isotype switching to an immunosuppressive IgG4 antibody response through upregulation of IL-10 in triple negative breast cancers. *J Transl Med* **20**, 1–15 (2022).

299. Karagiannis, P. *et al.* IgG4 subclass antibodies impair antitumor immunity in melanoma. *J Clin Invest* **123**, 1457–1474 (2013).

300. Somasundaram, R. *et al.* Tumor-associated B-cells induce tumor heterogeneity and therapy resistance. *Nat Commun* **8**, (2017).

301. Sexauer, D., Gray, E. & Zaenker, P. Tumour- associated autoantibodies as prognostic cancer biomarkers- a review. *Autoimmun Rev* **21**, 103041 (2022).

302. Monroy-Iglesias, M. J. *et al.* Antibodies as biomarkers for cancer risk: a systematic review. *Clin Exp Immunol* **209**, 46–63 (2022).

303. Yang, B., Li, X., Ren, T. & Yin, Y. Autoantibodies as diagnostic biomarkers for lung cancer: A systematic review. *Cell Death Discovery 2019 5:1* **5**, 1–15 (2019).

304. Enhancement of Antibody Detection in Cancer Using Panel of Recombinant Tumor-associated Antigens1 | Cancer Epidemiology, Biomarkers & Prevention | American Association for Cancer Research. https://aacrjournals.org/cebp/article/12/2/136/253485/Enhancement-of-Antibody-Detection-in-Cancer-Using.

305. Lubin, R. *et al.* Serum p53 antibodies as early markers of lung cancer. *Nature Medicine 1995 1:7* **1**, 701–702 (1995).

306. Zhang, J. Y. *et al.* Antibody detection using tumor-associated antigen mini-array in immunodiagnosing human hepatocellular carcinoma. *J Hepatol* **46**, 107–114 (2007).

307. Jiang, D. *et al.* Discovering Panel of Autoantibodies for Early Detection of Lung Cancer Based on Focused Protein Array. *Front Immunol* **12**, 710 (2021).

308. p53 Antibodies in the Sera of Patients with Various Types of Cancer: A Review1 | Cancer Research | American Association for Cancer Research.

https://aacrjournals.org/cancerres/article/60/7/1777/507199/p53-Antibodies-in-the-Sera-of-Patients-with.

309. Kijanka, G. *et al.* Human IgG antibody profiles differentiate between symptomatic patients with and without colorectal cancer. *Gut* **59**, 69–78 (2010).

310. Garziera, M. *et al.* Prognostic Role of Serum Antibody Immunity to p53 Oncogenic Protein in Ovarian Cancer: A Systematic Review and a Meta-Analysis. *PLoS One* **10**, e0140351 (2015).

311. Anderson, K. S. *et al.* p53 autoantibodies as potential detection and prognostic biomarkers in serous ovarian cancer. *Cancer Epidemiology Biomarkers and Prevention* **19**, 859–868 (2010).

312. Sobhani, N. *et al.* p53 Antibodies as a Diagnostic Marker for Cancer: A Meta-Analysis. *Molecules* **26**, (2021).

313. Fiorentino, D. F. *et al.* Immune responses to CCAR1 and other dermatomyositis autoantigens are associated with attenuated cancer emergence. *J Clin Invest* **132**, (2022).

314. Carbonell, C. *et al.* Standardized incidence ratios and risk factors for cancer in patients with systemic sclerosis: Data from the Spanish Scleroderma Registry (RESCLE). *Autoimmun Rev* **21**, (2022).

315. Duarte, J. D. G. *et al.* Autoantibodies May Predict Immune-Related Toxicity: Results from a Phase I Study of Intralesional Bacillus Calmette-Guérin followed by Ipilimumab in Patients with Advanced Metastatic Melanoma. *Front Immunol* **9**, (2018).

316. de Moel, E. C. *et al.* Autoantibody development under treatment with immune-checkpoint inhibitors. *Cancer Immunol Res* **7**, 6–11 (2019).

317. Ghosh, N. *et al.* Lower baseline autoantibody levels are associated with immune-related adverse events from immune checkpoint inhibition. *J Immunother Cancer* **10**, e004008 (2022).

318. Alserawan, L. *et al.* Association between Changes in the Patterns of Antinuclear Autoantibodies during Immune Checkpoint Inhibition Therapy and the Development of Severe Immune Related Adverse Events. *Int J Mol Sci* **23**, 12641 (2022).

319. Study of Anti-PD-L1 in Combination With Chemo(Radio)Therapy for Oesophageal Cancer - Full Text View - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT02735239.

320. Kato, K. *et al.* Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol* **20**, 1506–1517 (2019).

321. US Department of Health and Human Services. Common Terminology Criteria for Adverse Events (CTCAE) Version 4.0.3. (2010).

322. Kim, J. *et al.* Integrated genomic characterization of oesophageal carcinoma. *Nature 2017 541:7636* **541**, 169–175 (2017).

323. Seymour, L. *et al.* iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol* **18**, e143 (2017).

324. Mentzer, A. J. *et al.* Identification of host-pathogen-disease relationships using a scalable multiplex serology platform in UK Biobank. *Nat Commun* **13**, (2022).

325. Dejnirattisai, W. *et al.* A new class of highly potent, broadly neutralizing antibodies isolated from viremic patients infected with dengue virus. *Nature Immunology 2014 16:2* **16**, 170–177 (2014).

326. Wardemann, H. & Busse, C. E. Expression cloning of antibodies from single human B cells. *Methods in Molecular Biology* **1956**, 105–125 (2019).

327. Scrucca, L., Fop, M., Murphy, T. B. & Raftery, A. E. mclust 5: Clustering, Classification and Density Estimation Using Gaussian Finite Mixture Models. *R J* **8**, 289 (2016).

328. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at https://www.R-project.org/ (2021).

329. Wickham, H. ggplot2. (2016) doi:10.1007/978-3-319-24277-4.

330. Ashland & Becton Dickinson and Company. FlowJo™ Software (for Windows). Preprint at https://www.flowjo.com/citing-flowjo (2021).

331. Azad, A., Rajwa, B. & Pothen, A. flowVS: Channel-specific variance stabilization in flow cytometry. *BMC Bioinformatics* **17**, 1–14 (2016).

332. Emmaneel, A. *et al.* PeacoQC: Peak-based selection of high quality cytometry data. *Cytometry Part A* **101**, 325–338 (2022).

333. van Gassen, S., Gaudilliere, B., Angst, M. S., Saeys, Y. & Aghaeepour, N. CytoNorm: A Normalization Algorithm for Cytometry Data. *Cytometry Part A* **97**, 268–278 (2020).

334. Nowicka, M. *et al.* CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. *F1000Research 2019 6:748* **6**, 748 (2019).

335. Quintelier, K. *et al.* Analyzing high-dimensional cytometry data using FlowSOM. *Nat Protoc* doi:10.1038/s41596-021-00550-0.

336. Weber, L. M., Nowicka, M., Soneson, C. & Robinson, M. D. diffcyt: Differential discovery in high-dimensional cytometry via high-resolution clustering. *Communications Biology 2019 2:1* **2**, 1–11 (2019).

337. di Tommaso, P. *et al.* Nextflow enables reproducible computational workflows. *Nature Biotechnology 2017 35:4* **35**, 316–319 (2017).

338. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

339. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

340. Becht, E. *et al.* Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* **17**, 1–20 (2016).

341. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

342. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* **20**, 1–15 (2019).

343. Haghverdi, L., Lun, A. T. L., Morgan, M. D. & Marioni, J. C. Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat Biotechnol* **36**, 421–427 (2018).

344. Gupta, N. T. *et al.* Change-O: a toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics* **31**, 3356–3358 (2015).

345. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (1995).

346. Mietzner, B. *et al.* Autoreactive IgG memory antibodies in patients with systemic lupus erythematosus arise from nonreactive and polyreactive precursors. *Proceedings of the National Academy of Sciences* **105**, 9727–9732 (2008).

347. Richardson, C. *et al.* Molecular basis of 9G4 B cell autoreactivity in human systemic lupus erythematosus. *J Immunol* **191**, 4926–4939 (2013).

348. Tipton, C. M. *et al.* Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nature Immunology 2015 16:7* **16**, 755–765 (2015).

349. Yurasov, S. *et al.* Defective B cell tolerance checkpoints in systemic lupus erythematosus. *Journal of Experimental Medicine* **201**, 703–711 (2005).

350. Wardemann, H. *et al.* Predominant autoantibody production by early human B cell precursors. *Science (1979)* **301**, 1374–1377 (2003).

351. Meffre, E. & O'Connor, K. C. Impaired B-cell tolerance checkpoints promote the development of autoimmune diseases and pathogenic autoantibodies. *Immunol Rev* **292**, 90–101 (2019).

352. Moody, M. A. *et al.* Immune perturbations in HIV-1-infected individuals who make broadly neutralizing antibodies. *Sci Immunol* **1**, (2016).

353. Campochiaro, C. *et al.* Autoantibody positivity predicts severity of rheumatic immune-related adverse events to immune-checkpoint inhibitors. *Eur J Intern Med* **103**, 95–99 (2022).

354. Barth, D. A. *et al.* Evaluation of autoantibodies as predictors of treatment response and immune-related adverse events during the treatment with immune checkpoint inhibitors: A prospective longitudinal pan-cancer study. *Cancer Med* **11**, 3074–3083 (2022).

355. Wang, E. Y. *et al.* High-throughput identification of autoantibodies that target the human exoproteome. (2022) doi:10.1016/j.crmeth.2022.100172.

356. Ayoglu, B., Schwenk, J. M. & Nilsson, P. Antigen arrays for profiling autoantibody repertoires. *Bioanalysis* **8**, 1105–1126 (2016).

357. Schirwitz, C. *et al.* Sensing Immune Responses with Customized Peptide Microarrays. *Biointerphases* **7**, 47 (2012).

358. Larman, H. B. *et al.* Autoantigen discovery with a synthetic human peptidome. *Nat Biotechnol* **29**, 535–541 (2011).

359. Autoimmunity and Serology Profiling - SciLifeLab. https://www.scilifelab.se/units/autoimmunity-profiling/.

360. Martins, T. B. *et al.* Evaluation of multiplexed fluorescent microsphere immunoassay for detection of autoantibodies to nuclear antigens. *Clin Diagn Lab Immunol* **11**, 1054–1059 (2004).

361. Martins, T. B., Litwin, C. M. & Hill, H. R. Evaluation of a Multiplex Fluorescent Microsphere Immunoassay for the Determination of Epstein-Barr Virus Serologic Status. *Am J Clin Pathol* **129**, 34–41 (2008).

362. Jia, J. *et al.* Development of a Multiplex Autoantibody Test for Detection of Lung Cancer. *PLoS One* **9**, e95444 (2014).

363. Neiman, M. *et al.* Individual and stable autoantibody repertoires in healthy individuals. *Autoimmunity* **52**, 1–11 (2019).

364. Zaenker, P., Gray, E. S. & Ziman, M. R. Autoantibody Production in Cancer—The Humoral Immune Response toward Autologous Antigens in Cancer Patients. *Autoimmun Rev* **15**, 477–483 (2016).

365. Burbelo, P. D., Iadarola, M. J., Keller, J. M. & Warner, B. M. Autoantibodies Targeting Intracellular and Extracellular Proteins in Autoimmunity. *Front Immunol* **12**, 5 (2021).

366. Hua, Z. & Hou, B. TLR signaling in B-cell development and activation. *Cellular & Molecular Immunology 2013 10:2* **10**, 103–106 (2012).

367. Munoz, L. E. *et al.* Apoptosis in the pathogenesis of systemic lupus erythematosus. *Lupus* **17**, 371–375 (2008).

368. Brown, G. J. *et al.* TLR7 gain-of-function genetic variation causes human lupus. *Nature 2022 605:7909* **605**, 349–356 (2022).

369. Welsh, J. B. *et al.* Large-scale delineation of secreted protein biomarkers overexpressed in cancer tissue and serum. *Proc Natl Acad Sci U S A* **100**, 3410–3415 (2003).

370. Karsch-Bluman, A. *et al.* Tissue necrosis and its role in cancer progression. *Oncogene 2018 38:11* **38**, 1920–1935 (2018).

371. Clendenen, T. v. *et al.* Temporal reproducibility of IgG and IgM autoantibodies in serum from healthy women. *Scientific Reports 2022 12:1* **12**, 1–9 (2022).

372. Wu, M. *et al.* High-throughput Luminex xMAP assay for simultaneous detection of antibodies against rabbit hemorrhagic disease virus, Sendai virus and rabbit rotavirus. *Arch Virol* **164**, 1639–1646 (2019).

373. Ubillos, I., Campo, J. J., Jiménez, A. & Dobaño, C. Development of a high-throughput flexible quantitative suspension array assay for IgG against multiple Plasmodium falciparum antigens. *Malar J* **17**, 1–15 (2018).

374. Chen, Z., Dodig-Crnković, T., Schwenk, J. M. & Tao, S.-C. Current applications of antibody microarrays. *Clin Proteomics* **15**, 7 (2014).

375. Chan, Y. *et al.* Determining seropositivity—A review of approaches to define population seroprevalence when using multiplex bead assays to assess burden of tropical diseases. *PLoS Negl Trop Dis* **15**, e0009457 (2021).

376. Domingues, T. D., Mouriño, H. & Sepúlveda, N. Analysis of antibody data using Finite Mixture Models based on Scale Mixtures of Skew-Normal distributions. *medRxiv* 2021.03.08.21252807 (2021) doi:10.1101/2021.03.08.21252807.

377. Joseph, H. *et al.* Investigation of Mixture Modelling Algorithms as a Tool for Determining the Statistical Likelihood of Serological Exposure to Filariasis Utilizing Historical Data from the Lymphatic Filariasis Surveillance Program in Vanuatu. *Trop Med Infect Dis* **4**, 45 (2019).

378. Ster, I. C. Inference for serological surveys investigating past exposures to infections resulting in long-lasting immunity – an approach using finite mixture models with concomitant information. *http://dx.doi.org/10.1080/02664763.2012.722608* **39**, 2523–2542 (2012).

379. Bottomley, C. *et al.* Quantifying previous SARS-CoV-2 infection through mixture modelling of antibody levels. *Nat Commun* **12**, (2021).

380. Verma, R. *et al.* Lymphocyte depletion and repopulation after chemotherapy for primary breast cancer. *Breast Cancer Research* **18**, 1–12 (2016).

381. Wijayahadi, N., Haron, M. R., Stanslas, J. & Yusuf, Z. Changes in Cellular Immunity during Chemotherapy for Primary Breast Cancer with Anthracycline Regimens. *http://dx.doi.org/10.1179/joc.2007.19.6.716* **19**, 716–723 (2013).

382. p53 Antibodies in the Sera of Patients with Various Types of Cancer: A Review1 | Cancer Research | American Association for Cancer Research. https://aacrjournals.org/cancerres/article/60/7/1777/507199/p53-Antibodies-in-the-Sera-of-Patients-with.

383. Cusick, M. F., Libbey, J. E. & Fujinami, R. S. Molecular Mimicry as a Mechanism of Autoimmune Disease. *Clin Rev Allergy Immunol* **42**, 102 (2012).

384. Bjornevik, K. *et al.* Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science (1979)* **375**, 296–301 (2022).

385. Lanz, T. v. *et al.* Clonally expanded B cells in multiple sclerosis bind EBV EBNA1 and GlialCAM. *Nature 2022 603:7900* **603**, 321–327 (2022).

386. Shome, M. *et al.* Serum autoantibodyome reveals that healthy individuals share common autoantibodies. *Cell Rep* **39**, (2022).

387. Suppiah, A. & Greenman, J. Clinical utility of anti-p53 auto-antibody: Systematic review and focus on colorectal cancer. *World Journal of Gastroenterology : WJG* **19**, 4651 (2013).

388. Saeys, Y., van Gassen, S. & Lambrecht, B. N. Computational flow cytometry: helping to make sense of high-dimensional immunology data. *Nature Reviews Immunology 2016 16:7* **16**, 449–462 (2016).

389. Ahern, D. J. *et al.* A blood atlas of COVID-19 defines hallmarks of disease severity and specificity. *Cell* **185**, 916-938.e58 (2022).

390. CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets. https://www.bioconductor.org/packages/release/workflows/vignettes/cytofWorkflow/inst/doc/cytofWorkflow.html.

391. Jensen, G. *et al.* Selective Expression of CD45 Isoforms Defines CALLA+ Monoclonal B-Lineage Cells in Peripheral Blood From Myeloma Patients as Late Stage B Cells. *Blood* **78**, 711–719 (1991).

392. Brohée, D. & Higuet, N. In vitro stimulation of peripheral blood mononuclear cells by phytohaemagglutinin A induces CD45RA expression on monocytes. *Cytobios* **71**, 105–111 (1992).

393. Nagasawa, T. *et al.* Expression of CD14, CD16 and CD45RA on monocytes from periodontitis patients. *J Periodontal Res* **39**, 72–78 (2004).

394. Collin, M. & Bigley, V. Human dendritic cell subsets: an update. *Immunology* **154**, 3–20 (2018).

395. Henneken, M., Dörner, T., Burmester, G.-R. & Berek, C. Differential expression of chemokine receptors on peripheral blood B cells from patients with rheumatoid arthritis and systemic lupus erythematosus. *Arthritis Res Ther* **7**, R1001 (2005).

396. Wu, Z. *et al.* CD3+CD4-CD8- (Double-Negative) T Cells in Inflammation, Immune Disorders and Cancer. *Front Immunol* **13**, 388 (2022).

397. Overgaard, N. H., Jung, J.-W., Steptoe, R. J. & Wells, J. W. CD4+/CD8+ double-positive T cells: more than just a developmental stage? *J Leukoc Biol* **97**, 31–38 (2015).

398. Milush, J. M. *et al.* Functionally distinct subsets of human NK cells and monocyte/DC-like cells identified by coexpression of CD56, CD7, and CD4. *Blood* **114**, 4823 (2009).

399. Guo, R. *et al.* Lymphocyte mass cytometry identifies a CD3–CD4+ cell subset with a potential role in psoriasis. *JCI Insight* **4**, (2019).

400. Sanz, I. *et al.* Challenges and opportunities for consistent classification of human b cell and plasma cell populations. *Front Immunol* **10**, 2458 (2019).

401. Bhat, N. M., Kshirsagar, M. A., Bieber, M. M. & Teng, N. N. H. IgG Subclasses and Isotypes of VH4-34 Encoded Antibodies. *Immunol Invest* **44**, 400–410 (2015).

402. Schickel, J. N. *et al.* Self-reactive VH4-34–expressing IgG B cells recognize commensal bacteria. *J Exp Med* **214**, 1991 (2017).

403. Myles, A., Sanz, I. & Cancro, M. P. T-bet+ B cells: A common denominator in protective and autoreactive antibody responses? *Curr Opin Immunol* **57**, 40–45 (2019).

404. Geginat, J. *et al.* Plasticity of human CD4 T cell subsets. *Front Immunol* **5**, 630 (2014).

405. Schmitt, N., Bentebibel, S. E. & Ueno, H. Phenotype and functions of memory Tfh cells in human blood. *Trends Immunol* **35**, 436–442 (2014).

406. Gosselin, A. *et al.* Peripheral blood CCR4+CCR6+ and CXCR3+CCR6+CD4+ T cells are highly permissive to HIV-1 infection. *J Immunol* **184**, 1604–1616 (2010).

407. Anaparti, V. *et al.* Increased frequency of TIGIT+ CD4 T Cell subset in autoantibody-positive first-degree relatives of patients with rheumatoid arthritis. *Front Immunol* **13**, 4218 (2022).

408. Takeuchi, A. & Saito, T. CD4 CTL, a Cytotoxic Subset of CD4+ T Cells, Their Differentiation and Function. *Front Immunol* **8**, (2017).

409. Groom, J. R. & Luster, A. D. CXCR3 in T cell function. *Exp Cell Res* **317**, 620 (2011).

410. Tian, Y. *et al.* Unique phenotypes and clonal expansions of human CD4 effector memory T cells re-expressing CD45RA. *Nat Commun* **8**, (2017).

411. Churlaud, G. *et al.* Human and Mouse CD8+CD25+FOXP3+ Regulatory T Cells at Steady State and during Interleukin-2 Therapy. *Front Immunol* **6**, (2015).

412. Nolz, J. C. Molecular Mechanisms of CD8+ T cell Trafficking and Localization. *Cell Mol Life Sci* **72**, 2461 (2015).

413. Elzein, S. M., Zimmerer, J. M., Han, J. L., Ringwald, B. A. & Bumgardner, G. L. CXCR5+CD8+ T Cells: A Review of Their Antibody Regulatory Functions and Clinical Correlations. *The Journal of Immunology* **206**, 2775–2783 (2021).

414. Appay, V. *et al.* Memory CD8+ T cells vary in differentiation phenotype in different persistent virus infections. *Nature Medicine 2002 8:4* **8**, 379–385 (2002).

415. Martin, M. D. & Badovinac, V. P. Defining memory CD8 T cell. *Front Immunol* **9**, 2692 (2018).

416. Mahnke, Y. D., Brodie, T. M., Sallusto, F., Roederer, M. & Lugli, E. The who's who of T-cell differentiation: Human memory T-cell subsets. *Eur J Immunol* **43**, 2797–2809 (2013).

417. Lee, S. W. *et al.* CD8+ TILs in NSCLC differentiate into TEMRA via a bifurcated trajectory: deciphering immunogenicity of tumor antigens. *J Immunother Cancer* **9**, e002709 (2021).

418. Koch, S. *et al.* Multiparameter flow cytometric analysis of CD4 and CD8 T cell subsets in young and old people. *Immunity and Ageing* **5**, 1–12 (2008).

419. Perdomo-Celis, F., Taborda, N. A. & Rugeles, M. T. Circulating CXCR5-expressing CD8+ T-cells are major producers of IL-21 and associate with limited HIV replication. *J Acquir Immune Defic Syndr (1988)* **78**, 473–482 (2018).

420. Valentine, K. M. & Hoyer, K. K. CXCR5+ CD8 T cells: Protective or pathogenic? *Front Immunol* **10**, 1322 (2019).

421. Kondo, T. & Takiguchi, M. Human memory CCR4+CD8+ T cell subset has the ability to produce multiple cytokines. *Int Immunol* **21**, 523–532 (2009).

422. Srenathan, U., Steel, K. & Taams, L. S. IL-17+ CD8+ T cells: Differentiation, phenotype and role in inflammatory disease. *Immunol Lett* **178**, 20 (2016).

423. Kondo, T., Takata, H. & Takiguchi, M. Functional expression of chemokine receptor CCR6 on human effector memory CD8+ T cells. *Eur J Immunol* **37**, 54–65 (2007).

424. Sheu, B. C. *et al.* Up-regulation of Inhibitory Natural Killer Receptors CD94/NKG2A with Suppressed Intracellular Perforin Expression of Tumor-Infiltrating CD8+ T Lymphocytes in Human Cervical Carcinoma. *Cancer Res* **65**, 2921–2929 (2005).

425. Speiser, D. E. *et al.* CD28-negative cytolytic effector T cells frequently express NK receptors and are present at variable proportions in circulating lymphocytes from healthy donors and melanoma patients. doi:10.1002/(SICI)1521-4141(199906)29:06.

426. Pereira, B. I. *et al.* Senescent cells evade immune clearance via HLA-E-mediated NK and CD8+ T cell inhibition. *Nature Communications 2019 10:1* **10**, 1–13 (2019).

427. Borst, L. *et al.* NKG2A is a late immune checkpoint on CD8 T cells and marks repeated stimulation and cell division. *Int J Cancer* **150**, 688–704 (2022).

428. Sottile, R. *et al.* Human cytomegalovirus expands a CD8+ T cell population with loss of BCL11B expression and gain of NK cell identity. *Sci Immunol* **6**, (2021).

429. Sottile, R. *et al.* Human cytomegalovirus expands a CD8+ T cell population with loss of BCL11B expression and gain of NK cell identity. *Sci Immunol* **6**, 6968 (2021).

430. Kovalenko, E. I. *et al.* Surface NKG2C Identifies Differentiated αβT-Cell Clones Expanded in Peripheral Blood. *Front Immunol* **11**, 3875 (2021).

431. Savage, A. K. *et al.* The Transcription Factor PLZF Directs the Effector Program of the NKT Cell Lineage. *Immunity* **29**, 391–403 (2008).

432. Abel, A. M., Yang, C., Thakar, M. S. & Malarkannan, S. Natural killer cells: Development, maturation, and clinical utilization. *Front Immunol* **9**, 1869 (2018).

433. Shao, J. Y. *et al.* Siglec-7 Defines a Highly Functional Natural Killer Cell Subset and Inhibits Cell-Mediated Activities. *Scand J Immunol* **84**, 182–190 (2016).

434. Freud, A. G., Mundy-Bosse, B. L., Yu, J. & Caligiuri, M. A. The broad spectrum of human natural killer cell diversity. *Immunity* **47**, 820 (2017).

435. Caligiuri, M. A. Human natural killer cells. *Blood* **112**, 461–469 (2008).

436. Björkström, N. K., Ljunggren, H. G. & Sandberg, J. K. CD56 negative NK cells: origin, function, and role in chronic viral disease. *Trends Immunol* **31**, 401–406 (2010).

437. di Vito, C., Mikulak, J. & Mavilio, D. On the way to become a natural killer cell. *Front Immunol* **10**, 1812 (2019).

438. Béziat, V., Descours, B., Parizot, C., Debré, P. & Vieillard, V. NK Cell Terminal Differentiation: Correlated Stepwise Decrease of NKG2A and Acquisition of KIRs. *PLoS One* **5**, e11966 (2010).

439. Moreira, A. *et al.* Adaptive Features of Natural Killer Cells in Multiple Sclerosis. *Front Immunol* **10**, 2403 (2019).

440. Lopez-Vergès, S. *et al.* Expansion of a unique CD57 +NKG2C hi natural killer cell subset during acute human cytomegalovirus infection. *Proc Natl Acad Sci U S A* **108**, 14725–14732 (2011).

441. Gumá, M. *et al.* Imprint of human cytomegalovirus infection on the NK cell receptor repertoire. *Blood* **104**, 3664–3671 (2004).

442. Béziat, V. *et al.* CMV drives clonal expansion of NKG2C+ NK cells expressing self-specific KIRs in chronic hepatitis patients. *Eur J Immunol* **42**, 447–457 (2012).

443. Ziegler-Heitbrock, L. *et al.* Nomenclature of monocytes and dendritic cells in blood. *Blood* **116**, e74–e80 (2010).

444. Mengos, A. E., Gastineau, D. A. & Gustafson, M. P. The CD14+HLA-DrlO/NEG monocyte: An immunosuppressive phenotype that restrains responses to cancer immunotherapy. *Front Immunol* **10**, 1147 (2019).

445. Rhodes, J. W., Tong, O., Harman, A. N. & Turville, S. G. Human dendritic cell subsets, ontogeny, and impact on HIV infection. *Front Immunol* **10**, 1088 (2019).

446. Burns, M. *et al.* Dysregulated CD38 Expression on Peripheral Blood Immune Cell Subsets in SLE. *Int J Mol Sci* **22**, 1–19 (2021).

447. Fedele, G. *et al.* CD38 is expressed on human mature monocyte-derived dendritic cells and is functionally involved in CD83 expression and IL-12 induction. *Eur J Immunol* **34**, 1342–1350 (2004).

448. Tipton, C. M. *et al.* Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nature Immunology 2015 16:7* **16**, 755–765 (2015).

449. Ruschil, C. *et al.* Specific Induction of Double Negative B Cells During Protective and Pathogenic Immune Responses. *Front Immunol* **11**, 3304 (2020).

450. Koncz, G. & Hueber, A. O. The Fas/CD95 receptor regulates the death of autoreactive B cells and the selection of antigen-specific B cells. *Front Immunol* **3**, 207 (2012).

451. Clark, E. A. & Giltiay, N. v. CD22: A Regulator of Innate and Adaptive B Cell Responses and Autoimmunity. *Front Immunol* **9**, 2235 (2018).

452. Meyer, S. J. *et al.* CD22 Controls Germinal Center B Cell Receptor Signaling, Which Influences Plasma Cell and Memory B Cell Output. *J Immunol* **207**, 1018–1032 (2021).

453. Monteiro, R. C. & van de Winkel, J. G. J. IgA Fc Receptors. *https://doi.org/10.1146/annurev.immunol.21.120601.141011* **21**, 177–204 (2003).

454. Bakema, J. E. & van Egmond, M. The human immunoglobulin A Fc receptor FcαRI: a multifaceted regulator of mucosal immunity. *Mucosal Immunology 2011 4:6* **4**, 612–624 (2011).

455. Junker, F., Gordon, J. & Qureshi, O. Fc Gamma Receptors and Their Role in Antigen Uptake, Presentation, and T Cell Activation. *Front Immunol* **11**, 1393 (2020).

456. Jenks, S. A. *et al.* Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus. *Immunity* **49**, 725-739.e6 (2018).

457. Wangriatisak, K. *et al.* The expansion of activated naive DNA autoreactive B cells and its association with disease activity in systemic lupus erythematosus patients. *Arthritis Res Ther* **23**, 1–11 (2021).

458. Larbi, A. & Fulop, T. From “truly naïve” to “exhausted senescent” T cells: When markers predict functionality. *Cytometry Part A* **85**, 25–35 (2014).

459. Tomiyama, H., Takata, H., Matsuda, T. & Takiguchi, M. Phenotypic classification of human CD8+ T cells reflecting their function: inverse correlation between quantitative expression of CD27 and cytotoxic effector function. *Eur J Immunol* **34**, 999–1010 (2004).

460. Boelen, L. *et al.* Inhibitory killer-cell immunoglobulin-like receptors strengthen CD8+ T cell-mediated control of HIV-1, HCV and HTLV-1. *Sci Immunol* **3**, (2018).

461. Speiser, D. E. *et al.* Human CD8 + T cells expressing HLA-DR and CD28 show telomerase activity and are distinct from cytolytic effector T cells. doi:10.1002/1521-4141.

462. Hao, Z. *et al.* Fas Receptor Expression in Germinal-Center B Cells Is Essential for T and B Lymphocyte Homeostasis. *Immunity* **29**, 615 (2008).

463. Pipi, E. *et al.* Tertiary lymphoid structures: Autoimmunity goes local. *Front Immunol* **9**, 1952 (2018).

464. Fridman, W. H. *et al.* B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nature Reviews Clinical Oncology 2022 19:7* **19**, 441–457 (2022).

465. Simon, Q. *et al.* In-depth characterization of CD24highCD38high transitional human B cells reveals different regulatory profiles. *Journal of Allergy and Clinical Immunology* **137**, 1577-1584.e10 (2016).

466. Blair, P. A. *et al.* CD19+CD24hiCD38hi B Cells Exhibit Regulatory Capacity in Healthy Individuals but Are Functionally Impaired in Systemic Lupus Erythematosus Patients. *Immunity* **32**, 129–140 (2010).

467. Patel, A. J. *et al.* Regulatory B cell repertoire defects predispose lung cancer patients to immune-related toxicity following checkpoint blockade. *Nature Communications 2022 13:1* **13**, 1–16 (2022).

468. Das, R. *et al.* Early B cell changes predict autoimmunity following combination immune checkpoint blockade. *J Clin Invest* **128**, 715–720 (2018).

469. Thomas, G. J., Hart, I. R., Speight, P. M. & Marshall, J. F. Binding of TGF-β1 latency-associated peptide (LAP) to αvβ6 integrin modulates behaviour of squamous carcinoma cells. *British Journal of Cancer 2002 87:8* **87**, 859–867 (2002).

470. Taylor, A. W. Review of the activation of TGF-β in immunity. *J Leukoc Biol* **85**, 29 (2009).

471. Yang, L., Pang, Y. & Moses, H. L. TGF-β and immune cells: an important regulatory axis in the tumor microenvironment and progression. *Trends Immunol* **31**, 220 (2010).

472. Tschernia, N. P. & Gulley, J. L. Tumor in the Crossfire: Inhibiting TGF-β to Enhance Cancer Immunotherapy. *Biodrugs* **36**, 153 (2022).

473. Oh, D. Y. & Fong, L. Cytotoxic CD4+ T cells in cancer: Expanding the immune effector toolbox. *Immunity* **54**, 2701–2711 (2021).

474. Kar, A., Mehrotra, S. & Chatterjee, S. CD38: T Cell Immuno-Metabolic Modulator. *Cells* **9**, (2020).

475. Hu, Z. *et al.* The Role of KLRG1 in Human CD4+ T-Cell Immunity Against Tuberculosis. *J Infect Dis* **217**, 1491–1503 (2018).

476. Wang, J. *et al.* LAG-3 Represents a Marker of CD4+ T Cells with Regulatory Activity in Patients with Bone Fracture. *Immunol Invest* **47**, 492–503 (2018).

477. Miyara, M. *et al.* Sialyl Lewis x (CD15s) identifies highly differentiated and most suppressive FOXP3high regulatory T cells in humans. *Proc Natl Acad Sci U S A* **112**, 7225–7230 (2015).

478. Dei Zotti, F., Moriconi, C., Qiu, A., Miller, A. & Hudson, K. E. Distinct CD4+ T cell signature in ANA-positive young adult patients. *Front Immunol* **13**, (2022).

479. Bakema, J. E. & van Egmond, M. The human immunoglobulin A Fc receptor FcαRI: a multifaceted regulator of mucosal immunity. *Mucosal Immunology 2011 4:6* **4**, 612–624 (2011).

480. Moons, L. M. G. *et al.* Barrett's oesophagus is characterized by a predominantly humoral inflammatory response. *J Pathol* **207**, 269–276 (2005).

481. Rosenberg, C. E. *et al.* Esophageal IgG4 Levels Correlate with Histopathologic and Transcriptomic Features in Eosinophilic Esophagitis. *Allergy* **73**, 1892 (2018).

482. Barrett's Esophagus in Eosinophilic Esophagitis, A Coinciden... : Official journal of the American College of Gastroenterology | ACG. https://journals.lww.com/ajg/fulltext/2011/10002/barrett_s_esophagus_in_eosinophilic_esophagitis,_a.20.aspx.

483. Gilani, N., Gerkin, R. D., Ramirez, F. C., Hakim, S. & Randolph, A. C. Prevalence of Barrett's esophagus in patients with moderate to severe erosive esophagitis. *World Journal of Gastroenterology : WJG* **14**, 3518 (2008).

484. Barros, L. R. C. *et al.* High infiltration of B cells in tertiary lymphoid structures, TCR oligoclonality, and neoantigens are part of esophageal squamous cell carcinoma microenvironment. *J Leukoc Biol* **108**, 1307–1318 (2020).

485. Deguchi, S. *et al.* Clinical relevance of tertiary lymphoid structures in esophageal squamous cell carcinoma. *BMC Cancer* **22**, 1–12 (2022).

486. Li, R. *et al.* Tertiary lymphoid structures favor outcome in resected esophageal squamous cell carcinoma. *J Pathol Clin Res* **8**, 422–435 (2022).

487. Schlößer, H. A. *et al.* B cells in esophago-gastric adenocarcinoma are highly differentiated, organize in tertiary lymphoid structures and produce tumor-specific antibodies. *Oncoimmunology* **8**, (2019).

488. DeFalco, J. *et al.* Non-progressing cancer patients have persistent B cell responses expressing shared antibody paratopes that target public tumor antigens. *Clinical Immunology* **187**, 37–45 (2018).

489. Wieland, A. *et al.* Defining HPV-specific B cell responses in patients with head and neck cancer. *Nature 2020 597:7875* **597**, 274–278 (2020).

490. Kroeger, D. R., Milne, K. & Nelson, B. H. Tumor-infiltrating plasma cells are associated with tertiary lymphoid structures, cytolytic T-cell responses, and superior prognosis in ovarian cancer. *Clinical Cancer Research* **22**, 3005–3015 (2016).

491. Weiner, A. B. *et al.* Plasma cells are enriched in localized prostate cancer in Black men and are associated with improved outcomes. *Nature Communications 2021 12:1* **12**, 1–10 (2021).

492. Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P. & de Preter, K. Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nature Communications 2020 11:1* **11**, 1–14 (2020).

493. Chu, T., Wang, Z., Pe'er, D. & Danko, C. G. Cell type and gene expression deconvolution with BayesPrism enables Bayesian integrative analysis across bulk and single-cell RNA sequencing in oncology. *Nature Cancer 2022 3:4* **3**, 505–517 (2022).

494. Bashford-Rogers, R. J. M. *et al.* Analysis of the B cell receptor repertoire in six immune-mediated diseases. *Nature 2019 574:7776* **574**, 122–126 (2019).

495. Xu, J. L. & Davis, M. M. Diversity in the CDR3 Region of VH Is Sufficient for Most Antibody Specificities. *Immunity* **13**, 37–45 (2000).

496. Pickman, Y., Dunn-Walters, D. & Mehr, R. BCR CDR3 length distributions differ between blood and spleen and between old and young patients, and TCR distributions can be used to detect myelodysplastic syndrome. *Phys Biol* **10**, 056001 (2013).

497. Vandyk, L. & Meek, K. Assembly of IgH CDR3: Mechanism, Regulation, and Influence on Antibody Diversity. *http://dx.doi.org/10.3109/08830189209055568* **8**, 123–133 (2009).

498. Bose, B. & Sinha, S. Problems in using statistical analysis of replacement and silent mutations in antibody genes for determining antigen-driven affinity selection. *Immunology* **116**, 172–183 (2005).

499. Zhou, Z. H. *et al.* Polyreactive Antibodies Plus Complement Enhance the Phagocytosis of Cells Made Apoptotic by UV-Light or HIV. *Scientific Reports 2013 3:1* **3**, 1–11 (2013).

500. Tennant, I. *et al.* Innate recognition of apoptotic cells: novel apoptotic cell-associated molecular patterns revealed by crossreactivity of anti-LPS antibodies. *Cell Death & Differentiation 2013 20:5* **20**, 698–708 (2013).

501. Cocca, B. A. *et al.* Structural basis for autoantibody recognition of phosphatidylserine-2 glycoprotein I and apoptotic cells.

502. Shao, W. H. & Cohen, P. L. Disturbances of apoptotic cell clearance in systemic lupus erythematosus. *Arthritis Res Ther* **13**, 1–7 (2011).

503. Gordon, S. & Plüddemann, A. Macrophage clearance of apoptotic cells: A critical assessment. *Front Immunol* **9**, 127 (2018).

504. Rosen, L. A. C., Anhah, G. & Rosen, A. Autoantigens targeted in systemic lupus erythematosus are clustered in two populations of surface structures on apoptotic keratinocytes. *Journal of Experimental Medicine* **179**, 1317–1330 (1994).

505. Shero, J. H., Bordwell, B., Rothfield, N. F. & Earnshaw, W. C. High Titers of Autoantibodies to Topoisomerase I (Scl-70) in Sera from Scleroderma Patients. *Science (1979)* **231**, 737–740 (1986).

506. Maul, G. G., French, B. T., van Venrooij, W. J. & Jimenez, S. A. Topoisomerase I identified by scleroderma 70 antisera: Enrichment of topoisomerase I at the centromere in mouse mitotic cells before anaphase (kinetochore/mitosis/progressive systemic sclerosis). *Cell Biology* **83**, 5145–5149 (1986).

507. Monk, J. M. *et al.* Antigen receptor repertoire profiling from RNA-seq data. *Nature Biotechnology 2017 35:10* **35**, 908–911 (2017).

508. Hu, X. *et al.* Landscape of B cell immunity and related immune evasion in human cancers. *Nature Genetics 2019 51:3* **51**, 560–567 (2019).

509. Kitaura, K. *et al.* Different somatic hypermutation levels among antibody subclasses disclosed by a new next-generation sequencing-based antibody repertoire analysis. *Front Immunol* **8**, 389 (2017).

510. Yao, C. *et al.* Tumor-infiltrating plasma cells are the promising prognosis marker for esophageal squamous cell carcinoma. *Esophagus* **18**, 574–584 (2021).

511. Fristedt, R. *et al.* Prognostic impact of tumour-associated B cells and plasma cells in oesophageal and gastric adenocarcinoma. *J Gastrointest Oncol* **7**, 848–859 (2016).

512. Xia, J. *et al.* Single-cell landscape and clinical outcomes of infiltrating B cells in colorectal cancer. *Immunology* (2022) doi:10.1111/IMM.13568.

513. Ito, T. *et al.* Class distribution of immunoglobulin-containing plasma cells in the stroma of medullary carcinoma of breast. *Breast Cancer Research and Treatment 1986 7:2* **7**, 97–103 (1986).

514. Mattes, M. J., Michel, R. B., Goldenberg, D. M. & Sharkey, R. M. Induction of Apoptosis by Cross-Linking Antibodies Bound to Human B-Lymphoma Cells: Expression of Annexin V Binding Sites on the Antibody Cap. *Cancer Biother Radiopharm* **24**, 185 (2009).

515. Kourtzelis, I., Hajishengallis, G. & Chavakis, T. Phagocytosis of Apoptotic Cells in Resolution of Inflammation. *Front Immunol* **11**, 553 (2020).

516. Zhou, Y., Yao, Y., Deng, Y. & Shao, A. Regulation of efferocytosis as a novel cancer therapy. *Cell Communication and Signaling 2020 18:1* **18**, 1–11 (2020).

517. Erwig, L. P. & Henson, P. M. Clearance of apoptotic cells by phagocytes. *Cell Death & Differentiation 2008 15:2* **15**, 243–250 (2007).

518. Huynh, M.-L. N., Fadok, V. A. & Henson, P. M. Phosphatidylserine-dependent ingestion of apoptotic cells promotes TGF-β1 secretion and the resolution of inflammation. *J Clin Invest* **109**, 41–50 (2002).

519. Maderna, P. & Godson, C. Phagocytosis of apoptotic cells and the resolution of inflammation. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1639**, 141–151 (2003).

520. Duchemin, M., Tudor, D., Cottignies-Calamarte, A. & Bomsel, M. Antibody-Dependent Cellular Phagocytosis of HIV-1-Infected Cells Is Efficiently Triggered by IgA Targeting HIV-1 Envelope Subunit gp41. *Front Immunol* **11**, 1141 (2020).

521. Lara, S., Anania, J. C., Virtanen, A., Stenhammar, V. & Kleinau, S. Importance of antibody isotypes in antitumor immunity by monocytes and complement using human-immune tumor models. *Eur J Immunol* **51**, 1218–1233 (2021).

522. Bunn, C. C. & Black, C. M. Systemic sclerosis: an autoantibody mosaic. *Clin Exp Immunol* **117**, 207 (1999).

523. Newkirk, M. M., van Venrooij, W. J. & Marshall, G. S. Autoimmune response to U1 small nuclear ribonucleoprotein (U1 snRNP) associated with cytomegalovirus infection. *Arthritis Res* **3**, 253–258 (2001).

524. Ho, K. T. & Reveille, J. D. The clinical relevance of autoantibodies in scleroderma. *Arthritis Res Ther 2003 5:2* **5**, 1–14 (2003).

525. Maria, A. T. J. *et al.* Intriguing relationships between cancer and systemic sclerosis: Role of the immune system and other contributors. *Front Immunol* **10**, 3112 (2019).

526. Weeding, E., Casciola-Rosen, L. & Shah, A. A. Cancer and scleroderma. *Rheum Dis Clin North Am* **46**, 551 (2020).