UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Cancer Sciences

Investigation of cellular components of the anti-tumour adaptive immune response in humans

by

Oliver Wood

Thesis for the degree of Doctor of Philosophy (PhD)

December 2019

Abstract

The adaptive immune response to tumours plays a major role in determining outcome. Patients whose tumours contain high numbers of infiltrating lymphocytes (TIL), a marker of on-going immune attack, are likely to live longer. The improved understanding of immune cells in the tumour microenvironment has led to ground-breaking therapies (e.g. anti-CLTA4 and anti-PD1), in which targeting specific components of the anti-tumour immune response can lead to dramatically improved patient survival. The studies here focus on HNSCC (head and neck squamous cell carcinoma) and NSCLC (non-small cell lung cancer) to explore how differences in the adaptive immune response reflected in tumour infiltrating lymphocyte density, translates into differing efficacies and phenotypes of the different lymphocyte populations.

Monitoring of immunotherapy clinical trials and evaluation of changes relies on a stable signal over time. To address this, multiple tumour replicates were investigated between diagnostics biopsy and surgical resection in HNSCC. An immunologically focused transcriptomic and CD8 histology evaluation showed that tumours are stable, both over time and space. In HNSCC, HPV(+) and (-) disease display a significantly different patient survival, even among the TIL high group. Transcriptomic and histological assessment of these tumours at the bulk tissue level, revealed that this difference was driven by quantitative rather than qualitative differences in the T cells. However, a differential B cell signature emerged between HPV(+) and (-) tumours. Isolation of specific immune cell subsets (CD8+ T cells and B cells) from NSCLC and HNSCC enabled higher resolution analysis of immune cells in different disease settings (TIL density). Tissue resident memory T cells (T_{RM}) were enriched in tumours, the effector function of the T_{RM} cells were also found to be superior in TIL^{Hi} tumours. The T_{RM} cells were also prognostic, where higher numbers in the tumour conferred a survival benefit in NSCLC, and a trend towards improved survival in HNSCC. B cells from $CD8^{Hi}$ and $CD8^{Lo}$ NSCLC were assessed using single cell transcriptomics, and have been shown to exhibit different cellular characteristics depending on T cell density. B cells in immune-rich tumours ($CD8^{Hi}$) show a response to cytokine stimulus, the ability to present antigens and evidence of a germinal centre reaction. The B cells in $CD8^{Lo}$ tumours exhibit an alternative activation profile with increased markers of cellular stress and reduced interactions with co-stimulatory cells. The findings show that many of the cellular components of the immune infiltrate show significant qualitative differences which appear to associate with the density of CD8+ T cells. Future work needs to explore the control processes and cross-talk between the cells of the adaptive immune system that may be involved in determining these differences.

Table of contents

List of Tables

List of Figures

DECLARATION OF AUTHORSHIP

I, Oliver Wood declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Title: Investigation of cellular components of the anti-tumour adaptive immune response in humans

I confirm that:

1.This work was done wholly or mainly while in candidature for a research degree at this University;
2.Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3.Where I have consulted the published work of others, this is always clearly attributed;
4.Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5.I have acknowledged all main sources of help;
6.Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7.Parts of this work have been published as:

- Wood, O.*, et al.* Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget* **7**, 56781-56797 (2016

- Wood, O.*, et al.* Head and Neck Squamous Cell Carcinomas Are Characterized by a Stable Immune Signature Within the Primary Tumor Over Time and Space. *Clinical cancer research : an official journal of the American Association for Cancer Research* **23**, 7641-7649 (2017).

- Ganesan, A.P.*,* Clarke, J.C., Wood, O.*, et al.* Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer. *Nature immunology* **18**, 940-950 (2017).

Signed:

Date:.

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor Prof. Christian Ottensmeier for the continuous support of my study and related research, for his patience, motivation and expert knowledge. His guidance helped me during the time of researching and writing of this thesis. A special thanks also goes to Prof. Peter Friedmann for his continued wisdom, support and guidance. I would also like to thank the rest of my thesis co-supervisors:, Prof. Gareth Thomas, Dr Vijayanand Pandurangan and Dr Emma King, for their continued expert assistance, encouragement and guidance. My sincere thanks also go to the clinical support teams in Southampton and Poole hospital that make it possible for us to undertake this type of research. I would also like to thank the members of the ECMC and WISH lab for their laboratory support and training. Thanks also goes to Cancer Research UK and the associated donors for funding the majority of the work throughout this thesis. I thank my fellow laboratory colleagues and friends in Cancer Sciences at the University of Southampton. Last but not least, I would like to thank my family: firstly my wife and daughter, my parents and my brother for supporting me throughout this thesis and life in general.

Definitions and abbreviations

ACT - adoptive cell therapy
ADCC - antibody dependant cell cytotoxicity
APC - antigen presenting cells
BALT - bronchus associated lymphoid tissue
BCR - B cell receptor
CAFs - Cancer associated fibroblasts
CD – Cluster of differentiation
CGA - cancer genome data analysis
CL - cryopreserved non-involved lung
CSR - class switch recombination
CT - cryopreserved tumour
Ct - cycle time (qPCR)
DB - diagnostic biopsy
DEG - differentially expressed genes
DSL - dispersed lung
DST - dispersed tumour
ECM - extracellular matrix
FACS - fluorescence-activated cell sorting
FDR - false discovery rate
FFPE - formalin fixed, paraffin embedded
GO - gene ontology
H&E - haematoxylin and eosin
HIV-1 - Human immunodeficiency virus 1
HNSCC - head and neck squamous cell carcinoma
HPF - high-power field
HPV - Human papilloma virus
HSPs - heat shock proteins
IFC - integrated fluidics chip
IHC - Immunohistochemistry
ISGs - interferon stimulated genes
LN - lymph node
LPS - lipopolysaccharides
MAST - model based analysis of single cell transcriptomics
MHC - major histocompatibility complex
NLRs - nucleotide oligomerisation domain receptors
NSCLC - non-small cell lung cancer
PAMPs - pathogen-associated molecular patterns
PC - principal components
PCA - principal component analysis
PRRs - pattern-recognition receptors
RIN - RNA integrity number
RLRs - RIG-I like receptors
ROS - reactive oxygen species
RPKM – Reads per kilobase per million mapped reads
RT - Reverse transcription
SC - single cells
SCC - Squamous cell carcinomas
SD - standard deviation
SHM - somatic hypermutation
SR - surgical resection
t-SNE - T-distributed Stochastic Neighbour Embedding
TAA - Tumour associated antigens
TCR - T cell receptor
TIBs - tumour-infiltrating B lymphocytes
TIL - tumour-infiltrating lymphocytes
TLRs - Toll-like receptors
TLS - tertiary lymphoid structure
TMA - tissue microarray
TME - tumour microenvironment
TPM – transcripts per million mapped reads
UPR - unfolded protein response

Chapter 1: Introduction

The rapid expansion of research and development in the field of cancer immunology and immunotherapy has led to an improved understanding of both cancer biology and immunology. The key parts of the adaptive and innate immune system and how they interact with solid tumours are outlined, with a focus on the adaptive immune response in head and neck squamous cell carcinoma (HNSCC) and non-small cell lung cancer (NSCLC).

1.1 Head and Neck Squamous Cell Carcinoma (HNSCC)

1.1.1 HNSCC Epidemiology and Mortality

Head and neck squamous cell carcinoma (HNSCC) has an estimated incidence of 650,000 new cases and 350,000 deaths worldwide every year. This makes it the sixth most common type of cancer, accounting for approximately 6% of all cases. The 5-year survival data is about 60% for all stages combined on the basis of Surveillance Epidemiology and End Results (SEER) (Argiris et al., 2008, Parkin et al., 2005). Historically, the risk factors involved in HNSCC were predominantly smoking and alcohol, however changes in social behaviour have led to an increase in HPV (Human papilloma virus) associated HNSCC and an improved survival due to a shift in its aetiology (Chaturvedi et al., 2008). The published data indicates that between 50-75% of HNSCC are HPV-associated (Chaturvedi et al., 2011, Attner et al., 2010, Ward et al., 2014b) with HPV-positive [HPV(+)] patients displaying a more favourable prognosis compared to HPV-negative patients [HPV(-)]. In one study, the 3 year survival was reported as 84% for HPV(+) and 57% for HPV(-) (Ang et al., 2010), in another study, the 5 year survival was identified as 62% for HPV(+) and 26% for HPV(-) (Lassen et al., 2009).

1.1.2 HPV independent HNSCC

The main risk factors in HNSCC independent of HPV, are smoking and alcohol consumption. It would be highly attractive to be able to stratify patients based on information other than stage and basic histology. A comprehensive analysis of HNSCC (n= 172) at the genomic and transcriptomic level reveals that smoking-related HNSCC exhibits an almost universal loss of TP53 and CDKN2A, both key tumour suppressor genes (Cancer Genome Atlas, 2015). These large-scale datasets allow an unprecedented insight into the driving mechanism and tumour pathogenesis with the aim of identifying therapeutic targets. Analysis of HPV(-) HNSCC at a single cell level has taken this further with the identification of different subtypes based on the transcriptional profile

of the malignant cells and the interaction with epithelial cells, partial epithelial mesenchymal transition represented a key part of tumour metastasis, tumour grade and poor prognosis (Puram et al., 2017).

1.1.3 HPV driven HNSCC

1.1.3.1 HPV life cycle and events leading to cancer:

The HPV16 genome (Figure 1A) consists of a viral DNA organised into stages of expression defined as Early and Late: (E1-7, L1 and L2). The replication enzyme E1, and accessory protein E2, are involved in regulation of DNA replication, the E4 protein enables viral replication to take place and facilitates virion release, E5 protein is involved in maintaining the host cell in a proliferative state. The capsid proteins L1 and L2 are structural proteins that encapsulate the viral genomes and virions ready for release from the upper layers of the epithelium (Stanley, 2012, Moody and Laimins, 2010).

During an active HPV infection, the E6 and E7 oncoproteins play a major role in the transformation of the infected basal mucosal squamous cells into cancer cells, E6 and E7 proteins target critical cell cycle regulators leading to aberrant proliferation and cellular transformation. The E7 protein targets the retinoblastoma tumour suppressor (pRB) for degradation, this causes the cell to transition from G1 phase to S-phase of the cell cycle, leading to irregular proliferation (Dyson, 1998, Korzeniewski et al., 2011). Disruption of the pRB signalling pathways by E7 enables viral genome replication to occur in non-proliferating cells. Under normal cellular conditions, the unscheduled progression from G1 to S-phase would initiate pro-apoptotic cellular responses in the host cell. However, the E6 oncoprotein inhibits this process by targeting p53 for degradation via Ubiquitin-protein ligase E3A (UBE3A /E6AP), disrupting the apoptotic signalling cascade (Patel et al., 1999). Cellular immortalisation is another key step in the transition to cancer and occurs as a result of E6- dependent up-regulation of human telomerase reverse transcriptase (hTERT); this disrupts subsequent erosion of telomeres which would occur as a result of repeated cell division (Klingelhutz et al., 1996). HNSCC tumours display viral genome integration in 39% of cases with 61% displaying episomal viral sequences (Olthof et al., 2014). This study highlighted that it was constitutive expression of the high-risk proteins E6 and E7 that lead to cancer progression caused by the genomic instability as well as the integration on the viral genome.

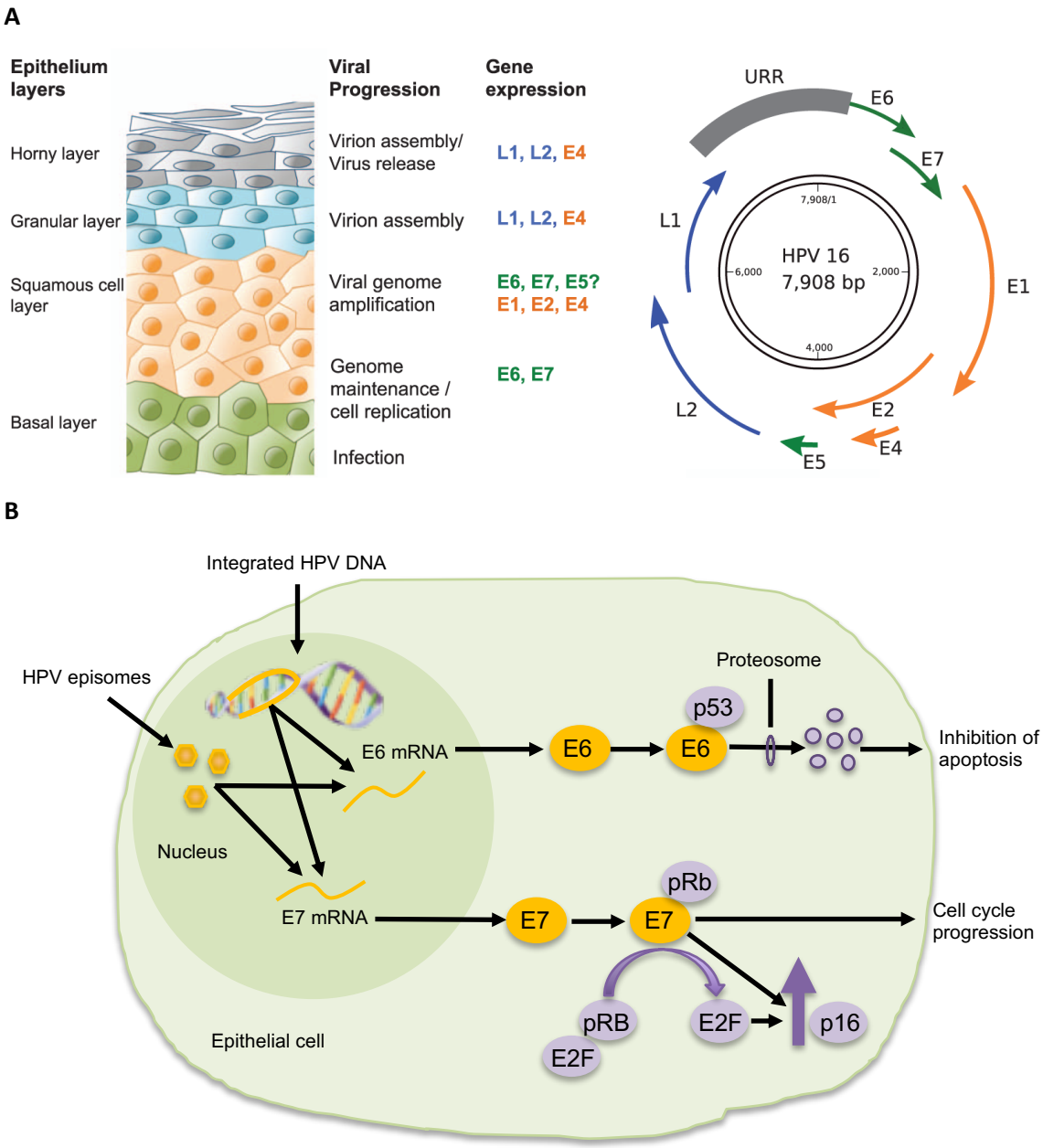


Figure 1. HPV life cycle, genomic organisation and cellular transformation.

(A) The key events that occur following infection and the stages of viral replication and the genes involved at each stage are depicted. Initially starting with the E6 and E7 oncoproteins in the basal epithelium at the site of infection. The organisation of the HPV16 genome is also shown highlighting the order in which the HPV-16 genes (E6, E7, E1, E2, E4, E5, L2 and L1) are expressed. (B). The cellular transformation of the infected epithelial cells takes place as a result of the E6 and E7 oncoproteins. E6 functions to target p53 for degradation, thus inhibiting apoptosis. The E7 protein interacts with the retinoblastoma (pRb) protein leading to its degradation and uncontrolled cellular proliferation. As a result of the lacking pRb the transcription factor E2F remains unbound and causes the increase of p16 (CDKN2A) a key histological marker in HPV Adapted from (Andersen et al., 2014).

1.2 Non-small cell lung cancer (NSCLC)

1.2.1 NSCLC epidemiology and mortality

Lung cancer is one of the leading causes of cancer worldwide with a 5-year survival of 20% (Siegel et al., 2018). NSCLC accounts for 90% of all lung cancer cases and is divided into the two subtypes Adenocarcinoma (50% of total) and Squamous cell carcinoma (40% of total)(Chen et al., 2014). The classification and treatment are primarily based on the TNM staging system (tumour size = T; nodal involvement = N and distal metastasis = M) (Mirsadraee et al., 2012), within this system tumours are classified into stages I-IV. The clinical outcomes can vary dramatically between patients that have the same stage tumours. A critical factor in the variability of patient outcome are the extent and type of immune cells that infiltrate the tumours (Domagala-Kulawik, 2015, Horne et al., 2011, Kilic et al., 2011).

1.2.2 Subtypes of NSCLC

Adenocarcinoma (AD) arises predominantly in the distal airways, they display features of glandular histology and express biomarkers that are specific to the distal region of the airways (Keratin 7- KRT7 and thyroid transcription factor 1 -TTF1). The alveolar epithelial type 2 cells (club cells) are proposed to be the cell of origin for adenocarcinoma (Thunnissen et al., 2012). Squamous cell carcinomas (SCC) arise in the more proximal regions of the airways and are associated with the chronic inflammation usually caused by smoking. They display squamous differentiation that is similar to the epithelium of the upper airways, the tracheal basal cell is proposed to be the cell of origin for SCC (Davidson et al., 2013).

1.3 The role of the immune system in cancer

The importance of the immune system and its potential role in cancer was first put forward in 1970, where the concept of immune surveillance of cancer was raised (Burnet, 1970). It was hypothesised that the potential frequency of mutations leading to malignancy would need mechanisms for elimination of these mutant cells; one was proposed to be immunological (Burnet, 1970). The evidence for the impact of the immune system on cancer can be seen in situations where it is compromised, such as inherited and drug-induced immunodeficiencies. Primary immunodeficiencies are genetic disorders that lead to severe impairment of the immune system leading to autoimmunity and increased risk of infections, as well an increased risk of cancer. The overall risk of children with primary immunodeficiency developing cancer is estimated to be 4-25% depending on the type of immunodeficiency, showing a link between the immune

system and the control of cancer cells (Salavoura et al., 2008). Another example is increased risk of cancer linked to virally induced immunodeficiency as is the case with HIV-1 (Human immunodeficiency virus type 1)/AIDS. These patients have reduced numbers of CD4 T cells and have been shown to have an elevated risk of cancer. This has been attributed to oncogenic viruses and also environmental factors like smoking; HIV has also been shown to be an independent risk factor in developing lung cancer (Sigel et al., 2012),(Clifford et al., 2005). Immunosuppressive drugs that are administered to organ transplant recipients to reduce the risk of transplant rejection also lead to the increase risk of cancer, with a 2- to 5-fold increase in risk for cancers of the colon, larynx, lung, bladder, prostate, and testis, an increased risk of 10- to 30-fold was observed in cancers of the lip, skin, kidney and endocrine glands (Birkeland et al., 1995). The role of the immune system in control of cancer is clearly evident from the increased risk of malignancy that arise as result of perturbations in the immune system.

The importance of the immune system can also be seen by the impact that certain cell types have on patient survival. The immune density of CD4 and CD8 cells within tumours predicts for survival in many cancers, where a denser and more functional infiltrate leads to a reduction in tumour burden and improved survival, these cell types and the core elements of the immune system are discussed with a focus on cancer biology (Fridman et al., 2012, Hiraoka et al., 2006, Mahmoud et al., 2011).

1.3.1 Immune response to tumours

The general structure of tumours includes a tumour core, an invasive margin and a surrounding stromal compartment; heterogeneous immune infiltrates can be observed dispersed or aggregated within the tumour microenvironment. Tumour cells develop from mutations and alterations in key genes, the main hallmarks of cancer are evading growth suppression, resistance to apoptosis, metastasis, angiogenesis, sustained proliferative capacity, modulation of cellular metabolism and the ability to avoid immune destruction (Hanahan and Weinberg, 2011, Hanahan and Weinberg, 2000). Most tumours will contain a heterogeneous immune infiltrate comprising dendritic cells, Natural killer cells, macrophages, mast cells and then the B cells and T cells. It is important to know how stable immune infiltrates are across a tumour sample; this becomes an important consideration when sampling small areas of the whole tumour.

The clinical benefit for patients with highly infiltrated tumours is well documented with one of the key pieces of research carried out in colorectal cancer, where it was shown that a high number of CD8+ lymphocytes both intratumoural and at the invasive margin, conferred a vast improvement in survival (Galon et al., 2006, Ward et al., 2014b). The immune landscape of tumours was characterised in detail using a network of gene expression and systems biology to show that

cellular infiltrates consisted of a functional immune response that changes over time with a reduction in T cells and an increase in B cells and T follicular helper (Tfh) cells as tumours progress (Bindea et al., 2013).

1.3.2 Immunoediting

Immunoediting refers to the interaction of the immune system and the tumour over time. Originally identified in mouse experiments using a combination of immunocompetent/ immunodeficient mice. Tumours forming in mice lacking an immune system were more immunogenic ("unedited") than those from immunocompetent mice ("edited")(Shankaran et al., 2001, Dunn et al., 2002). The immune system protects against tumour growth but also shapes tumour immunogenicity. This has been put into three distinct phases of elimination, equilibrium and escape, the process of elimination involves innate and adaptive immune cells that either eradicate tumour cells completely, or due to mutational variations, some tumour clones are not killed (Vesely et al., 2011). In equilibrium the mutational events (germline or epigenetic) lead to tumour cell variants escaping immune attack, this phase is considered the longest and where immunoediting occurs. It is also referred to as a type of dormancy where disease progression is limited due to immunological control (Aguirre-Ghiso, 2007). The escape phase refers to tumour cells that no longer elicit an immunoprotective response, either from (i) loss of antigenicity, (ii) insensitivity to effector mechanisms or (iii) an immunosuppressive state. Once these tumours enter this phase, which is essentially selection of poorly immunogenic tumour variants that are no longer visible to the immune system, tumour outgrowth occurs rapidly (Vesely et al., 2011).

1.3.3 Stromal features of tumours and immune evasion

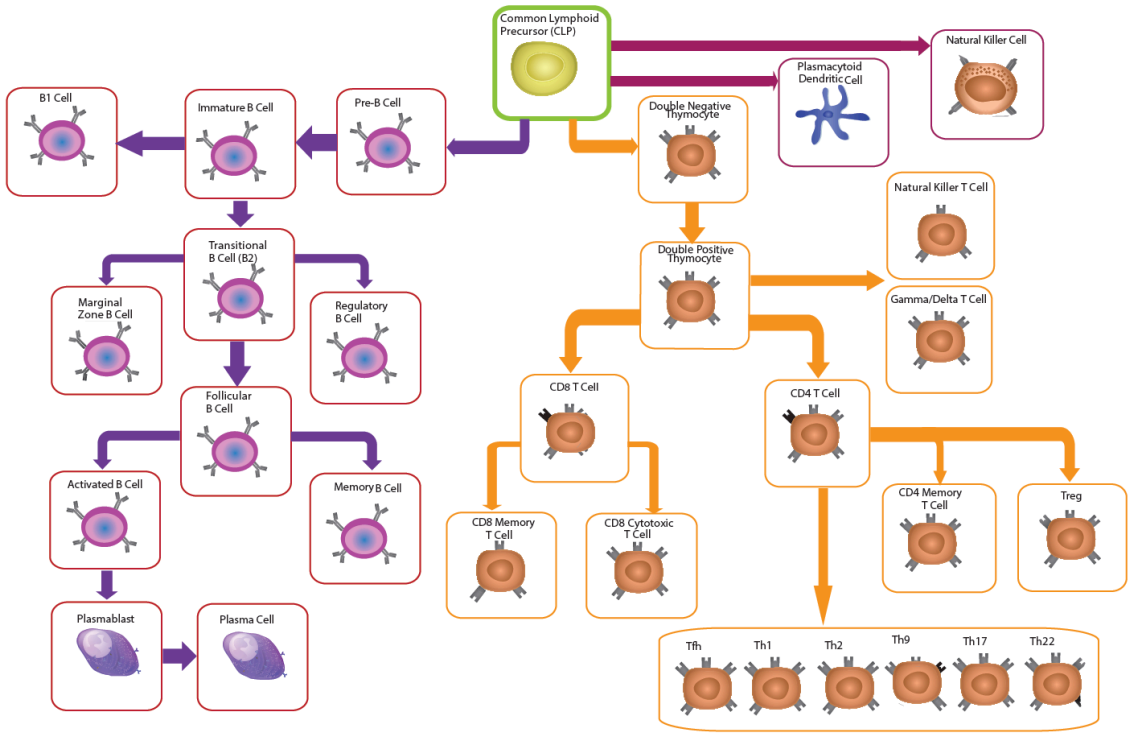
An important feature of all tumours and immune infiltrates is the tissue context or stromal composition that forms the tumour microenvironment (TME). Cancer associated fibroblasts (CAFs) are a key part of the TME, playing a role in tumour progression and also suppression of antitumor immunity (Shalapour et al., 2017). The activation of these via TGFβ leads to αSMA(+) myofibroblasts that contribute to cancer progression and a poor prognosis via extracellular matrix (ECM) re-modelling. The conversion of fibroblasts to myofibroblasts relies on the ROS (reactive oxygen species) generated by NOX4; inhibition of this pathway leads to a reduction in tumour growth and fewer αSMA(+) cells (Hanley et al., 2018). The impact on the immune cell component is one of impeded migration through the ECM, as shown in migration assays of CD8 cells in collagen rich tumours (Bougherara et al., 2015).

1.4 Adaptive Immune system

The fundamental role of the immune system is to distinguish self from non-self; at the core of this co-ordinated response are the innate and adaptive immune systems. The lymphocytes of the adaptive immune system (T cells and B cells) generate a response to unique antigens that are recognised as non-self with the aim of clearing the pathogen/ cells from the host. The key cell types of the adaptive immune system are shown in Figure 2. T cells mature in the thymus (primary lymphoid tissue) and make up cellular immunity, with the humoral immune response being fulfilled by B cells that develop in the bone marrow (primary lymphoid tissue); both groups of cells then migrate to the secondary lymphoid organs (lymph nodes and spleen) (Bendall et al., 2011).

T cells are activated by antigen presenting cells (e.g. dendritic cells) that process and present protein peptides to CD4+ and CD8+ cells. Activation of naïve helper T cells requires 3 signals in the form of TCR (T cell receptor) stimulus, co-stimulation from antigen presenting cells and environmental stimulus in the form of cytokines. Activation of naïve CD4 Th0 helper cells leads to a diverse range of cells that provide “help” to CD8+ cells and B cells as well as other cells of the immune system (e.g macrophages) via co-stimulation and cytokine release. CD8+ cytotoxic cells are activated by antigen presenting cells and function to kill infected/ damaged cells; T-helper cells also play a role in their activation. B cell activation occurs when antigen is bound by the unique BCR, co-stimulation is required from T helper cells as is the action of cytokines. This leads to the release of antibodies that bind antigen and can activate innate immune cells (e.g. natural killer and macrophages) that enable pathogen/ damaged cell clearance. The activated effectors cells (T cells and B cells) differentiate and expand, carry out their effector functions, followed by a contraction phase and the formation of long-lasting memory cells that respond rapidly to future antigenic challenge. CD4+ and CD8+ memory cells are formed as well as memory B cells and plasma cells. More detail on the different aspects of the adaptive immune response and its role in cancer biology are described below.

A



B

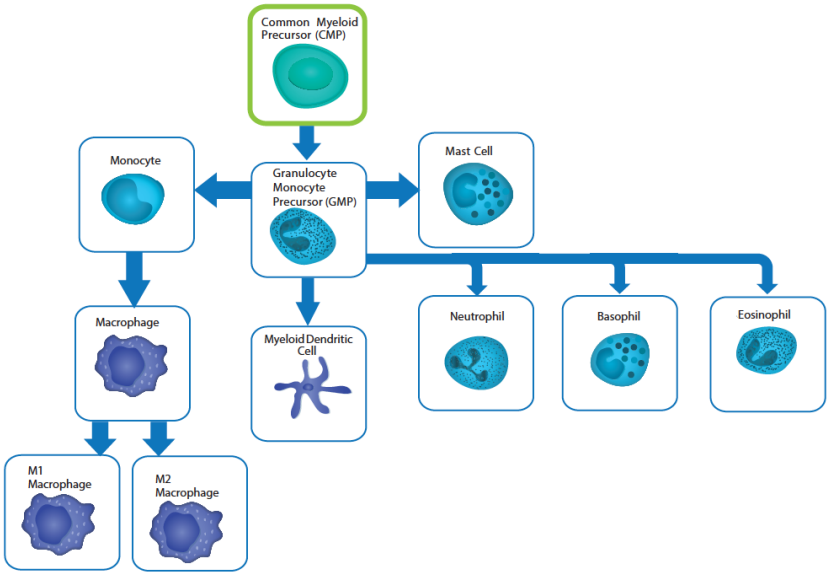


Figure 2. Cells of the lymphoid and myeloid cell lineages.

(A) Cells of the lymphoid lineage originate from a common lymphoid precursor and give rise to innate Natural killer cells, plasmacytoid dendritic cells and the T cells and B cells that then differentiate into the different functional cell populations. (B) The innate cells of the myeloid lineage originate from the common myeloid precursors and give rise to the granulocyte/ monocyte precursor and subsequently the macrophages, neutrophils, basophils, eosinophils and myeloid dendritic cells. Adapted from Human immune cell lineage and antigen expression database, Bio-Rad-antibodies.com, 2018.

1.4.1 Cellular immunity and T cells

1.4.1.1 T-cell development and receptor rearrangement

T-cells develop from common lymphoid progenitor cells originating in the bone marrow (Figure 2 shows an overview of the T cell linage). They then migrate to the thymus as CD8 and CD4 double negative cells. Here they develop further via rearrangement of the T-cell receptor (TCR) genes (Hedrick, 2008, Nimmo et al., 2015, Hedrick et al., 1984). The TCR is made up from the genes encoding the alpha, beta, gamma, and delta chains. The unique features of TCRs are encompassed by the complementarity determining regions CDR1, CDR2 and the hypervariable CDR3; these regions ultimately form the antigen binding site of the cell (Shetty and Schatz, 2015). The functional formed TCR is unique and has the ability to recognise a vast number of different targets including self-antigens, at this stage the thymocytes are CD4+ and CD8+ with the TCR complexed to CD3.

A process of positive and negative selection then occurs. Positive selection occurs when thymocytes interact through their TCR with peptide MHC-I and -II complexes on thymic epithelial cells in the medulla; this leads to selection of CD8+ cells that complex with pMHC-I (peptide-MHC) and CD4+ that complex with pMHC-II. Cells with very high avidity for self pMHC complexes are negatively selected against, cells receiving no signal are removed via apoptosis and a lack of survival signals (Capone et al., 2001, Van Laethem et al., 2013). The next stage of selection focuses on deletion of cells that interact with self-peptides that are normally expressed in distant organs, this is achieved through the gene AIRE (autoimmune regulator), which functions to stimulate gene expression of tissue restricted antigens (TRAs) in medullary thymic tissue (Mathis and Benoist, 2009, Danso-Abeam et al., 2011). Cells that exhibit high levels of self-reactivity are deleted as a mechanism of self-tolerance. The self-restricted (MHC-I/ II) non-autoreactive, negatively selected T cells (CD4/ CD8) downregulate CD69, and migrate to exit the medulla into the peripheral circulation via the expression of sphingosine-1-phospahte (S1P1) as mature CD4+ or CD8+ (CD62L+CD69-) cells (Drennan et al., 2009, Bunting et al., 2011).

1.4.1.2 Activation of T cells (signal 1)

Naïve T (Th0) cells reside in the paracortical areas of lymph nodes, awaiting the wake-up call that can only be given by Dendritic (professional) antigen presenting cells (APC). These APC sample and collect material from their local environment and potential antigenic peptides are processed and brought into physical association with MHC molecules which are then mounted on the APC surface where they can be presented to naïve T cells.

MHC class II (HLA-DR, HLA-DQ and HLA-DP) present peptides (11-33 amino acids) that originate from phagocytosis and endocytosis of exogenous proteins and interact with TCRs expressed by CD4+ Helper T cells (Neefjes et al., 2011, Matsui et al., 1994, Matsui et al., 1991). The MHC class I (HLA-A, HLA-B and HLA-C) present intracellular antigenic peptides (9-11 amino acids) in a complex with β2microglobulin (B2M) that are bound by TCRs of CD8+ cells. A single pMHC (peptide:MHC complex) interaction has been shown to be sufficient to trigger the activation of T cells via the TCR (Huang et al., 2013). The TCR signalling cascade consists of a complex set of interactions that are initiated upon a stable CD4/ CD8 TCR bound to pMHC-I/ II (signal 1), the CD3-TCR complex and co-receptor (CD4/ CD8) leads to activation of signal amplification transcriptional networks and cellular diversification modules (Malissen et al., 2014, Salek et al., 2013).

As stated previously the main site of T-cell priming takes place in the lymph nodes, it has now also been shown to take place in the tumour microenvironment. Antigen processing cells in the form of conventional dendritic cells (cDCs) that reside in the tumour microenvironment can acquire, process and present antigens to T-cells in-situ. This T-cell priming can initiate CD8 and CD4 responses outside of the lymph node (Sanchez-Paulete et al., 2017). One such study identified conventional DCs that reside in the tumour as a key cell type in the cross presentation of antigen to CD8+ T-cells in-situ, overall the CD8+ T-cell immune infiltrates were higher when more DCs were present (Broz et al., 2014).

1.4.1.3 T-cell co-stimulation (signals 2 and 3)

T cells activated via TCR engagement require additional signals via co-stimulation from antigen presenting cells (APC). The classical two signal hypothesis arose from studies of the co-stimulatory receptor CD28 on T-cells and its ligands CD80/ CD86 (B7.1 and B7.2) on APCs that act through MAPK, NFkB and PI3K pathways to promote cell growth, effector function and survival (Linsley et al., 1990). As part of the signalling programme, upregulation of CTLA4 occurs that downregulates CD28, it also competes with CD28 and the co-stimulatory receptor on the APC (Rudd et al., 2009). T-cells exhibit a wide range of co-signalling molecules that act as stimulatory and inhibitory checkpoints in T cell activation – so-called “Signal 2” summarised in Table 1. The 3rd signal that activated T cells require is delivered via APCs and the surrounding microenvironment. The innate sensing mechanism of APCs of PAMPs, DAMPs and PRRs (see section 1.3.3 Innate immunity) induce cytokine release that leads to polarisation of the T-cells. For example, release of IL-12 or IL-4 by APCs activates specific lineage transcriptional gene networks leading to differentiation of the cells into different effector cell types (e.g. CD4+ Th1 and Th2) (van Panhuys et al., 2014, O'Garra and Arai, 2000).

	Receptor	CD28	ICOS	CD27	4-1BB	0X40	GITR	PD1	CTLA4	TIM3	LAG3	TIGIT	CD80	CD86	PD-L1	PD-L2
$CD8^+$	Priming	+		+												
	Cell growth	+		+	+	+	+	-	-	-	-		-	-	-	-
	T_H cell differentiation															
	Effector function				+	+	+	-	-	-	-		-	-	-	-
	Survival	+	+	+	+	+		-	-							
	Memory	+	+	+	+	+		-	-	-						
	Receptor	CD28	ICOS	CD27	4-1BB	0X40	GITR	PD1	CTLA4	TIM3	LAG3	TIGIT	CD80	CD86	PD-L1	PD-L2
$CD4^+$	Priming	+		+												
	Cell growth	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-
	T_H cell differentiation		+	+		+				-						
	Effector function		+	+	+	+		-	-	-	-	-	-	-	-	-
	Survival	+	+	+	+	+		-	-							
	Memory	+	+	+	+	+			-	-						

Positive function Negative function

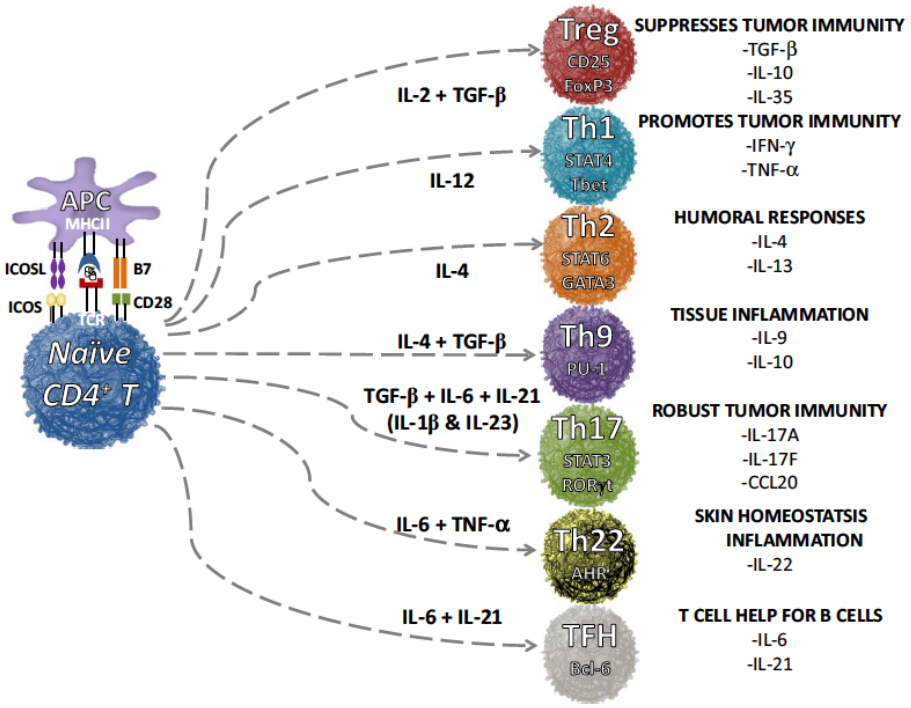
Table 1. T-cell co-stimulatory and inhibitory molecules.
An overview of the key co-stimulatory (+) and inhibitory molecules (-) for both CD4+ and CD8+ cells are shown with their functional effect (Cell priming, cell growth, Th cell differentiation, effector function, survival and memory) on that cell type. Adapted from (Chen and Flies, 2013, Schildberg et al., 2016).

1.4.1.4 CD4+ and CD8+ T cell phenotype and function.

1.4.1.4.1 CD4+ T helper cells

Naïve CD4 T helper cells (Th0), once activated, show extensive plasticity depending on the external stimuli that they receive. This leads to a diverse range of cell subsets with Th1, Th2, Th17, Th9, Th22, Tfh and Treg cells all developing from the same precursor cells; each cell type has specific transcriptional profiles and functions briefly outlined in Figure 3 (Zhou et al., 2009, Tripathi and Lahesmaa, 2014). Their primary function is to provide helper functions to CD8 cells and B cells via co-stimulation and the range of cytokines that each subset produces, in addition to this they function to inhibit excessive immune response via the immunosuppressive Treg populations. CD4+ Th cells also shift phenotype depending on the environmental cues, CD4+ Th17 inflammatory cells when exposed to the external stimulus of TGFβ are able to transdifferentiate into an anti-inflammatory Treg phenotype with the expression of FOXP3 (Treg) transcription factor (Gagliani et al., 2015, DuPage and Bluestone, 2016).

A



B

The subset	Factors inducing lineage	STAT activated	Lineage-specifying transcription factor	Effector cytokines produced	Functions
Th1	IL-12 IL-27	STAT4, STAT1	T-bet	IFNγ, lymphotoxin, TNFα	Cell-mediated immunity, delayed-type hypersensitivity responses, clearance of intracellular pathogens and tumor cells, opsonizing Ab production by B cell class-switching to IgG2a
Th2	IL-4 IDO	STAT6	Gata-3 c-MAF	IL-4, IL-5, IL-13, IL-10	Humoral immunity, clearance of extracellular bacteria and worms, B cell class-switching to IgE, allergic responses
Th9	IL-4 TGFβ	STAT6	BATF	IL-9, IL-10	Protection against parasitic worms/helminth infections
Th17	IL-6 MyD88 Low TGFβ IL-23	STAT3	RORγT, RORα	IL-17, IL-17F, IL-6, IL-22, TNFα, IL-10	Protection of mucosal surfaces, recruitment of neutrophils, clearance of Mycobacterium and tuberculosis
Th22	IL-10Rβ IL-6 TNFα	STAT3	Aryl hydrocarbon receptor	IL-22, IL-13, FGF, CCL15, CCL17, TNFα	Mucosal immunity, prevention of microbial translocation across epithelial surfaces, promotes wound repair.
Th25	IL-4, IL-25	Unknown	Act1	IL-25, IL-4, IL-5, IL-13	Mucosal immunity, stimulates nonlymphoid cells to produce IL-4, limits Th1 and Th17 induced inflammation, CD4 T cell memory (mouse)
TFH	Strong TCR signal, IL-12, CXCR5, IL-21, IL-4	STAT3	MAF (IL-21 transactivator)	IL-21, OX40, ICOS	Helps B cells produce high affinity, class-switched antibodies, guides migration into germinal centers
Treg	High TGFβ mTOR	STAT5	Foxp3	IL-10, TGFβ	Suppression of existing immune responses, maintains tolerance/protection against autoimmunity

Figure 3. CD4+ T- helper cell lineage commitment and cytokine functions.

(A) Diagrammatic representation of CD4+ T helper cell differentiation. Adapted from (Tripathi and Lahesmaa, 2014, Caza and Landas, 2015, Majchrzak et al., 2016). (B) The Functional and Phenotypic Plasticity of CD4+ T Cell Subsets is outline in brief with the key lineage determining cytokines and transcription factors associated with that lineage. The effector cytokines and overall function are also outlined. Adapted from (Tripathi and Lahesmaa, 2014, Caza and Landas, 2015, Majchrzak et al., 2016).

1.4.1.4.2 Cytotoxic CD8+ T cell (CTL)

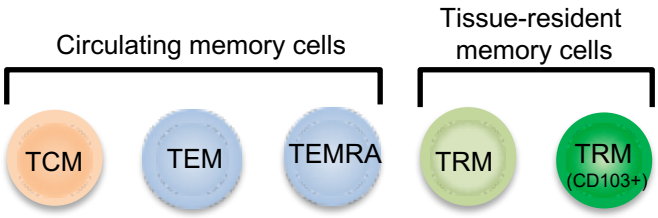
CD8+ T cells are the key defence against intracellular antigen(s) and function to kill infected cells (viruses/ bacteria) and malignant cells. This process is mediated via the initial activation of the cells via the TCR:pMHC-I interaction and the co-receptor CD8. Activation occurs when the CD8+ T cells are primed by APCs (e.g. Dendritic cells). In this process, the APC scans T cells (500+ T cells/ hr), once antigen engagement occurs its duration is influenced by the avidity of the interaction and can last for hours (Bousso and Robey, 2003). Following TCR activation (including co-stimulation as described previously), extensive proliferation of the CTL and short-lived effectors occurs that then reduce in number once the antigenic stimulus is cleared, leaving a memory population (Gerlach et al., 2010).

Maximal expansion of CTLs requires multiple signals in the form of co-stimulation, CD27, OX-40, 4-1BB and the cytokines IL-2, IL-12 and IFNγ all act to enhance CD8 expansion and function. The pro-inflammatory cytokine IL-12 acts via T-bet to induce CD8+ T cell differentiation via mTOR pathways (Rao et al. (2010)). IL-2 acts via CD25 (IL2R) and acts as a growth factor as well as enhancing the effector cell differentiation and functions. It functions via the STAT5, PI3K and MAPK pathways, inducing the transcriptional repressor BLIMP1 in a negative feedback loop (Pipkin et al., 2010, Malek and Castro, 2010). The effector function of the CTL is linked to an IL-2 dependent induction of the cytolytic components IFNγ, TNFα and the pore forming perforin and serine esterase granzymes (GZMA/ GZMB) that are essential for CTL activity (Pipkin et al., 2010, Harty et al., 2000, Zhang and Bevan, 2011).

1.4.1.5 Central memory and effector memory CD4+ and CD8+ T cells

Upon antigen exposure and co-stimulation, antigen-specific naïve CD4+ and CD8+ cells undergo cellular expansion and differentiation into the effector T cells as described above. During this process, precursors for the long-lived antigen-specific memory cells are formed, that upon subsequent antigen exposure in peripheral tissues, rapidly expand into effectors cells. After the initial expansion and contraction phases the different subsets of CD4 and CD8 cells are **effector memory (TEM), central memory (TCM)**, effector memory re-expressing CD45RA (**TEMRA**) and **T resident memory (T_{RM})**. This is based broadly on the expression of CD45RA, CCR7 and CD62L (see Figure 4) (Sallusto et al., 2004). The surface markers CD62L and CCR7 function to retain central memory cells in lymphoid tissues, where they differentiate into effector cells upon antigen re-encounter, conversely effector memory cells migrate to the peripheral tissues and respond to antigen with immediate effector functions. This then gives rise to additional subsets defined as resident memory cells (T_{RM}), often characterised by CD103 and CD69 expression, these cells are

retained in the tissue where they exert their effector functions (Farber et al., 2014, Sathaliyawala et al., 2013).



	TCM	TEM	TemRA	TRM	CD103 TRM
CD45RO	(+)	(+)	(-)	(+)	(+)
CD45RA	(-)	(-)	(+)	(-)	(-)
CCR7	(+)	(-)	(-)	(-)	(-)
CD62L	(+)	(-/+)	(-/+)	(-/+)	(-/+)
CD69	(-)	(-)	(+)	(+)	(+)
CD103	(-)	(-)	(-)	(-)	(+)
IL2	(++)	(+)	(+)	(-/+)	(-/+)
IFNγ	(+)	(++)	(++)	(++)	(++)
TNF	(+)	(++)	(++)	(++)	(++)

Figure 4. Central memory and effector memory T cells (CD4+ and CD8+).
Following T cell activation, expansion and contraction a memory population is generated that responds to future antigen stimulus. The phenotypic and functional characteristics of the cells are shown. T central memory (TCM) cells home to lymph nodes and are also found in circulation; conversely T effector (TEM) cells circulate and home to the tissues, the TEM expressing CD45RA (TEMRA) are terminally differentiated but express the CD45RA normally present on naïve cells. T-resident memory cells (T_{RM} and T_{RM} CD103+) are retained in the tissues (skin, lung, gastrointestinal tract and also in tumours).Adapted from (Farber et al., 2014 and Sallusto et al., 2004).

1.4.1.6 Gamma delta T cells (γδ T cells)

These are a diverse group cells that originate from the lymphoid thymocytes and recognise antigens in an MHC-independent manner and have the capacity to secrete cytokines. They are classified by their TCR composition and function. The TCR consists of a γ and δ chain and can be arranged in several combinations, four Vδ chains give rise to four main γδ T cells types with various Vγ chains. The V9γVδ2 T cells are the most abundant cells in the peripheral blood and recognise phospho-antigens (Dimova et al., 2015). Upon activation the cytokines IFNγ, TNFα and IL-17 can be released as well as the cytolytics perforin and granzyme. γδ T-cells can also serve as APCs, presenting antigen on MHC-I to CD8+ T cells (Brandes et al., 2009, Moens et al., 2011).

1.4.1.7 T cell inhibition

The highly potent immune response that ensues upon initial activation of the T cell system requires a way of self-limitation, this avoids excessive immune activation and the potential for harmful effects on the surrounding tissues. There are many co-inhibitory molecules expressed by T cells that dampen the immune response, some of these are summarised in Table 2. The interplay between some of the key molecules (PD1, PDL1, PDL2, LAG3, CTLA4 and TIM3) are also of importance as therapeutic targets for immunotherapy (see section 1.5.7 Immunotherapy). In the classical example of CD28 co-stimulation and T cell activation, the inhibitory molecule CTLA4 is induced on the T cell leading to interactions with the B7 receptors and competition for CD28 (CTLA4 has a greater affinity for B7 proteins than CD28). CTLA4 binding to B7 leads to activation of the phosphatase (PP2A), which in turn acts to inhibit the CD28/TCR activation of Akt, ultimately leading to a reduction in the immune response (Rudd et al., 2009).
The PD1 and PDL1/ 2 interaction is also a well characterised immune checkpoint. Expression of PD1 is primarily on mature T cells and limits effector function by binding to its main ligands PDL1 and PDL2. This leads to inactivation of ZAP70 (dephosphorylation), recruitment of SHP2 and inactivation by dephosphorylation of PI3K and subsequent inhibition of Akt, a reduction in cytokine release and cell survival also occurs (Keir et al., 2008). The use of these co-stimulatory and co-inhibitory molecules is of great interest in the field of immunotherapy where immunomodulatory agents can be used to overcome the negative regulators of the immune response.

1.4.1.8 T cell tolerance, anergy, exhaustion and senescence

Antigen-stimulated T cells normally differentiate into effector cells followed by a contraction phase yielding memory cells. Under certain circumstances, T cell dysfunction may occur leading to hyporesponsive/ unresponsive T cells. Peripheral tolerance refers to the ability to manage self-reactive T cells that escape central tolerance (central tolerance occurs in the thymus via deletion of autoreactive cells). Peripheral tolerance occurs by deletion of the autoreactive cells and suppression by regulatory cells (Tregs) and ultimately a terminal state of unresponsiveness (Redmond et al., 2005, Wing and Sakaguchi, 2010). **Anergic cells** represent cells that no longer respond appropriately to antigenic stimulus, originally identified in skin hypersensitivity reactions to recall antigens (Schier, 1954). This state represents cells stimulated through the TCR in the absence of co-stimulation, where an anergic transcriptional profile is induced, a key transcription factor in that program is EGR2 that regulates Ras/MAPK signalling and IL2 production (Zheng et al., 2012).

Immune Checkpoint Protein	Gene	Cellular Expression	Ligand	Mechanism of Action
CTLA-4 (suppressive)	CTLA4	Mostly naïve CD4 and CD8 T cells, Tregs	CD80 (B7-1) CD86 (B7-2)	❖ Negatively regulates T cell activation by neutralizing CD28 (outcompetes for to bind to CD80 and CD86) ❖ CD80 and CD86 ligands removed from APC surface via transendocytosis ❖ Reduces TCR signalling via impaired Akt activation
PD-1 (suppressive)	PDCD1	Activated T cells in peripheral tissue, B cells, professional APCs, NK cells	PD-L1 PD-L2	❖ Reduces T cell proliferation by Inhibiting T cell kinase ZAP70, recruitment of SHP2, dephosphorylation of PI3K and inhibition of Akt activation, leads to reduced cytokines and decreased expression of survival proteins
LAG-3 (suppressive)	LAG3	Activated T cells, NK cells	MHC class II	❖ Mechanism likely involves the intracellular KIEELE domain
TIM-3 (suppressive)	HAVCR2	Activated T cells in tissues	Galectin-9	❖ Interacts with Bat-3 to remain inactive when not bound to ligand. On ligand binding, Bat-3 dissociates, leading to decreased production of IFNγ and reduced T cell proliferation
KIRs (suppressive)	CD158	NK cells but also APCs and tumor associated CTLS	MHC class I molecules (HLAs)	❖ Inhibitory KIRs induce NK cell tolerance through licensing, in which KIR recognizes self MHC class I molecules thus preventing NK cell activation against self tissue and auto-antigens

Table 2. Checkpoint molecules in T cell inhibition.
The key checkpoints that are involved in T cell inhibition and their cellular effects as well as the ligands are shown in the table. These function to supress the immune response after antigen exposure and activation. Adapted from (Chen and Flies, 2013).

Persistent exposure to antigenic stimulus can occur during chronic viral infections and also within tumours, leading to the progressive loss of T cell functionality and a hyporesponsive state. Experiments using LCMV infection identified a sequential change in virus-specific CD8+ cells upon chronic antigen exposure, with first the loss of IL2 and proliferative capability, followed by loss of TNFα and IFNγ production. T cell exhaustion occurs over a longer period of time (weeks) and coincides with the increase in inhibitory surface receptors PD1, LAG3, TIM3 and CTLA4 (Wherry et al., 2003, Wherry, 2011b, Blackburn et al., 2009). These features of **T cell exhaustion** also exist in tumour-infiltrating lymphocytes (TIL), CD8+ cells in the tumour microenvironment express the same inhibitory receptors (PD1, LAG3, TIM3 and CTLA4), these can also be used as therapeutic targets to restore anti-tumour immunity (discussed in more detail in section 1.5.7 Immunotherapy), blockade of TIM3 and PD1 inhibitory receptors maintains anti-tumour T cell functionality (Sakuishi et al., 2010). The cellular states of **senescence** and **quiescence** refer to cells that are in a non-proliferative state, cells in a state of a permanent cell-cycle arrest are termed senescent where the Hayflick limit (maximum number of replicative events for a cell) has been reached and are often terminally differentiated functional cells, they are often defined by the

shortening of telomeres. Both CD4+ and CD8+ cells can enter this non-proliferative but functional state, which has implications for immune related therapies (Akbar et al., 2016, Hayflick and Moorhead, 1961).

Quiescent cells are in G_0 phase (non-cycling), naïve T cells are in this state of inactivity until they become activated by antigen stimulation and the correct co-stimulatory cues, naïve quiescent cells express low levels of CD5 which increases upon activation, conversely the transcription factor FOXP1 is linked to maintenance of the quiescent phenotype and naïve T cell control (CD8 cells) (Feng et al., 2011, Hamilton and Jameson, 2012).

1.4.2 Humoral immunity and B cells

1.4.2.1 B-cell development and receptor rearrangement

The development of B cells occurs within the bone marrow, an overview of the B cell subtypes is depicted in Figure 2. B cell maturation occurs through sequential steps in the bone marrow prior to release into the periphery, a key stage in B cell development is the formation of an immature B cell with a mature IgM B cell receptor (BCR). The BCR is composed of a heavy chain and light chain, different constant regions make up the different Ig isotypes (IgM, IgD, IgG1-4, IgE and IgA1-2). The unique binding ability is encompassed by the CDR3 leading to a large number of possible BCR combinations and the ability to recognise many antigenic targets including autoantigens (Melchers, 2005). Clonal deletion by apoptosis of strongly autoreactive B cells occurs prior to release from the bone marrow (central tolerance).

1.4.2.2 B cell activation and survival

The Immature B cells (functional IgM BCR) leave the bone marrow as transitional B cells. It has been shown that all B cells exiting the BM display some level of auto-reactivity with variable IgM and BCR signalling (Zikherman et al., 2012). These transitional (T1 and T2) cells enter the Spleen, where they differentiate into marginal zone B cells (MZ) and follicular B cells (FO) depending on the BCR signals. MZ B cells function in a **T-cell independent** manner, providing a defence against bloodborne pathogens by differentiating into extrafollicular plasma cells and secreting IgM. FO B cells are primarily $IgD^{hi}IgM^{lo}$ and require Bruton's tyrosine kinase (BTK) for their development, these cells are mature Naïve B cells and shuttle between blood and secondary lymphoid organs (Martin et al., 2001).

T cell dependent B cell activation requires the interaction of antigen with B cells expressing the corresponding BCR that recognise that antigen, for FO B cells this occurs within a germinal centre (GC) where the FO (naïve) cells express CXCR5 and are retained in the lymphoid tissue by CXCL13 (Ebisuno et al., 2003). The interaction with T helper cells provides co-stimulatory signals, which

occurs within areas defined by a dark zone and a light zone. At the border of the T and B zone the B cells interact with antigen-specific T follicular helper cells (Tfh), help is provided by co-stimulatory molecules and cytokines from Tfh cells (CD40L, ICOS, IL21, IL4 and IFNγ) (Kerfoot et al., 2011). Different B cell subsets (GZMB+ B cells) can also be produced in the presence of IL21 but in the absence of CD40L co-stimulation (Lindner et al., 2013).

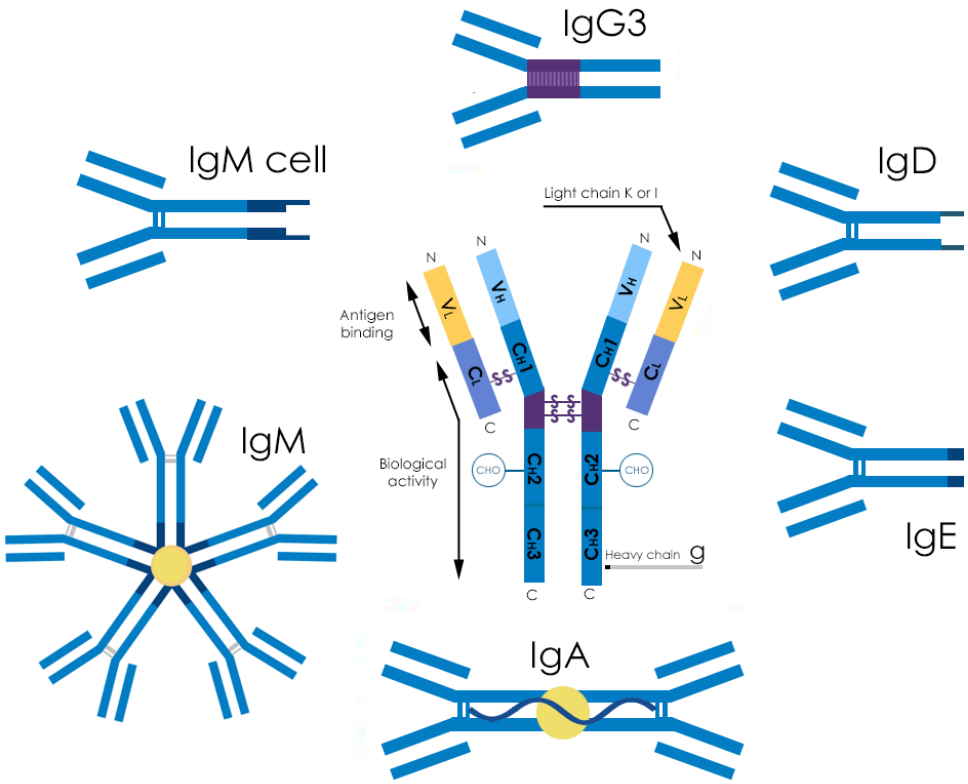
The signals from co-stimulation sustain B-cell proliferation and activate genes needed for class switch recombination (CSR) and somatic hypermutation (SHM), expression of activation induced cytidine deaminase (AID) is a crucial part of CSR and SHM leading to the introduction of point mutations in the Variable region of the Ig gene. This enables selection of high affinity clones as well as the replacement of genes associated with the isotype classes leading to class switching and the different effector functions of the resultant antibodies (IgM to IgG, IgE and IgA) (Chaudhuri and Alt, 2004). GC B cells that have been activated and undergone CSR (via Tfh co-stimulation) differentiate into plasma cells and memory B cells; these display high affinity class switched immunoglobulin and reside in the BM and secondary lymphoid organs (Shlomchik and Weisel, 2012). Extrafollicular activation can also occur leading to memory cell and plasmablast formation independent of the GC reaction (MacLennan et al., 2003). The key function of memory B cells is to remain in secondary lymphoid tissues and rapidly respond to exposure of the same antigen by differentiating into plasma cells and releasing large titres of high affinity class switched antibody (Kometani et al., 2013). This occurs both in an extrafollicular manner and GC-dependant manner, these cells can also undergo further SHM and CSR by re-entering the GC leading to higher affinity BCRs (McHeyzer-Williams et al., 2015, Pape et al., 2011). Long lived plasma cells home to the bone marrow via CXCR4, where they are supported by stromal cells secreting CXCL12 and the cytokine IL6, here they function to produce antibodies (Nutt et al., 2015).

1.4.2.3 B cell effector functions

Effectors functions of B cells are divided into the primary function of producing highly specific unique immunoglobulins, and cellular functions. Secreted immunoglobulins are qualitatively the same as the BCR, but in a non-membrane bound form and, as mentioned, are divided into the different isotypes IgD, IgM, IgG1-4, IgE and IgA1-2 (Fc regions); these regions mediate the effector function and are summarised in Figure 5. The key function is neutralisation of the target and activation of other immune cells such as macrophages and NK cells. This occurs through binding of the Fc region to Fc receptors (FcR) and activation of complement by the binding of C1q. The IgM and IgG3 isotypes are more effective at activating complement, IgG1 also initiates complement activation and activates macrophage and IgE acts on Mast cells (Bruhns and Jonsson, 2015). The cellular functions of B cells consist of antigen presentation, co-stimulation, tertiary lymphoid organ formation and cytokine release, both inhibitory and stimulatory. B cells present antigens by

internalising the antigen:BCR complex and processing it in to peptides where it is then bound to MHC-I and II, this allows B cells to act as antigen presenting cells to CD8+ and CD4+ cells (Yuseff et al., 2013).

A



B

Ig subtype	Serum (%)	Structure	Complement Fixation	Opsonizing	Cross Placenta	Other Functions	FcR
IgG	75%	Monomer	+	+++	+	For all IgG: subclasses	FcγR
-IgG1	67% IgG	Monomer	Yes	Yes	+	Secondary response	I,II,III
-IgG2	22% IgG	Monomer	Yes	Yes	+	Neutralize toxins	II
-IgG3	7% IgG	Monomer	Yes	Yes	+	Neutralize virus	I,II,III
-IgG4	4% IgG	Monomer	No	No	+		I,II
IgM	10%	Pentamer	+++	+	–	Primary response	
IgA	15%	Monomer, Dimer	–	–	–	Mucosal response	FcαR (CD89)
-IgA1		Monomer, Dimer	–	–	–		
-IgA2		Monomer, Dimer	–	–	–		
IgD	<0.5%	Monomer	–	–	–	homeostasis	FcδR
IgE	<0.01%	Monomer	–	–	–	Allergy	FcεR I,II

Figure 5.Immunoglobulin structure and function.

(A) The structure of the different sub classes of immunoglobulin; IgM, IgG, IgA, IgD and IgE adapted from (Bruhns and Jonsson, 2015). (B) This highlights the different roles of each of the antibody subtypes and the relative percentages found in the serum. Adapted from (Bruhns and Jonsson, 2015, Schroeder and Cavacini, 2010)

B cells also play a key role in the formation of lymphoid organs and germinal centres by the release of lymphotoxin alpha (LTA), thus directly influencing the co-localisation of follicular dendritic cells and CD4+ cells (Tfh cells) (Fu et al., 1998). B cells can also exert effector functions through cytokines that act on CD4+ T cells and modify their polarisation state. Thus, production of

TNFα and CCL3 promotes Th1 cell differentiation leading to increased IFNγ from Th1 cells. IL-6 from B cells supports Th1, Tfh and Th17 cells, IL-17 promotes the production and accumulation of neutrophils and IL-4 supports a Th2 response (Vazquez et al., 2015, Shen and Fillatreau, 2015). The other important role of B cells is as negative regulators of the immune response, these cells are termed Bregs. The immunosuppressive role of Bregs falls into the release of cytokines in the form of IL35, IL10 and TGFβ that act to impede the Th1 and Th17 responses as well as increase Treg generation (IL10 mediated) (Olkhanud et al., 2011). The cellular interactions of Bregs via PD-L1, OX40L, CD80/CD86 and FasL also act to impair T cell mediated immunity (Zhou et al., 2016). Regulatory B cells fall into different subsets depending on their mechanism of suppression with B10 plasmablast cells expressing IL10, GZMB B cells (GZMB, IL10 and IDO), BR1 cells (IL10 and IgG4), induced iBregs (TGFβ and IDO) and i35Bregs (IL35) (Rosser and Mauri, 2015, Mauri and Menon, 2015).

1.4.3 Innate Immunity

The Innate immune system works in partnership with the unique antigen recognition of the adaptive immune system. Innate immunity acts as the first line of defence against infection, it includes an array of specialised cells and molecular mechanisms that can be used to eliminate infections. This starts with the anatomical barriers that are the epithelial surfaces of skin, mouth, stomach, intestines and the airways that all play an integral role in controlling infections and are a key component of the innate immune system (Turvey and Broide, 2010). The structural epithelial cells of barrier tissues are equipped with some basic protective mechanisms that enable them to sense and respond to potential danger in the form of invasive microbes. In addition, there is a range of specialised cells in close anatomical proximity to the barriers, with a greater range of defensive and protective functions, these are regarded as forming the core of the innate immune defences.

1.4.3.1 Pattern recognition receptors of Innate Immunity

The innate immune system detects the presence of potentially harmful organisms/ cells (e.g. bacteria and viruses) through the pattern-recognition receptors (PRRs). These are divided into the Toll-like receptors 1-10 (TLRs), C-type lectin receptors (CLRs), nucleotide oligomerisation domain receptors (NLRs) domain, RIG-I like receptors (RLRs) and AIM2-like receptors. They recognise pathogen-associated molecular patterns (PAMPs), which are conserved molecular structures specific to particular microorganisms (Kawai and Akira, 2010). TLRs and CLRs are predominantly located at the cell surface and in endocytic compartments, the NLRs, RLRs and ALRs are unbound

and located in the cytoplasm(Yang et al., 2017), Table 3 summarises the various PRRs and PAMPs of the innate immune system.

1.4.3.2 Cells of the innate immune system

The main specialised cellular components of innate immunity are shown in Figure 2, they consist of macrophages, dendritic cells, natural killer cells and granulocytes (neutrophils, basophils and eosinophils).

Monocytes are a subset of circulating cells with the ability to migrate to areas of tissue damage and inflammation, this occurs via the expression of various chemokine receptors (e.g. CCR2, CXCR1, CCR1 and CCR5) and adhesion molecules (e.g. L-selectin, LFA1 and VLA4). These allow the cells to attach and migrate to sites of tissue damage and inflammation (Shi and Pamer, 2011). At sites of inflammation monocytes have the ability to differentiate into plasmacytoid dendritic cells (pDC), pre-cursor DC and Macrophages (Geissmann et al., 2010).

Macrophages: Once at the site of tissue damage and in the presence of the correct environmental cues differentiate and polarise into different subsets of cells depending on the type of signals. These have distinct cytokine, cell-surface markers and expression profiles on which they are broadly divided into the classical M1 and the alternative M2 macrophages (Mantovani et al., 2002). The M1 phenotype is associated with the Th1 cytokine IFNγ and bacterial stimulus such as lipopolysaccharides (LPS), they produce the pro-inflammatory cytokines IL-6, IL-12 and IL-23 as well as TNFα. The alternative M2 macrophages are more pleiotropic and differentiate into M2a (IL-4 and IL-13 stimulus); M2b (LPS and Ig immune complex stimulus); M2c (TGFβ and IL-10 stimulus) and the M2d which occur in response to tumour-derived factors (Movahedi et al., 2010). The M2 phenotype is broadly associated with a pro-angiogenesis and anti-inflammatory response, this includes the production of TGFβ, IL-10, CCL17, CCL22 and CCL24 (Qian and Pollard, 2010, Roszer, 2015).

Dendritic cells are specialised cells that function as potent activators of the immune response, they are divided into conventional DCs (cDCs) and plasmacytoid DCs (pDCs). They are derived from common lymphoid and myeloid progenitor cells in the bone marrow, pre-DCs migrate through the blood to peripheral tissue where they differentiate into cDCs (either CD8α+/CD103+ or CD11b+). The highly migratory cDCs are found at peripheral tissue sites where they function as specialised antigen presenting cells that capture and process antigen for presentation to cells of the adaptive immune response in the lymphoid organs, this subsequently leads to activation of an immune response. pDCs terminally differentiate in the bone morrow and migrate to lymphoid organs and tissue compartments. The role of pDCs is to respond to viral infection via TLR7 and 9, upon stimulation large amounts of IFNα and INFβ are produced (Schraml and Reis e Sousa, 2015, Geissmann et al., 2010).

PAMP Detection by TLRs and Other PRRs			
Species	PAMPs	TLR	PRRs Involved in Recognition
Bacteria, mycobacteria	LPS	TLR4	
	lipoproteins, LTA, PGN,	TLR2/1, TLR2/6	NOD1, NOD2, NALP3, NALP1
	DNA	TLR9	AIM2
	RNA	TLR7	NALP3
Viruses	DNA	TLR9	AIM2, DAI, IFI16
	RNA	TLR3, TLR7, TLR8	RIG-I, MDA5, NALP3
	structural protein	TLR2, TLR4	
Fungus	zymosan, β-glucan	TLR2, TLR6	Dectin-1, NALP3
	Mannan	TLR2, TLR4	
	DNA	TLR9	
	RNA	TLR7	
Parasites	tGPI-mutin (Trypanosoma)	TLR2	
	glycoinositolphospholipids (Trypanosoma)	TLR4	
	DNA	TLR9	
	hemozoin (Plasmodium)	TLR9	NALP3
	profilin-like molecule (Toxoplasma gondii)	TLR11	

Table 3. The recognition of pathogens by the innate immune system.
Pathogen-associated molecular patterns (PAMPs) are recognised by Pattern-recognition receptors (PRRs) on immune cells. Innate immune cells in particular use the PRRs to react quickly to these antigens. Adapted from (Brubaker et al., 2015).

Natural killer cells (NK) have the ability to bind specific regions of Ig FC receptor via the expression of CD16a/b (FCGR3A/B) and CD32 (FCGRIIC, a low affinity IgG receptor)(Anderson et al., 1990). The binding of FcR by NK cells leads to antibody dependant cell cytotoxicity (ADCC), where degranulation of cytotoxic particles (perforin and granzyme) occurs causing cell cytotoxicity (Smyth et al., 2005).

Natural Killer T-cells (NKT) are innate T lymphocytes that undergo thymic selection and possess a functional T cell receptor (TCR), they recognise lipid antigens presented by CD1d as well as cytokine receptors such as IL-12, IL-18, IL-25 and IL-23 (Cohen et al., 2013). They are divided into two main subtypes; Type I NKT or invariant NKT (iNKT) and Type II NKT or variant NKT (Exley et al., 2001).

Granulocytes (Neutrophils basophils and eosinophils) play an important role in the immune response clearing pathogens and immunoregulation, they originate from myeloid progenitor cells in the bone marrow and are released into the circulation fully differentiated.

Neutrophils are recruited to sites of infection and inflammation and enter tissues via a cascade of cell migration-where cytokines and chemo-attractants signals originate (Sundd et al., 2013), (Chou et al., 2010). Activation can occur via an array of different signals that consist of pattern recognition receptors (PRR), G protein-coupled receptors (GPCRs) and opsonising receptors in the form of FcγR and complement. One of the main PRRs in neutrophils are the Toll like receptors (TLR1, 2, 4, 5, 6, 8, 9 and 10), which recognise DNA, RNA, Lipids, carbohydrates and peptides (Trinchieri and Sher, 2007). Opsonising receptors expressed by neutrophils include part of the complement system; CR3, CR4 and CR1 as well as high- and low-affinity IgG FC receptors (Bruhns, 2012) Once activated neutrophils are able to exert killing effects via receptor mediated uptake, this involves phagocytosis and the formation of a phagosome with opsonised IgG targets, production of reactive oxygen species (ROS), as well as release of three types of cytotoxic granules during degranulation. (Vethanayagam et al., 2011).

Basophils are recruited to sites of inflammation and are activated by FCERI aggregating with IgE, this is enhanced in the presence of IL-3, IL-5 and GM-CSF, TLR2 and TLR4 leading to activation. Once activated they release preformed mediators stored in granules, these contain heparin, histamine and leukotrienes (lipid mediators). The cytokines produced included IL-4, IL-13 and GM-CSF, in particular IL-4 which is also involved in Th2 cell differentiation (Stone et al., 2010, Sullivan and Locksley, 2009).

Eosinophils are released from the bone marrow in response to IL-5 and traffic to the tissues and sites of infection (helminth) and inflammation. They are often located at sites where a Th2 response has dominated, IL-4 and IL-13 play a crucial role in the trafficking of the cells to the tissues. Activation of Eosinophils can occur via IgG and IgE. Once activated they release proinflammatory granules that contain major basic protein, a highly cationic protein that interacts with the cell membrane on parasites, they also release lipid mediators (leukotrienes) (Uhm et al., 2012, Rosenberg et al., 2007, Hogan et al., 2008).

1.5 Immune response to viral infection

During a viral infection the innate immune system reacts immediately (hours), sensing the molecular patterns associated with the pathogen via the PAMP receptors. The adaptive immune response develops over a longer period of time (7 days) and is in part triggered by the innate response. Ultimately the goal is to clear the pathogen and prevent further infection by the production of immunological memory.

1.5.1 Innate immune reaction to virus

The innate non-specific early antiviral response involves pathogen recognition (mostly DCs), followed by IFN and cytokine release and recruitment of other innate cells (NK cells, neutrophils and macrophages). The sensors of the innate system consist of TLRs with endosomal nucleic acids being detected by TLR3 (dsRNA), TLR7 and TLR8 (ssRNA) and TLR9 (CpG and DNA viruses), viral envelope glycoproteins are detected by TLR2 and TLR4 (Finberg et al., 2007, Barbalat et al., 2009). Ultimately the activation of the TLR leads to NFkB activation and type I IFN production (Takeuchi and Akira, 2009). Intracellular sensing of viral RNA by NLR and DNA by ALR leads to the release of IL-1β (Rehwinkel et al., 2010). The stimulator of IFN genes (STING) play a key role in RLR function, increased expression of STING is linked with subsequent increases IFN release (Ishikawa and Barber, 2008). Increased production of IFN type I (α/ β) leads to an antiviral response, pDCs are specialised IFN producers and sense viral RNA and DNA via TLRs (TLR7, 8 and 9) (de Veer et al., 2001, Haller et al., 2006).

1.5.2 Adaptive immune response to virus

During an adaptive immune response to virus, a co-ordinated reaction takes place involving the processing and presentation of the viral proteins on MHC-I and MHC-II by APCs (e.g. DCs) as well as the infected cells in the case of MHC-I. This leads to the activation of cell mediated immunity via CD8+ cells (MHC-I) and CD4+ cells (MHC-II), with a prominent CD4+ Th1 response. The Th1 response is characterised by production of the cytokines IL-2, IL-12, IL-23, IL-27 and IFNγ, this in turn acts to help drive clonal expansion of CD8+ cells, as well as helping virus specific B cells switch to IgG1 isotypes involved in ADCC (Strutt et al., 2013, Trinchieri, 2003). CD8+ cells (CTLs) function as the key defence mechanism against virus infected cells by releasing perforin and granzymes that induce apoptosis in the infected cells (Lawrence et al., 2005). Immune response to HPV-16 is covered in section 1.6.7 (Immune response to HPV driven cancer), where the oncogenic viral proteins E6 and E7 play a key role in the adaptive immune response to the virus by generating both an adaptive T cell (E6/E7 specific CD8 cells) as well as a B cell response (HPV16(+) antibodies) to the virus (Frazer, 2009).

1.6 Interaction of the host immune system and cancer

1.6.1 Tumour Infiltrating Lymphocytes (TIL) and survival

The presence of tumour infiltrating lymphocytes (TIL) is regarded as reflecting the immune response against the tumour. The density of TILs is linked to an improved prognosis and is

associated with an active immune response within the tumour (Galon et al., 2006). A key piece of work in colorectal tumours showed a significant clinical benefit of immune cells infiltrating the tumour, both at the invasive margin and within the tumour. Tumours with high densities of immune infiltrate (CD3) show a much better prognosis when compared to immune low tumours (Galon et al., 2006). The type of cells that are involved in the infiltrate are also pivotal, with CD8+ and Th1 cells linked to the best prognosis, whereas Th2, Th17 and Treg numbers are generally linked to a poorer outcome (Fridman et al., 2012). In metastatic melanoma, a similar pattern has been observed, with low, moderate and high immunotypes as determined by H and E. This showed that the most prognostic immune cell as well as the most numerous, was the CD8+ T cell, again with higher global immune densities showing the best survival (Erdag et al., 2012).

The same clinical benefits were observed in HNSCC when tumours were characterised as immune cell high, moderate or low as determined by H and E staining, where higher immune cell densities conferred an improved survival. This classification was able accurately to group cases in the same way as using cell-specific markers (CD4+ and CD8+). This study also shows a highly improved survival in HPV(+) tumours compared to HPV(-) HNSCC (Ward et al., 2014b). Patients that display high levels of CD8+ T-cell tumour invasion have increased disease free survival, with the level of functional CD8+ CTL (GZMB+) infiltrate being directly correlated to tumour burden and prognosis (Bindea et al., 2011, Galon et al., 2006).The presence of TLS (tertiary lymphoid structures) in tumours broadly predicts for a better outcome, these are often marked by high densities of B cells that aggregate to form the TLS in the tumour, it is indicative of an on-going immune reaction involving the major cell types (CD4+ and CD8+) that encounter antigen stimulus and respond (Germain et al., 2014, Sakimura et al., 2017).

1.6.2 Immune response and survival in HNSCC

HPV(+) tumours are now treated as a separate group to HPV(-) tumours due to differences in presentation and prognosis. HPV(-) tumours usually arise on the background of alcohol and smoking, conversely the HPV(+) tumours as discussed previously arise due to the cellular transformation linked to the viral oncoproteins, the clinical presentation and prognosis is also different between the two tumour groups (Pai and Westra, 2009). HPV(-) tumours have a lower lymphocytic infiltration of CD8+ T-cells and B-cells compared to HPV(+) (Russell et al., 2013). The CD8 infiltrate in HPV16 disease in one study was 6 times higher than HPV(-) with a greatly improved 3-year overall survival (Nordfors et al., 2013). The prognostic value of TIL in HNSCC shows a direct relationship between the level of immune infiltration and survival, in one study the presence of a high/ moderate global TIL conferred improved prognosis compared to low

infiltration. The 3-year survival of HNSCC patients was 82% in HPV(+) vs 56% in HPV(-), the level of TIL was strongly correlated to HPV(+) tumours with 85% containing high/ moderate TIL. The 3-year survival in patients with HPV(+)/ TIL^{Hi} tumours was 96%, HPV(+)/TIL^{Mod} was 76% and HPV(+)/TIL^{Lo} was 59% (Ward et al., 2014b). The HPV(+)/ TIL^{Lo} displayed a similar survival to that of the HPV(-) at 56%, this is an interesting group that may allow insights into the cellular and molecular state that is required for tumour infiltration when compared to the HPV(+)/TIL^{Hi}. The predictive values of individual cell types were also assessed in terms of total TIL, CD4, CD8, Foxp3, CD4:CD8 ratio and Foxp3:CD8 ratio, the prognostic model indicated that the overall TIL (High/ Mod and Low) were equally as informative a prognostic maker as the individually scored cell populations Ward et al. (2014b). The difference in survival between patients with HPV(+) and HPV(-) tumours and the level of immune infiltrate highlights the possibility of a different immune response between virally driven and virus-independent tumours; analysis of immune rich tumours and comparison of immune functions in HPV(+) and HPV(-) tumours is therefore important as it may give insight into these survival differences.

The overall immune composition of HNSCC using flow cytometry was characterised by T-cell infiltrates that display an effector memory phenotype (70% of CD3 cells = $CD45RA^{-}/CCR7^{-}$), an increase in T-regulatory cells (12% of CD3 cells = $CD25^{+}/CD127^{Low}/CD4^{+}$) and higher expression of immune check point markers PD1 and CTLA4 relative to PBMC controls. The overall tumour infiltrating lymphocyte level was also highly variable and displayed higher levels in HPV(+) tumours relative to HPV(-) (Lechner et al., 2017). In HPV(-) HNSCC single cell transcriptomics identified a varied immune infiltrate, this consisted of T-cells, macrophages and dendritic cells. The T-cells were divided into cytotoxic CD8 T-cells (exhausted and non-exhausted), with the CD4 cells divided into Tregs and CD4 T helper cells. The cell types were expressed across multiple patients but varied in their overall proportions, interestingly there were little or no B cells identified (Puram et al., 2017, Cillo et al., 2020).

1.6.3 Immune infiltrates and survival in NSCLC

The extent of the immune cell infiltrate in NSCLC is also directly linked to patient survival, where the overall 5-year survival for TIL^{Hi} tumours was 60% compared to 21% in TIL^{Lo} tumours, as determined by global TIL scoring of haematoxylin and eosin stained tumours (Horne et al., 2011). The Patient survival also differs between Adenocarcinoma and Squamous cell carcinoma, where 5-year overall survival was 42% and 63% respectively (Moldvay et al., 2000). A more focused immune score using CD8+ TIL yielded similar results with 5-year survival at 61% (CD8+ TIL^{Hi}) and 41% (CD8+ TIL^{Lo}) (Donnem et al., 2015). Tumours are heterogenous, the location of the CD8+ immune cells (CD8+ TIL^{Hi}, ^{Mod} and ^{Lo}) can have an impact on survival. Whether the whole tumour,

representative areas, central tumour, invasive margins, dense lymphoid aggregates or random area sampling are used affects the overall survival. The most concordant survival data was derived from random sampling 20% of the tumour, central tumour, conversely the invasive margin led to poor concordance and significance (Obeid et al., 2017).

Immune profiling of NSCLC shows that the tumour infiltrating lymphocytes consist of a wide range of immune cells that recapitulates an on-going adaptive immune response T-cells (CD4 and CD8) B cells, NK cells, and macrophages are all identified at varying proportions across a patient cohort see Table 4 (Lizotte et al., 2016) (Stankovic et al., 2018).

Cell type	Gating strategy	Cell % in NSCLC	Number of samples
Live Leukocytes	$CD45^+PI^-$	100	n = 68
T cells	$CD45^+PI^-CD3^+CD19^-$	46.5	n = 43
$CD4^+$ T cells	$CD45^+PI^-CD3^+CD19^-CD8^-CD4^+$	25.9	n = 43
$CD8^+$ T cell	$CD45^+PI^-CD3^+CD19^-CD8^+CD4^-$	21.7	n = 30
$CD4^-$ $CD8^-$ T cells	$CD45^+PI^-CD3^+CD19^-CD8^-CD4^-$	1.4	n = 30
B cells	$CD45^+PI^-CD3^-CD14^-CD19^+$	15.9	n = 55
Naive B cells	$CD45^+PI^-CD3^-CD14^-CD19^+IgM^+IgD^+$ $CD27^-$ $CD38^{+/-}$	1.6	n = 23
$CD27^+CD38^{+/-}$ B cells	$CD45^+PI^-CD3^-CD14^-CD19^+IgM^-IgD^-$ $CD27^+CD38^{+/-}$	4.2	n = 23
Plasma cells	$CD45^+PI^-CD3^-CD14^-CD19^+IgM^-IgD^-$ $CD27^+CD38^{++}$	0.8	n = 23
IgM^+IgD^- B cells	$CD45^+PI^-CD3^-CD14^-CD19^+IgM^+IgD^-$	2.2	n = 23
Macrophages	$CD45^+PI^-CD19^-CD14^+$HLA-DR^+	4.7	n = 33
pDCs	$CD45^+PI^-CD19^-CD14^-$HLA-$DR^+CD11c^-CD123^+$	1.2	n = 29
Classical DCs	$CD45^+PI^-CD19^-CD14^-$HLA-DR^+CD11c^+	1.6	n = 29
$CD1c^+$ DCs	$CD45^+PI^-CD19^-CD14^-$ HLA-DR^+CD11c^+ $CD1c^+CD141^-$	0.8	n = 29
$CD141^+$ DCs	$CD45^+PI^-CD19^-CD14^-$ HLA-DR^+ $CD11c^+CD1c^-CD141^+$	0.1	n = 29
"DN DCs"	$CD45^+PI^-CD19^-CD14^-$HLA-DR^+CD11c^+ $CD1c^-CD141^-$	0.5	n = 29
NK cells	$CD45^+PI^-CD19^-CD14^-CD3^-CD56^+$	4.5	n = 18
$CD16^+$ NK cells	$CD45^+PI^-CD19^-CD14^-CD3^-CD56^+CD16^+$	2.3	n = 18
$CD16^-$ NK cells	$CD45^+PI^-CD19^-CD14^-CD3^-CD56^+CD16^-$	2.2	n = 18
Neutrophils	$CD45^+PI^-CD19^-CD3^-C14^-CD11b^+CD15^+CD49d^-$	8.6	n = 18
Basophils	$CD45^+PI^-CD19^-CD3^-C14^-CD11b^-Fc\varepsilon R1\alpha^+CD117^-CD49d^+$	0.4	n = 18
Eosinophils	$CD45^+PI^-CD19^-CD3^-C14^-CD11b^+CD15^+CD49d^+$	0.3	n = 18
Mast cells	$CD45^+PI^-CD19^-CD3^-C14^-CD11b^-Fc\varepsilon R1\alpha^+CD117^+CD49d^+$	1.4	n = 18

Table 4. The Immune cell composition in NSCLC.
The tumour infiltrating lymphocytes in NSCLC are diverse and represent a wide range of cells associated with and on-going adaptive immune response. Adapted from Stankovic et al., 2018

Further evaluation of the tumour ecosystem in NSCLC using single cell transcriptomics, again shows a highly heterogenous immune infiltrate that represents immune response to the tumour. B cells, macrophages and T cells were all elevated in the tumour relative to control tissue. CD8 T-cells exhibited cytolytic (GZMA/ B and PRF1) features, as well as exhaustion features (PD1 and CTLA4) , CD4 cells were divided into helper and regulatory. B cells displayed signs of activation and class switching indicative of germinal centre reaction (Chung et al., 2017, Peng et al., 2019, Sade-Feldman et al., 2018, Lambrechts et al., 2018, Tirosh et al., 2016, Puram et al., 2017).

1.6.4 Innate response to tumour

The key innate cells that elicit a response to tumour are NK cells, NKT cells, γδ cells, DCs and macrophages that all play a role in the response via DAMPs. NK cells in the context of the tumour environment can be activated by the MICA/ B (MHC-I chain related molecules A/B) ligands that are expressed on tumour cells under stress, these are recognised by the NKG2D receptors on NK cells and lead to direct activation and release of cytotoxic granules and expression of FASL/ TRAIL (Gasser et al., 2005, Cheng et al., 2013). IFNγ from NK cells also functions to activate CD8 T cells and skew the CD4+ response towards Th1 (Martin-Fontecha et al., 2004). NKT (type I) cells can be activated by glycolipids like tumour associated ganglioside GD3 that are recognised by CD1d, the activated NKT cells can rapidly produce cytokines and cytotoxic activity, expressing IL-2, IL-4, and IFNγ (Wu et al., 2003, Coquet et al., 2008). NKT (type II) cells have been shown to play a negative role via the action of IL-13 from type II NKT cells, this down regulated tumour immune surveillance in a mouse model via the IL4R- STAT6 pathway (Terabe et al., 2000). γδ cells function to recognise phospho-antigens that are expressed on tumour cells, accumulation of mevalonate pathway metabolites in tumours act as ligands for TCRs on γδ cells leading to activation and direct killing of the tumours cells via cytotoxic granule release and FasL/ TRAIL. Activated γδ cells also produce IFNγ and TNFα contributing to the control of tumour growth (Gober et al., 2003, Kondo et al., 2008).

1.6.5 Cellular adaptive response to tumours

In order to elicit an adaptive immune response to a tumour the immune system must become activated to antigen, in the tumour this can take a variety of different forms (Table 5). Tumour associated antigens (TAA) are presented to T-cells via MHC-I/II and to B cells via the BCR, with the activation process mirroring what is seen in normal secondary lymphoid organs and peripheral tissues. Some of the classical TAAs are the MAGE (melanoma antigens) proteins that are normally confined to the testis, CEA (carcinoembryonic antigen) often overexpressed in colorectal cancer, PSA (prostate specific antigen) expressed in the normal prostate and elevated in prostate cancer and Her2 which is over expressed in breast cancer.

A different type of TAA occurs as a result of genetic mutations yielding proteins that are foreign to the immune system, these can be in the form of fusion proteins like BCR-ABL in AML and CLL, while others arise from splicing mutations such as K-RAS mutations (Novellino et al., 2005). Further classification of these mutations has led to the term neo-antigens, where novel peptides arise as result of frameshifts and point mutations in normal proteins, this yields a protein that is then perceived as foreign to the immune system eliciting an immune response. The number of these mutations is linked to a higher immune infiltrate and increased CTL action as well as

improved patient survival (Brown et al., 2014). This was expanded further by the assessment of 18 tumour types linking MHC-I neo-antigens to cytolytic activity (GZMA and PRF1), cervical cancer (HPV(+)) displayed the highest CTL+ predicted neo-antigen activity, thyroid on the other hand showed the least (Rooney et al., 2015). Targeting of these mutations has led to successful immune responses that can be used as a custom therapeutic target, where neo-antigen specific RNA vaccines can yield peptide specific responses in human patients yielding an improvement in disease control (melanoma)(Sahin et al., 2017, Ott et al., 2017).

Tumour-associated antigens that are recognised by T cells		
Classification of tumour antigen	**Mechanism of immune activation**	**Example**
Cancer-testis antigens	Normal expression found in spermatocytes in testis and occasionally placenta (immune-privileged site), triggering T cell response.	MAGE (melanoma antigen) BAGE (B antigen) GAGE (G antigen)
Differentiation antigens	Antigen is expressed by the tumour and the normal tissue from which it arose.	CEA – expressed in embryonic tissue and over-expressed in colorectal cancer. Gp100 – expressed in melanocytes and melanoma. PSA – expressed in normal prostate and over-expressed in prostate cancer.
Over-expressed tumour associated antigens	Level of expression in normal tissue is below the threshold for T cell activation. Over- expression by cancers cells triggers T cell activation.	Her2 – over-expressed in breast cancer. AFP – over-expressed in hepatocellular cancer and certain germ cell tumours.
Tumour-specific antigens	These arise from genetic mutations or splicing aberrations, generating a protein that is foreign to the host immune system.	Neoantigens unique to each patient tumour.
Fusion proteins	Chromosomal translocation results in fusion of distant genes and expression of an abnormal fusion protein that is foreign to the host immune system.	BCR-ABL in CML and some ALL. EML4-ALK in non-small cell lung cancer.

Table 5. T cells reactivity to Tumour antigens.
This shows some of the various types of tumour antigens that are able to generate an immune response and the mechanism by which it occurs. Adapted from (Novellino et al., 2005).

1.6.6 Humoral adaptive response to tumours

The humoral response to tumours requires engagement of the CD4+ T helper compartment to mount an effective B cell mediated response to tumour. The role of tumour infiltrating B cells rests with an interplay between both anti-tumour and pro-tumour activity, positive roles of B cells in tumour clearance are associated with ADCC, Granzyme B producing B cells and antigen presentation. The negative role of B cells are linked to regulatory cells that function to inhibit the immune response (Breg), inhibitory Immunoglobulin classes (IgA) and expression of negative regulators of the immune response like PDL1 (Olkhanud et al., 2011) (Park et al., 2010).

1.6.6.1 Positive role of B cells in tumour immunity

The process of initiating a protective immune response to tumour requires the interplay of the key cells of the immune response, both the CD4+ and CD8+ T cells as well as B cells providing co-stimulatory cues both direct and in-direct (CD40-CD40L and cytokine release). One of the key features that appears to outline an ongoing and sustained immune response to antigenic tumours is the formation on TLS (tertiary lymphoid structures) within the tumour; these resemble the anatomical structures observed in the secondary lymphoid organs. A higher number of TLS and increased B cell densities confer a survival benefit in NSCLC, this is also seen in HNSCC (Germain et al., 2014).

The role of ADCC and tumour auto-antibodies also plays a role in anti-tumour activity, HER2 expressing breast and ovarian tumours can be targeted by endogenous anti-HER2 antibodies (Montgomery et al., 2005); mutated proteins such calreticulin in hepatocellular cancer also generate an antibody response. These tumour-associated antigens are currently being used as biomarkers of on-going disease rather than for any therapeutic applications (Le Naour et al., 2002). B cells that have not received sufficient co-stimulation from CD40L on CD4+ Th cells in the presence of IL21 shift from plasma cell formation, to granzyme B+ B cell production. These cells are not able to function in the same manner as CTLs as they lack perforin, although the mode of action potentially revolves around antiviral/ cytolytic function, auto-regulation and regulatory and suppressive effects on other immune cells(Hagn et al., 2012, Hagn and Jahrsdorfer, 2012, Lindner et al., 2013). B cells also function as antigen presenting cells and express both MHC-I and MHC-II, presentation of antigen to CD4+ cells via MHC-II leads to CD4+ cell priming and a Th1 response (Rodriguez-Pinto, 2005, Bruno et al., 2017). The primary function of a B cell is to produce antibody (IgG1-4, IgA1-2 and IgM), with IgG1 and IgG3 functioning through opsonisation, NK cell activation and complement activation, IgM also activates the complement system, most treatment regimens (e.g. Herceptin) that use antibodies use the IgG1 isotype (Beers et al., 2016, Vidarsson et al., 2014, Shi et al., 2015).

1.6.6.2 Negative role of B cells in tumour immunity

Regulatory cells within the tumour microenvironment impede the immune response, B regulatory cells (Bregs) are those that arise from the B cell lineage and have a negative effect on the immune response. This is commonly attributed to the secretion of the suppressive cytokines IL10, IL35 and TGFβ. IL10 plays a role in suppressing IFNγ and IL17 production from Th1 and Th17 cells, it also induces the formation of Treg (FOXP3+) cells (Olkhanud et al., 2011). Bregs have also been shown to reduce secretion of IFNγ in cytotoxic CD8 cells. The secretion of TGFβ by B cells functions to modulate CD4+ cell proliferation, CD8+ cell cytotoxicity and Treg formation (Sarvaria et al., 2017,

Zhou et al., 2016, Flores-Borja et al., 2013). Bregs also play a role in macrophage differentiation, depletion of B cells in SCC re-programs the tumour associated macrophage (TAM) phenotype from a Th2-TAM expressing FC gamma receptors that respond to the immune complex deposited by B cells, to a Th1-TAM antitumor phenotype that express anti-angiogenic chemokines CXCL10/ 11, expression of CCR5 also leads to a greater infiltrate of CD8 and improved survival in mice (Affara et al., 2014).

The chronic antigen exposure that occurs in tumours may also lead to a hypofunctional state and an exhausted B cell phenotype, this has been extensively studied in persistent viral infections (HIV). In this situation B cells are CD21lo and CD27lo with the expression of an inhibitory tissue marker FCRL4, the loss of CD21 is a marker of exhaustion, coupled with a lack of CD95 expression in the presence of persistent antigen (Ehrhardt et al., 2005, Moir et al., 2008), (Sciaranghella et al., 2013). Antigen presentation by B cells to CD4+ cells can also alter their phenotype, this depends on the functional state of the B cells. Activated B cells are associated with a Th1 response, whereas B cells with an exhausted phenotype are associated with the formation of CD4+ regulatory cells (FOXP3(+)) (Sciaranghella et al., 2013). The particular Ig isotype produced by the plasmablast also plays an important role in antitumor immunity, IgG4 subclasses accumulate in tumours with an IL10 driven Th2 pro-inflammatory response (IL10, IL4, IL13), this class switch was linked to a poorer patient outcome and impaired antitumor activity in Melanoma (Karagiannis et al., 2013). Immunosuppressive IgA+IL10+PDL1+ plasmocytes (induced by TGFβ) have been shown to impede treatment of prostate cancer (mouse model and human), with subsequent depletion of these cells leading to effective tumour eradication. Inflammation induced IgA+PDL1+ B cells have also been shown to impede the activity of CD8 CTL function directly, these cells directly suppressed CD8+ cells in liver cancer. The disruption of either IgA or PDL1 led to eradication of the tumour and regression of tumours through the released activity of CD8+ CTL in hepatocellular carcinoma (Shalapour et al., 2015, Shalapour et al., 2017).

1.6.7 Immune response to HPV driven cancer

1.6.7.1 Immune response HPV16

HPV driven HNSCC exhibits unique tumour characteristics with known antigenic components that can generate an immune response. The innate immune response triggered by HPV consists of TLR activation involving TLR9 (unmethylated CpG dsDNA), TLR3, 7 and 8 (dsDNA) leading to the production of IFN-α, IFN-β and IFN-γ (Akira and Takeda, 2004). Tumour infiltrating macrophages can also play a role in both tumour suppression and promotion, the activity type of the macrophage is dependent on the differentiation pathway with classically-activated M1 (IFN-γ and LPS) leading to anti-tumour activity via immune stimulation and production of inflammatory

cytokines; alternatively-activated macrophages M2 (IL-4, IL13, IL10) cause tumour promotion via tissue repair and angiogenesis (Quatromoni and Eruslanov, 2012, Tang et al., 2013).

During an on-going infection with HPV16, T cell responses are found in the form of circulating HPV-reactive CD4+ and CD8+ cells, 63% of patients are reactive to the E6 and E7 protein in patients with HPV16(+) driven HNSCC. However, E6/E7 reactive cells have also been identified in 24% of HPV16(-) lesions arising as a result of a previous exposure to the virus (Heusinkveld et al., 2012). A local T cell response to HPV16(+) tumours has been found in most patients and is determined by the number of tumour infiltrating lymphocytes that are specific to HPV16. This consists of CD4+ T-helper (Th1 and Th2), CD8+ T-cells and CD4+ T-regulatory cells (Treg) that are specific to E6 and E7 epitopes. These also produce a range of cytokines in the form of IFNγ, TNFα, IL-5 and IL-4 (Heusinkveld et al., 2012). The balance between CD8+ T-cells and Tregs (CD4, CD25, Foxp3) has also been linked to a better prognosis; thus, in addition to the total CD8 T-cell infiltration, the CD8:Treg ratio can also be used as a prognostic indicator (Nasman et al., 2012, Ward et al., 2014b).

In cervical cancer (HPV driven) serum antibodies to the HPV16 virus can be detected in 30% to 53% of low to high-grade cervical lesions respectively (control subjects = 16%), this indicates that both a T-cell and B-cell response to the HPV16 virus can be generated (Wideroff et al., 1995). Data from HPV16 driven CIN (cervical intraepithelial neoplasia) mirrors what is seen in HNSCC with a T-cell response to HPV16 proteins (E6, E7, E4, L1 and L2) in 78% cases, in this particular study there was a correlation between T-helper cells and a response that is able to clear an HPV lesion, CD4+ T-cell responders were also less likely to have progressive disease. The study also showed that 80% of cases have a CD8+ T-cell response to HPV16 (E6) (Steele et al., 2005). The interplay between the CD4 subtypes and the cytokines released plays an important role in controlling the tumour, Th1 responses (IL-2 and IFNγ, TNFα) being expected from an intracellular pathogen like HPV leading to cancer suppression, whereas a Th2 (IL-4 and IL-10) response would lead to tumour progression and suppression of the immune response. The presence of Treg cells producing IL-10, TGFβ1 and TGFβ2 are associated with a more severe cancer (Peghini et al., 2012b). As the disease severity increases from low grade cervical intraepithelial lesions (LSILs), to high grade cervical intraepithelial lesions (HSILs) and finally invasive cancer (CA), a shift in cytokine profile occurs with a reducing Th1 (IL-2, IFNγ and TNFα) response, conversely the Th2 and Treg response profiles increase with disease severity (Peghini et al., 2012a). This supports the Th1 cytokine profile inducing a tumour suppressive state, in contrast to the Th2 cytokine profile that promotes tumour progression and immune suppression.

1.6.7.2 HPV16 HNSCC Immune evasion tactics

A feature of HPV is the chronic infection caused by ineffective clearance of the virus due to an inadequate immune response. HPV is a non-lytic virus causing little cellular damage that would initiate an inflammatory response. Finally, the expression of oncogenes at various stages of the viral lifecycle play a role in the immune evasion, with the late production of the immunogenic capsid proteins (Stanley, 2010, Chow et al., 2010). The ability of the HPV16 proteins E6 and E7 to bind and inhibit key elements of the interferon signalling pathway lead to an impeded interferon antiviral response at several levels. The E6 protein subverts apoptotic signalling normally caused by the action of TNF-α, it also interacts with FAS-associated protein with death domain (FADD) and caspase 8, blocking FAS and TNF-related apoptosis-inducing ligand (TRAIL) signals ultimately leading to interruption of apoptosis and cell survival (Garnett and Duerksen-Hughes, 2006). The E5 protein is linked to down-regulation of the HLA class I molecule and reduced peptide presentation of the viral peptides (Ashrafi et al., 2005). The E6 and E7 proteins are key elements both in the tumour progression but also in the process of subverting the immune response.

1.6.8 Immunotherapy

Immune cell infiltration predicts for a better prognosis in many cancers, this is linked to both the number and functional state of the immune cells. This leads to the concept of using the immune system as a therapeutic option in cancer treatments. Immunotherapy looks to modulate the immune system in order to improve the effectiveness of tumour infiltrating lymphocytes.

1.6.8.1 Cancer vaccines

Vaccines that target TAAs or CTAs are a way of improving an immune response to antigens in the tumour using a systemic vaccine approach. Many clinical trials have been carried out with varying success (Melero et al., 2014). In lung cancer a MUC1 peptide vaccine yielded an improvement of 3 months (vaccine only) and 10 months when concurrent rather than sequential chemotherapy was administered; GP100 (peptide vaccine) has been used in Melanoma with an increased survival of 6.5 months (Butts et al., 2014, Schwartzentruber et al., 2011). On-going Trials using vaccines are focusing on unique antigens present in the tumour (neoantigens) that are delivered using an RNA vector loaded with multiple predicted peptides (Brown et al., 2014, Sahin et al., 2017, Ott et al., 2017). HPV E6 and E7 antigens that are present in head and neck cancer and Cervical cancer are used as targets in an on-going trial using an RNA vaccine (HARE-40 - A trial of a vaccine for cancers that tested positive to HPV, EudraCT reference number:2014-002061-30). Vaccination regimens are also employed using DNA fusion vaccines with immunogenic tetanus toxoid domains (pDOM) linked to carcinoembryonic antigen (CAP-1); this generated effective CAP1 CD8 T-cells responses.

In addition to this, gastrointestinal adverse events were identified and linked to a vaccine response in 48% of patients(McCann et al., 2016).

1.6.8.1 Adoptive cell transfer

Another method of eliciting an immune response is by introducing cells that are derived from the host immune system, this is done using adoptive cell therapy (ACT) where immune cells with anti-cancer activity are introduced back to the patient (Rosenberg and Restifo, 2015, Restifo et al., 2012). Tumour infiltrating lymphocytes are isolated from patients and then expanded in culture with IL2, followed by reintroduction of the autologous immune cells back into the patient, in a Melanoma study a response rate of 40% was observed (Besser et al., 2013). An alternative to this is reintroducing a specific cell population by co-culture with APCs from the tumour that contain tumour antigens. Tumour reactive CD8 identified by 4-1BB (TCR engagement) are further expanded in-vitro prior to reintroduction to the patient (Seliktar-Ofir et al., 2017). This has been used in a phase 2 clinical trial (Immunotherapy using 4-1BB Selected Tumour Infiltrating Lymphocytes for Patients with Metastatic Melanoma-NCT02111863), where tumour-antigen specific expanded 4-1BB positive cells are given to the patients.

The use of dendritic cell (DC) vaccines focuses on cultured DC cells (expanded from CD14 cells), they are loaded with immuno-reactive tumour antigens, further manipulation ensures the correct homing and co-stimulatory molecules are expressed, ultimately leading to tumour antigen presentation to CD4+/ CD8+ cells when reintroduced to the patient. The safety of this has been shown in a prostate cancer phase I clinical trial (Garg et al., 2017). Also, in patients with metastatic melanoma treated with DCs loaded with TAAs gp100 and/or tyrosinase, effective CD8+ T cell responses have also been observed and the patients show an improved progression-free survival (Schreibelt et al., 2016). Many of these studies assume a stable immune signature and tumour transcriptome across the time course of the clinical trial. However, in reality, natural change in the tumour may make the outcomes more difficult to interpret. An example is a recent trial to assess the efficacy of an inhibitor of PI3Kδ in HNSCC - (AMG319 study in HPV positive and negative HNSCC: NCT02540928). This isoform of PI3Kδ is expressed predominantly on leukocytes, the inhibition of PI3Kδ leads to a selective inactivation of Treg cells. This in turn releases cytotoxic CD8 cells and induces tumour regression (breast and melanoma mouse tumour model) (Ali et al., 2014). This is a so-called “window of opportunity” study, where the effects of the drug are evaluated by administering it between the diagnostic biopsy and surgical resection (approx. 30 days) prior to conventional treatment. It is important to know with confidence that the nature

and distribution of the immune infiltrates in tumours are stable over the course of a clinical trial and are well represented in the information obtained from small biopsies.

1.6.8.2 Immune checkpoint blockade

The use of immune cell checkpoint blockade has transformed the field of immunotherapy with dramatic results and improved patient survival; two of the most studied checkpoint inhibitors are the anti-CTLA4 (Ipilimumab) and anti-PD1 (nivolumab/pembrolizumab) therapies. Both of these checkpoint inhibitors target the self-regulating mechanism of the immune system that ordinarily dampen the immune response post infection, these antibodies block this function abrogating the negative, suppressive effects.

CTLA4 expressed on CD4+, Tregs and CD8+ cells competes for CD28 on APCs and has an increased binding affinity, thus affecting the co-stimulatory action of CD28. Anti-CTLA4 treatment in a phase-III trial was able to improve survival (10 months) in metastatic melanoma compared to gp100 peptide vaccine alone and in combination (gp100+ anti-CTLA4) (Hodi et al., 2010). The activity of anti-CTLA4 is able to increase the number of tumour specific CD8 cells in melanoma compared to pre-treatment, broadening the T cell repertoire via potentially enhanced T cell priming, in contrast pre-existing immune reactivities remained unaltered (inc. virus specific control responses)(Kvistborg et al., 2014).

Anti-PD1 therapy utilises the blockade of PD1 interactions with PDL1 and PDL2, which when engaged, inhibit T cell responses. The survival benefit of Pembrolizumab versus Ipilimumab in advanced Melanoma was dramatic, with much less toxicity (fewer adverse events) (Robert et al., 2015). Treatments of NSCLC with anti-PD1 also exhibited a radical reduction in tumour growth and improved survival, this was correlated to the expression of tumour PDL1, with low PDL1 expressing tumours having an inferior response to the therapy (Robert et al., 2015). The action of anti-PD1 releases pre-existing immune infiltrates that have been negatively regulated due to the PD1-PDL1 interactions, the subsequent released CD8 response is also more clonally restricted (Tumeh et al., 2014, Rizvi et al., 2015).

The highly effective nature of checkpoint inhibitors in mono-therapy has led to combination therapies with multiple checkpoint inhibitors (CTLA4 and PD1), as well as combinations of vaccines and checkpoint blockade, which can improve responses in personalised neoantigen vaccines when combined with anti-PD1 (Sahin et al., 2017, Wolchok et al., 2013).

As the nature of the immune infiltrates in tumours is better understood, additional targets are identified and explored, **Table 6** shows some of the on-going immunotherapies in HNSCC and NSCLC. This is not and exhaustive list of clinical trials, it does however demonstrate that many immune checkpoints are being targeted with overlaps between the two cancer types (Mei et al., 2020, Ghanizada et al., 2019, Herbst et al., 2018, Chen et al., 2020).

One of the main issues with immune checkpoint blockades has been the poor response rates, with 20-30% displaying little or no benefit (Sharma and Allison, 2015). It argues that selecting patients based on information about the tumour could be critical in improving responses. This requires stratification of patients into groups that are more likely to respond, selecting patients with high tumour immune infiltrates that are $PD1^{Hi}$ prior to treatment, will likely yield better responses. Conversely, patients with sparse immune infiltrates that are $PD1^{Lo}$ may benefit from alternative approach that are not immune focused.

HNSCC

Drug	Mechanism of action	Clinical trial	Phase
Nivolumab, nivolumab+ipilimumab or BMS-986016 (anti-LAG-3 antibody) (neoadjuvant and metastatic cohort)	PD-1, CTLA-4 LAG-3	NCT02488759	I/II
LAG525, PDR001	LAG-3, PD-1	NCT02460224	I/II
Varlilumab, nivolumab	CD27 PD-1	NCT02335918	I/II
PDR001	PD-1	NCT02404441	I/II
MEDI6383, durvalumab	OX-40 PD-L1	NCT02221960a	I
Atezolizumab	PD-1	NCT01375842	I
Pembrolizumab	PD-1	NCT01848834a	Ib
Pembrolizumab	PD-1	NCT02255097a	II
Nivolumab vs investigators choice	PD-1	NCT02105636a	III
Durvalumab	PD-L1	NCT01693562	I/II

NSCLC

Drug	Mechanism of action	Clinical trial	Phase
REGN4659	Anti-CTLA-4	NCT03580694	I
Cemiplimab	Anti-PD-1	NCT02383212	I
Tislelizumab	Anti-PD-1	NCT03432598	II
NC318	Siglec-15 antibody	NCT03665285	I
TSR-022,TSR 042, TSR033	Anti-TIM-3, Anti-PD-1, and anti-LGA3	NCT02817633	I
Eftilagimod alpha	LAG-3 fusion protein	NCT03625323	II
NKTR-214	CD122-biased agonist	NCT02983045	I/II
NKTR-214	CD122-Biased Cytokine	NCT03138889	I
APX005M	CD40 agonistic antibody	NCT03123783	I/II
SEA-CD40	CD40 Antibody	NCT02376699	I
ALX148	CD47 blocker	NCT03013218	I
IO102	IDO vaccine	NCT03562871	I/II
Epacadostat	IDO1 inhibitor	NCT02178722	I/II

Table 6. A Summary of on-going Immunotherapies in HNSCC and NSCLC.

This shows a non-exhaustive list of immune related therapies in HNSCC and NSCLC. It demonstrates that a wide range overlapping immune checkpoints are targeted in both cancer types. Adapted from (Mei et al., 2020, Ghanizada et al., 2019, Herbst et al., 2018, Chen et al., 2020).

1.7 Study summary, hypotheses and objectives

1.7.1 Study summary

The rapid expansion of immunotherapy being used in the treatment of cancer, has led to clinical trials using novel approaches with the aim of modulating the immune response. On-going clinical trials in HNSCC plan to use high resolution transcriptomics (RNA-Seq), to try and understand the impact of immunomodulatory agents on the immune cells. The AMG319 trial is a "window of opportunity" trial, where the pre-treatment diagnostic biopsy will be compared to surgical resection samples after treatment. This poses the problem of differences that exist due to different spatial and temporal sampling. Heterogeneity is a known feature of tumours at a genomic level but has not been assessed at the transcriptomic level.

Tumour infiltrating lymphocytes (TIL) in cancer play an important role in patient survival, with higher numbers linked to improved prognosis. HNSCC arises from conventional triggers (smoking, alcohol etc.) as well as a large number being caused by oncogenic HPV and the oncoproteins E6 and E7. The difference in survival between patients with HPV(+) and HPV(-) tumours is significant, even more so when the level of immune infiltration is taken in to account, where HPV(+) TIL rich tumours have the best prognosis even when compared to HPV(-) TIL rich tumours. This presents an interesting comparison group to study the potential differences in immune attack between HPV(+) and HPV(-) tumours. Where the presence of the immunogenic HPV proteins may elicit a different immune response leading to improved survival. Isolation and transcriptomic investigation of immune cells from solid tumours (primarily HSNCC and NSCLC) aims to profile CD8+ T-cells and B cells in greater detail. The comparison of purified immune cell subsets in different tissues and disease settings (TIL density) will allow unique features of these cells to be uncovered.

1.7.2 Study hypotheses

1.7.2.1 Hypothesis 1

Hypothesis 1a) the general tumour transcriptomic and immune signature is stably distributed across the tumour; **Hypothesis 1b)** the overall tumour transcriptomic and immune signature also remains stable over time.

1.7.2.2 Hypothesis 2

HNSCC tumours with rich immune infiltrates ($TIL^{high/mod}$) have different outcomes depending on whether they are HPV(+) or (-). **Hypothesis 2a)** is that there is a difference driven by the presence

of the virus between HPV(+) and (-) tumours regarding the functional phenotypes of the infiltrating T cells: CD8+ and CD4+. This assumes that in HPV(-) tumours, the T cells are somehow less functionally active or are suppressed. **Hypothesis 2b)** is that the difference in outcome is due to purely quantitative differences in the numbers of TILS between HPV(+) and HPV(-) tumours. This assumes that the abundance of TAAs are the driving force in the anti-tumour responses, where HPV(+) tumours have more CTLs due to ubiquitous expression of E6 and E7.

1.7.2.3 Hypothesis 3

Hypothesis 3a) Immune infiltrates of CD8+ T and B cells in tumours (head, neck and lung cancers) display a distinct profile relative to a control tissue, but remain consistent between the tumour types. **Hypothesis 3b)** The functional characteristics of the immune cells (CD8+ T and B cells) in tumours with varying levels of infiltration (TIL^{Hi} to TIL^{Lo}) are different and consist of different cellular subsets, that lead to differences in survival.

1.7.3 Objectives

Hypothesis 1a) will be tested by taking multiple, replicate biopsies from different locations within the tumour and comparing the transcriptomic and immunological signature within the replicates; **hypothesis 1b)** will be tested by taking serial biopsies over time between the diagnostic biopsy and surgical resection and comparing the transcriptomic signature and immunological signature within the replicates.

Hypotheses 2a and 2b) will be tested through different approaches: first numbers, types and gross phenotypic features will be determined by use of IHC. Functional indicators such as production of key cytokines, cytotoxic molecules and indicators of TCR engagement will be examined. Secondly, differential gene expression will be determined in whole tumour transcriptomes from HPV(+) and HPV(-) tumours. The transcripts for immune cell markers may need to be used to normalise the data and account for differences in immune cell number between $HPV(+)^{high/mod}$ and $HPV(-)^{high/mod}$ tumours. However, the analysis of whole tumours and the encompassed cells is a crude form of analysis. Therefore, it may be necessary to move to more sensitive transcriptomic analysis of isolated specific immune cell populations (CD8+ cells) to further interrogate phenotypic differences between HPV(+) and (-) tumours.

Hypothesis 3a) will involve the isolation and evaluation of specific immune cells (CD8+ T and B cells) from suspensions derived from tumours (HNSCC and NSCLC), control tissue and peripheral blood. Transcriptomic gene expression profiles (RNA-Seq) and cellular phenotypes (flow cytometry) will be evaluated between tissue and tumours types. **Hypothesis 3b)** will be tested by comparing the transcriptomic profiles of CD8+ T and B cells in TIL^{Hi} and ^{Lo} tumours, survival

differences will be evaluated using a retrospective cohort of patients in a tissue microarray by immunohistochemistry.

Chapter 2: Materials and Methods

2.1 Outline of collaborative work and data generation

An outline of the individuals that have supported the completion of this thesis are outlined in Table 7, the appropriate section is also noted in the table.

Support and guidance - Highlights key individuals that assisted with scientific guidance during the analysis and writing of the thesis.

Clinical samples and data - A large number of clinical samples were obtained and analysed throughout this thesis. None of this would be possible without access to the tumour specimens enabled by the ethical approval and input from the clinical staff. Dr Emma King was instrumental in allowing access to head and neck tumours, also enabling me to collate the clinical data for these cases. The Lung tumour collections were made possible by Dr Serena Chee and the surgeons Mr A. Alzetani and Mr E. Woo. The clinical trial assistant team facilitated collection of the lung tumours, including access to anonymised patient clinical data.

The two tissue microarray cohorts (TMA) in head and neck cancer and lung cancer, were prepared by the Research histology team based in the Histopathology department, Southampton Hospital. The head and neck meta data was originally curated by Dr Matthew Ward, with the lung cancer data originally prepared by Dr Serena Chee.

Experimental assistance - The immunohistochemistry (IHC) for markers of interest were carried out using the service from Research Histology, Histopathology department, Southampton Hospita (Monette Lopez and Maria Machado). Dr Sonya James carried out training on the Leica Cryostat, enabling the development of a process for obtaining RNA from frozen tumour tissue.

Training in the use of flow cytometry was carried out by Dr Angelica Cazaly (BD FACS Canto II). Initial training on cell sorting using the BD FACS Aria I and II was carried out by Dr Carolann McGuire and Richard Jewell. This was then built on to develop an immune cell isolation strategy from solid tumours. The RNA-Seq of whole tumour, bulk immune cells and single cells was carried out at the sequencing core facility, La Jolla Institute for Allergy & Immunology (Vijay Pandurangan).

Bioinformatics support – The RNA-Seq mapping and feature counting were carried out by collaborators Dr Jeongmin Woo (Whole tumour), Bharat Panwar (bulk CD8) and Crio Suastegui (single cell). Dr Jeongmin Woo also performed the differential gene testing on the whole tumour RNA-Seq using EdgeR. Guidance and assistance with custom R scripts was also given by Dr James Clarke and Dr Bharat Panwar.

Name	Category	Details	Section of thesis
Prof. Christian Ottensmeier	Support and guidance	Scientific and experimental	All sections
Prof. Peter Friedmann			
Dr Emma King			
Prof. Gareth Thomas			
Dr Vijayanand Pandurangan			
Dr Tilman Sanchez-Elsner			
Dr Katy McCann			
Lindsey Chudley			
Dr Angelica Cazaly			
Dr Serena Chee			
Dr James Clarke			
Dr Emma king	Clinical samples and data	Access to HNSCC surgical resection samples from Poole	3.2, 3.3, 4.2
Prof. Gareth Thomas		Access to HNSCC surgical resection samples from Southampton	3.2.3, 3.3.3, 4.2.3
Monnette Lopez and Maria Machado		Research histology - Access to pathology blocks, routine histology marker staining and formation of tissue microarrays (TMA)	3.2.3.7, 3.3.3.9, 4.2.3.1, 4.2.3.4, 4.3.4.5
Mr Aiman Alzetani, Mr Edwin		Access to surgical resection samples in NSCLC	4.3
Ben Johnson		Clinical support for NSCLC collection	4.2, 4.3
Dr Matthew Ward		Clinical database for retrospective HNSCC cohort	4.2.3.4
Dr Serena Chee		Clinical database for retrospective NSCLC cohort	4.2.3.4
Monnette Lopez and Maria Machado	Experimental	Research histology - routine histology marker staining and formation of tissue microarrays (TMA)	3.2.3.7, 3.3.3.9, 4.2.3.1, 4.2.3.4, 4.3.4.5
Dr Sonya James		Training on the cryostat, Histology and microscope support	3.2, 3.3, 4.3
Richard Jewell and Carolann McGuire		Training on the FACS ARIA for cell sorting	3.2, 3.3, 4.3
Dr Vijayand Parandurgan and Dr Greg Seumois		La Jolla Sequencing core facility - Whole tumour RNA-Seq	3.2.3, 3.3.3
Dr Vijayand Parandurgan and Dr Greg Seumois		La Jolla Sequencing core facility - Microscaled RNA-Seq	4.1.3, 4.2.3, 4.3.3
Dr Vijayand Parandurgan and Dr Greg Seumois		La Jolla Sequencing core facility - 10x single cell RNA-Seq in La Jolla	4.3.4
Dr James Clarke		Cell sorting for B cells for Single cell transcriptomics (10x genomics) in La Jolla	4.3.4
Christopher Woelk	Bioinformatics	Bioinformatics support - alignment, mapping and data normalisation of RNA-Seq data	3.2.3.1, 3.3.3.1, 3.3.3.2
Dr Jeongmin Woo		Bioinformatics support - alignment, mapping and data normalisation of RNA-Seq data	3.2.3.1, 3.3.3.1, 3.3.3.2
Dr Jeongmin Woo		Differential gene analysis - EdgeR	3.3.3.2
Dr Bharat Panwar and Dr Greg Seumois		Alignment, mapping and data normalisation of CD8+ T cell RNA-Seq data.	4.2.3
Ciro Ramirez Suastegui		Alignment, mapping and data normalisation of Single cell (B cell) RNA-Seq data	4.3.3
Dr Bharat Panwar		Bioinformatics support - help with DESEQ2 r scripts	4.2.3.3
Dr James Clarke		Single cell analysis support - r script help	4.3.3

Table 7. Acknowledgments

Outline of individuals that supported the completions of this thesis. It is divided into general guidance, clinical samples and data, experimental work and bioinformatics.

2.2 Study Subjects and sample collection

2.2.1 Ethics and approval

Patients diagnosed with HNSCC or NSCLC were consented and recruited by clinical staff using a standardised consent form. This allowed collection of material surplus to diagnostic requirements to be used for research purposes following surgical resections. Ethical approval for the collection and use of HNSCC and NSCLC samples for research was obtained from the NRES Committee South Central for HNSCC REC reference number: 09/H0501/90 and for NSCLC REC Reference Number: 14/SC/0186. Patient samples were anonymised in line with GDPR (general data protection regulation) and stored in line with the Human Tissue Act at the University of Southampton (Human Tissue Act License: 12009 Southampton and South West Hampshire Research Ethics Committee: 280/99).

2.2.2 Tissue collection strategy

The collection of tumours samples for both whole tumour RNA-Seq and also immune cell isolation by FACS was carried out concurrently, the flow diagram in Figure 6 highlights the standardised tissue collection strategy that was employed for collection of HNSCC and NSCLC tumours. The collection of tumours followed a prioritisation of always obtaining snap frozen and cryopreserved material, followed by immune cell isolation using FACS and then cryopreservation of cell suspensions if any surplus remained.

2.2.3 Summary of sample cohorts and data

Several different patient cohorts were used to generate the data presented within this thesis, Table 8 provides a summary, full details of each cohort are outlined in Appendix A.1

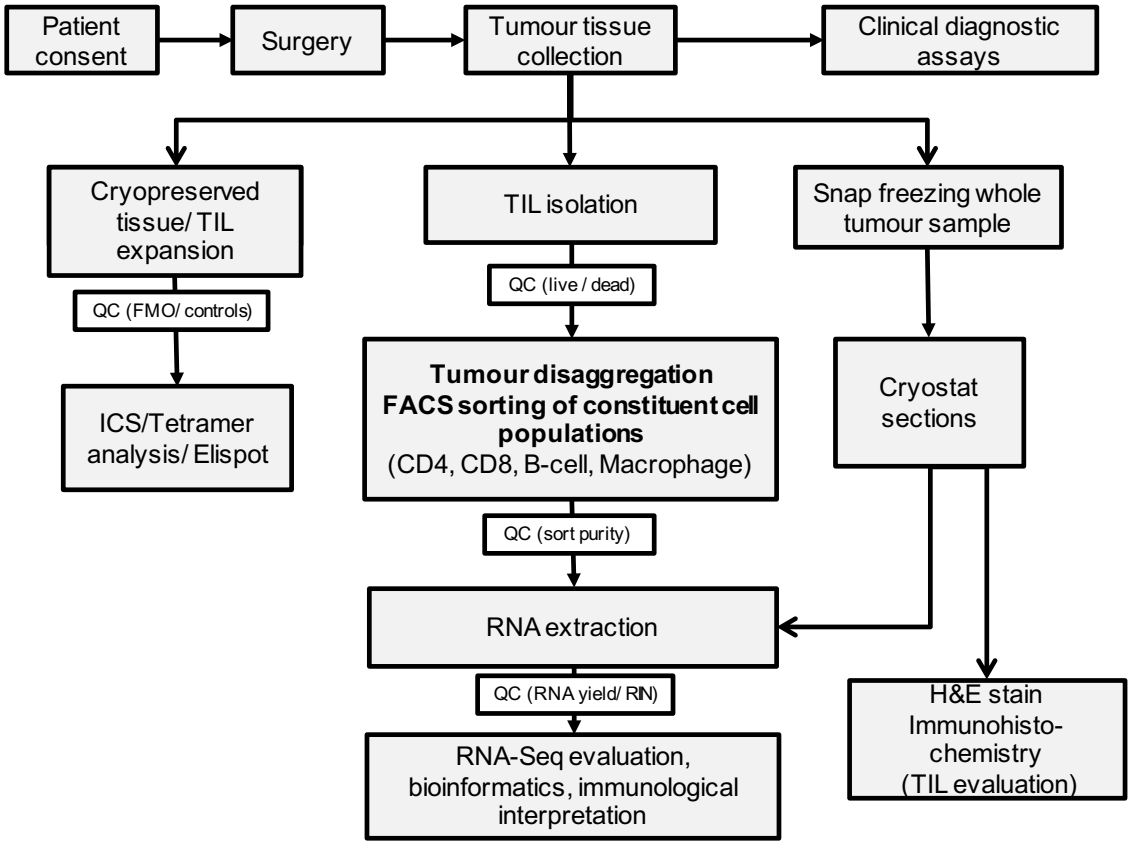


Figure 6. Standardized collection of human tumour tissue.

The figure demonstrates the sample collection and processing prioritisations for all patient tumours.

Patient cohort	tissue type	no.	data type	Thesis section
Tumour replicate dataset TRTS	HNSCC	44	Whole tumour RNA-Seq	3.2.3
Tumour replicate	HNSCC	41	CD8 IHC	3.2.3.7
Immuno-oncology panel - Biomark	HNSCC	12	IFC based qPCR	3.3.3
HPV(+) compared to (-) dataset HPVN	HNSCC	39	Whole tumour RNA-Seq	3.4.3
HNSCC CGA data	HNSCC	518	RNA-Seq	3.4.3.7
NSCLC CGA data	NSCLC	1004	RNA-Seq	3.4.3.7
HPV(+) compared to (-) dataset HPVN	HNSCC	16	qPCR on Purified B cells	3.4.3.8
Tissue processing method comparison	NSCLC and Melanoma	6	FACS and Micro-scaled RNA-Seq (CD8+ T cells)	4.1.3
Purified CD8 T cell RNA-Seq from tumours	HNSCC, NSCLC and non-involved lung	55, 48 and	FACS and Micro-scaled RNA-Seq	4.3.3
TRM flow cytometry	NSCLC	16	FACS surface and inta-cellular staining	4.3.3
Comparison of CD8+PD1 Hi and CD8+PD1 Lo T cells	NSCLC and Melanoma	7	Micro-scaled RNA-Seq (CD8+ T cell)	4.3.3
NSCLC retrospective survival cohort	NSCLC	689	CD8 and CD103 IHC	4.3.3
HNSCC retrospective survival cohort	HNSCC	123	CD103 IHC	4.3.3
Single cell B lymphocyte dataset	NSCLC	15	Single cell RNA-Seq (10x) and IHC	4.3.4

Table 8. Summary of patient cohorts and data.
A summary of the different patient cohorts and the analysis that has been undertaken.

2.3 Isolation of RNA from whole tumour samples

All snap frozen tumour samples were stored at -80°C. Snap frozen tumour samples were sectioned using a cryostat (Leica) with blade and sample block set at -20°C. Tumour tissue was kept on dry ice and put immediately into the cryostat before partial embedding in OCT (optimum cutting temperature) compound (Leica), cryospray (Bright instruments) was used at this stage to ensure the tissue remained frozen. Cryosections n=15-30 (10µm) were cut and used for RNA isolation with the RNeasy Mini Kit (Qiagen Ltd., Manchester, UK). This procedure is currently being put into an SOP for wider use.

2.4 Whole tumour RNA-Seq analysis of tumour replicates across time and space (Data set TRTS)

Multi-region tumour samples (replicates) were collected under general anaesthesia and were snap frozen immediately at diagnostic biopsy and surgical resection. Pre-resection replicates were

taken before vascular ligature to minimize hypoxic time. Spatial heterogeneity was assessed by collection of multiple replicates at either diagnostic sampling or at resection. Samples were taken at least 10mm apart to maximize the chance of capturing spatial heterogeneity. The comparison between replicates collected at diagnostic sampling to those replicates from the same patient at resection were further used to assess change over time (temporal heterogeneity), in the recognition of the fact that this comparison may also be shaped by any spatial heterogeneity in any given cancer. A total of 16 patients were consented, with 44 samples collected, the full analysis included 14 patients and n=37 tumour replicates. Single timepoint replicates (n=15) were collected from 6 patients, 8 patients had replicates (n=22) from two timepoints, both diagnosis and resection, with an average of 26 days between. Patient demographics, clinical details and number of replicates are shown in Appendix A.1.1.

2.4.1 Histology and immunohistochemistry for RNA-Seq dataset TRTS (Analysis of tumour replicates across time and space)

Tumour grade and differentiation were recorded from formalin fixed, paraffin embedded (FFPE) tissue by an accredited pathologist [G.J.T] using the diagnostic H and E sections. Immunohistochemical (IHC) staining for CD8a (anti-CD8a antibody (clone: C8/144B, DAKO) was performed on each frozen tumour tissue replicate, for each case, using the region immediately adjacent to that used for RNA-Seq analysis. This enabled spatially distinct samples to be compared within the same patient, this did however limit the number of analytes that we could evaluate. CD8 was chosen as the key metric due to its known prognostic importance in tumour biology (see section 1.6.1). IHC was undertaken by the research pathology team, Southampton Hospital (Monette Lopez and Maria Machado). An insufficient amount of frozen material remained for cases #6, #11 and #13. Tumour-infiltrating CD8 T-cells were quantified using a Zeiss AxioCam MRc5 microscope (Zeiss, Cambridge, UK) and Zeiss Axiovision software (version 4.8.1.0; Zeiss) with an average of 10 high-power (x400) fields; an average intratumoural score per high-power field (HPF) was calculated (Appendix A1.1).

2.4.1.1 Comparison of IHC to RNA-Seq data for Immune gene markers

Spearman correlation coefficients were calculated between CD8A gene expression and CD8 immune cell IHC count using Graphpad Prism (v8.0).

2.4.1 RNA-Seq methodology for Dataset TRTS (Analysis of tumour replicates across time and space)

RNA quality was assessed using the Agilent 2100 Bioanalyser (Agilent Technologies UK Ltd., Stockport, UK); an average RNA quality number (RIN) of 8.51±0.90 was observed across all tumour samples. Total RNA was converted into a library for sequencing on the HiSeq 2000 (Illumina Inc., San Diego, USA) using the TruSeq™ stranded mRNA Sample Preparation Kit (Illumina Inc.). Briefly, poly-A mRNA was purified from total RNA (100ng) using the Poly(A) Purist Mag Kit (Life Technologies Ltd., Paisley, UK), according to the manufacturer's instructions. The mRNA was then amplified and converted into cDNA, which was purified and used to construct libraries that were hybridized to the flow cell for single end (SE 35bp) sequencing. The RNA sequencing was carried out in collaboration with the La Jolla Institute for Allergy & Immunology using the sequencing core facility.

2.4.2 Bioinformatics and data processing for TRTS dataset (Analysis of tumour replicates across time and space)

2.4.2.1 RNA-Seq Dataset TRTS analysis pipeline

The quality of raw SE read data in FASTQ files was assessed and reads of low quality were trimmed or removed. SE reads were then mapped to the human genome (hg19) using TopHat (version 2.0.9) (Trapnell et al., 2009) and, following the removal of multi-mapping reads, converted to gene-specific read counts for 20,825 annotated genes using HTSeq-count (version 0.5.4) (Anders et al., 2014). Raw counts from the RNA-Seq were processed in Bioconductor package DESeq2, variance was estimated and size factor normalized. Bioinformatics was performed by Dr Jeongmin Woo at Southampton University (Bioinformatics core).

2.4.2.2 RNA-Seq data visualisation (heatmaps and PCA)

Qlucore Omics software (3.2) was used to visualize the normalized RNA-Seq data using Principle component analysis and hierarchical clustering. These were performed on data using the setting: mean=0, variance=1 normalization in the Qlucore Omics software (3.2). Where the clustering is agglomerative, average linkage criterion was used to determine the distance between sets of samples and hierarchical clustering is based on Euclidean distances, where the Euclidean distance will be equal to sqrt(1-r), where r is the Pearson correlation coefficient. In order to identify the 4000 most variable genes, variance was set at 0.2879, this value was then adjusted to yield a decreasing number of gene features. Briefly genes with a standard deviation (SD) less than a specified variance cut off from the maximal SD were removed using Qlucore Omics Explorer (3.2)

as previously described (Engel et al., 2016). RNA-Seq data were visualised in heat maps where each row represents normalized gene expression values for a given gene; each column represents the gene expression for a given tumour: red shading denotes greater gene expression; blue shading denotes lower gene expression.

2.4.2.3 Euclidean distance comparisons

As a measure of similarity and dissimilarity for intrapatient and interpatient sample replicates, Euclidean distances were quantified for the 4000 most variable genes in the R statistical environment (3.3) between all replicates from the same patient to all other replicates, where the Euclidean distance is equal to the sqrt(1-Pearson)^2. This measures distances between samples in the hierarchical tree, with smaller distances between replicates indicating a closer relationship. The median distances were calculated for all replicates to all other replicates (interpatient) and compared to the median distances from intrapatient replicates, e.g. case #14 median distance for intrapatient derived from n=6 and compared to median distance to all other cases n=31 (each replicate comparison is plotted). The Wilcoxon matched-pairs signed rank test was used to compare median distances between replicates from the same patient to those from different patients.

2.4.2.4 Immune gene heatmap and dot plots

A curated list of immune genes relating to immune cell markers, effector function and exhaustion/ regulatory function were derived from the gene ontology terms GO:0002250 adaptive immune response, GO:0002449 lymphocyte mediated immunity and GO:0002456 T cell mediated immunity were visualised in a heatmap. Reads per kilobase per million mapped reads (RPKM) were used to display expression of genes between replicates in dot plot comparisons.

2.4.2.5 Correlation analysis

Spearman correlation analysis was performed in Corrplot 0.73, in an R statistical environment (3.3) and visualised in in Graphpad Prism (v8.0). Correlation analysis was carried out on all genes and the top 4000 most variable genes. Median, minimum and maximum correlations were also calculated for intrapatient and interpatient replicates.

2.4.2.1 Data repository for TRTS - Tumour replicates

RNA-Seq data for all samples including clinical data has been deposited at ArrayExpress accession E-MTAB-4546.

2.5 Patient stratification using expression of immune genes.

2.5.1 Immune gene expression evaluation in HNSCC and NSCLC using RNA-Seq

The HNSCC HPVN cohort (Appendix A1.2) (n=39), HNSCC CGA (n=518) and NSCLC CGA (n=1004) RNA-Seq data were used to visualise a curated list of core immune genes, stromal and glycolysis markers.

2.5.1.1 RNA-Seq data

Our own HNSCC cohort HPVN RNA-Seq data (Wood et al., 2016a). The Cancer Genome Atlas (TCGA) Genome Data Analysis Centre (GDAC) Firehose website was used to access HNSCC RNA-Seq data (http://gdac.broadinstitute.org/runs/stddata_2015_11_01/data/HNSC/20151101/)(Cancer Genome Atlas, 2015), and the NSCLC RNA-Seq data (http://gdac.broadinstitute.org/runs/stddata_2016_01_28/data/LUSC/20160128 and _2016_01_28/data/LUAD/20160128)((Cancer Genome Atlas Research, 2014) (Campbell et al., 2016).

2.5.1.1 Data visualisation and heatmaps

A curated list of immune genes relating to immune cell markers, effector function and exhaustion/ regulatory function were derived from the gene ontology terms GO:0002250 adaptive immune response, GO:0002449 lymphocyte mediated immunity and GO:0002456 T cell mediated immunity. Additionally, the stromal genes ACTA2 (Smooth muscle actin), POSTN (Periostin) and COL1A1 (collagen, type 1, alpha 1), and glycolysis genes SLC2A1 (GLUT1), ALDOA (Aldolase) and PKM2 (Pyruvate kinase M1/2) were included. Qlucore Omics software (3.2) was used to visualize the normalized RNA-Seq data using hierarchical clustering as stated previously. Briefly, the clustering is agglomerative, average linkage criterion was used to determine the distance between sets of samples and hierarchical clustering is based on Euclidean distances (Pearson correlation coefficient). Red shading denotes greater gene expression; blue shading denotes lower gene expression. Immune infiltration as determined by H and E is annotated on the HPVN data, CGA data displays manual categorisation of immune infiltration based on gene expression.

2.5.2 Advanta Immuno-Oncology gene expression assay

The Immuno-Oncology gene expression assay 24.192 IFC (integrated fluidics chip) (PN 101-7088) from fluidigm was used to assess gene expression in n=9 (See Appendices B.5) HNSCC samples

that had previously been analysed with RNA-Seq. The samples selected had a range of immune gene expression from ”hot ” to “cold” (n=4 Immune^{Hi}, n=2 Immune^{Mod} and n=4 Immune^{Lo}) determined by the RNA-Seq. A positive control was used to confirm that all of the assay targets were working, this was also used as a reference point for the analysis. The Advanta Immuno-Oncology Panel A consists of 91 target genes, and panel B consists of 74 target genes and 5 reference genes (see table in Appendix B.5.1 for details).

2.5.2.1 Integrated fluidics chip (IFC) and RT-qPCR

Full details of the Immuno-Oncology gene expression assay protocol 24.192 IFC (PN 101-7088) are outlined in Appendices B.5.2. A summary of the reverse transcription (RT) reaction is shown in Table 9. Pre-amplification of gene targets were carried out using pooled assays from panel A and B, followed by a 1:5 dilution (Table 10).

RT pre-mix			
Component	**Vol (µL)**	**RT conditions**	
Reverse Transcription Master Mix (Fluidigm)	1	25°C	5min
PCR Water	3	42°C	30min
RNA (75ng)	1	85°C	5min
Total	5	4°C	-

Table 9. Reverse transcription of RNA to cDNA

Conditions used for conversion of RNA to cDNA for the immune-oncology panel.

Panel A and B assay mixes are diluted from 10x to 2x prior to running on the IFC (Table 11). Once the assays and samples are ready they are loaded into the IFC controller (Juno) see Figure 7 for assay loading map. Following this the IFC is run on the Biomark HD system using a GE 24x192 standard v1 thermal protocol, with the settings Application = Gene expression, passive reference = ROX, Assay = Single probe and Probes = FAM-MGB.

Preamplification reaction mix:	
Component	**Vol (µL)**
Preamp Master Mix	1
Advanta IO Gene Expression Assay Preamp Pool—Panel A	1.25
Advanta IO Gene Expression Assay Preamp Pool—Panel B	1.25
PCR Water	0.25
Total	3.75

Preamplification reaction:		**RT conditions**	
Preamplification pre-mix	3.75	95°C – 2min	
cDNA	1.25	95°C – 15sec	16 cycles
Total	5	60°C – 4min	
		4°C - hold	

Sample dilution mix	**Vol (µL)**
20x GE sample loading reagent	35
DNA Dilution Reagent	525
Total	560
1:5 dilution of Preamplification reaction	
Sample dilution mix	20
Preamplification reaction	5

Table 10. Preamplification of cDNA and subsequent dilution of the cDNA.

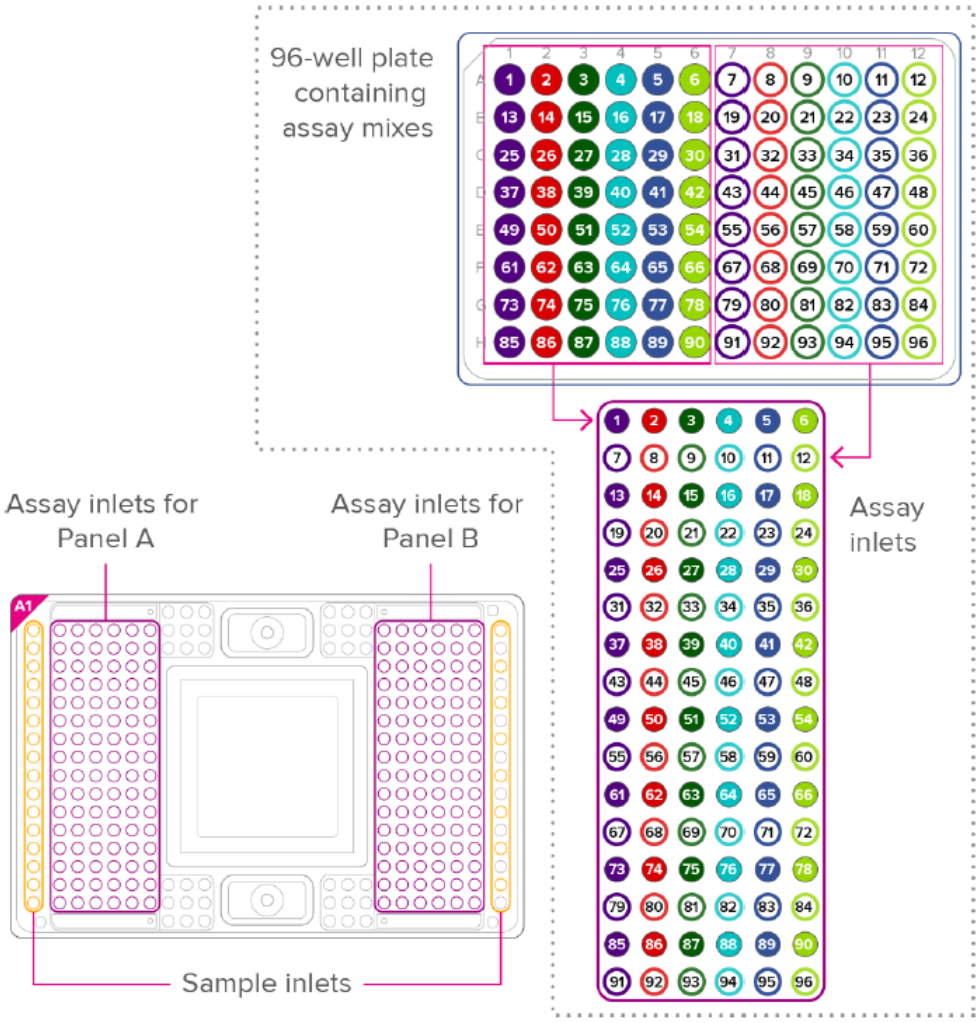
Details of the pre-amplification and cDNA dilution for the immune-oncology panel.

Assay dilution mix		
Component	**Vol (µL)**	
20x GE sample loading reagent	250	
DNA Dilution Reagent	1,750	
Total	2,000	
Assay dilution (10x to 2x)		
Component	**Vol (µL)**	
10x assay Plate A and B	2	
DNA Dilution Reagent	8	
Total	10	
Final assay mix		
Component	**Vol (µL)**	
2x Gene expression master mix	3	
2x assay mix (plate A and B)	3	
Total	10	

Table 11. Assay preparation for Panel A and B.

Details of the assay plate preparation for immune-oncology panels A and B.

(A) Assay mix loading map:



(B) Sample mix loading map:

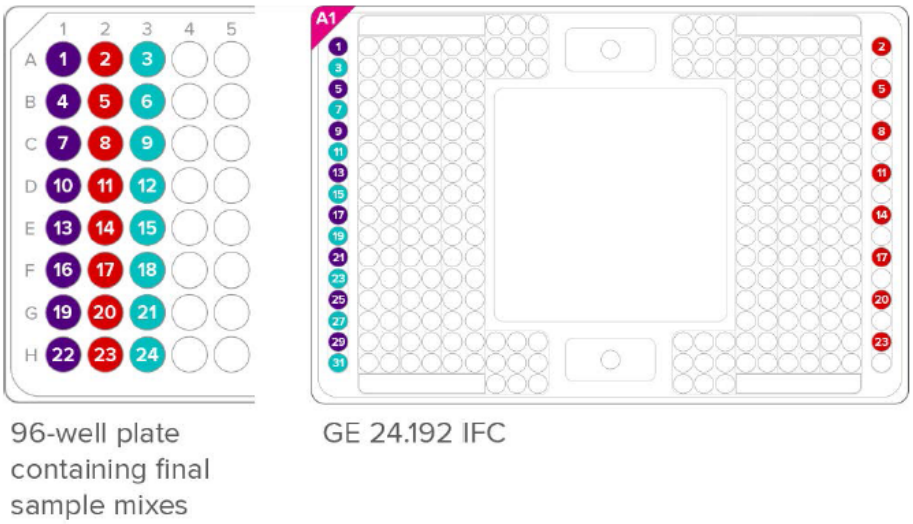


Figure 7. Loading the integrated fluidics chip (IFC).

(A) Displaying the assay loading map for panel A and B. (B) The sample loading map for the 24.192 IFC.

2.5.2.2 Data processing and calculation of relative gene expression.

The raw Ct (cycle time) values from the Biomark HD qPCR machine were processed in the Biomark real-time PCR analysis software (Fluidigm), Ct values for genes of interest (GOI) were normalized against the mean expression of the two reference genes (ACTB and B2M) and presented as ΔΔCt values relative to the positive control (pos.ctrl). The final relative expression values were generated by subtracting the ΔΔCt from the total number of cycles (40), this meant the test results were positively correlated, a format that facilitates interpretation for clinical decision making (Laible et al., 2016, Varga et al., 2017).

Relative expression: (Ct) $40 - \Delta\Delta Ct(GOI)_S =$

$40 - ((Ct[GOI]_{sample} - meanCt[REF]_{sample}) - (Ct[GOI]_{pos.ctrl} - meanCt[REF]_{pos.ctrl}))$

The relative expression values were then visualised as heatmaps in Qlucore Omics software (3.2) detailed in section 2.4.3.3. Red shading denotes greater gene expression; blue shading denotes lower gene expression. Comparison of RNA-Seq and qPCR expression for a selected number of targets was performed using Spearman's correlation coefficient (r) in Graphpad Prism (v8.0).

2.6 Whole tumour RNA-Seq Comparing HPV(+) and HPV(-) TIL rich tumours (Dataset HPVN)

39 consecutive HNSCC samples were obtained from patients at three centres (Southampton, n=22; Poole, n=15; Liverpool, n=2) from 2010-2014. Tumour samples were collected at least 10mm apart, following general anaesthesia but before surgical resection, and were snap frozen immediately. Appendix A.1.2 shows the patient demographics, tumour characteristics and tumour sampling/processing information for the HPV(+) and HPV(-) patient cohorts.

2.6.1 RNA-Seq methodology for HPVN dataset (HPV(+) vs HPV(-) TIL rich tumours)

The RNA-Seq for the HPVN dataset was single end (SE 35bp) sequencing as per TRTS data (Analysis of tumour replicates across time and space **-** Section 2.3.1).

2.6.2 Histology and immunohistochemistry for RNA-Seq dataset HPVN (HPV(+) vs HPV(-) TIL rich comparison).

Frozen tumour sections taken immediately adjacent to the tissue analysed by RNA-Seq were stained with hematoxylin and eosin (H&E) and viewed under low-power magnification (x2.5 objective) as described previously (Marsh et al., 2011); TIL^{Hi}: diffuse, present in >80% of tumour/stroma; TIL^{Mod}: patchy, present in 20–80% of tumour/stroma; TIL^{Lo}: weak/absent, present

in <20% of tumour/stroma by an accredited pathologist [G.J.T]. Data regarding the percentage tumour cells, tumour grade and pattern of invasion were also recorded. IHC was performed on FFPE tumour sections against CD3, CD4, CD8 and CD20 (all from Novocastra, Milton Keynes, UK). IHC was performed by Research histology, Southampton University Hospital. TILs were quantified using a Zeiss AxioCam MRc5 microscope (Zeiss, Cambridge, UK) and Zeiss Axiovision software (version 4.8.1.0; Zeiss) in an average of 10 high-power (x400) fields across representative areas of each tumour to allow for intratumoural heterogeneity; an average intratumoural TIL score per high-power field was calculated, CD8 is shown as an example in Figure 8. Additionally, IHC was performed against the antigenic targets, CD200 (Sigma-Aldrich Company Ltd., Gillingham, UK) and CD23 (Abcam, Cambridge, UK). HPV status was evaluated by IHC against p16 (CINtec, Roche, Burgess Hill, UK) and scored as HPV(+) (>50% tumour cells positive) or HPV(-) (<50% tumour cells positive). HPV status was also confirmed by evaluation of E6 and E7 RNA transcript levels from the RNA-Seq data Appendix A.1.2.

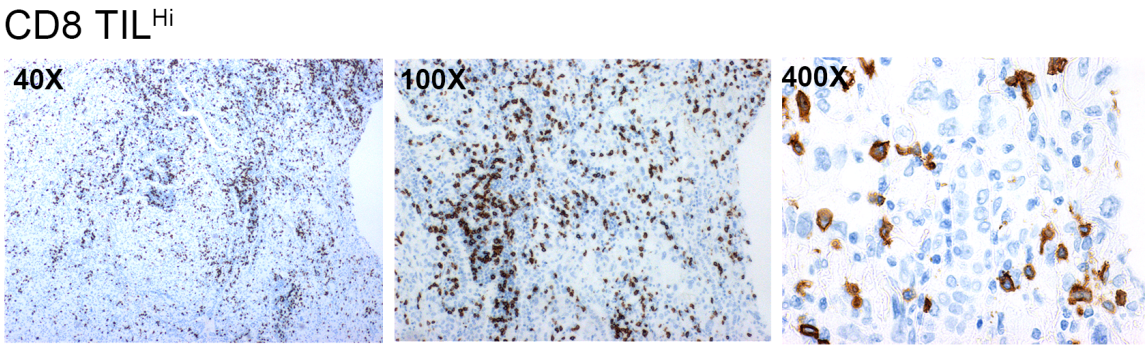


Figure 8. Representative images for CD8 TIL^{Hi} at 40x, 100x and 400x magnification.
For TIL scoring, 10 high power fields are counted at 400x. (IHC slides prepared by research histology – Monette Lopez and Maria Machado).

2.6.3 Bioinformatic analysis of HPVN dataset (HPV(+) vs HPV(-) TIL rich tumours)

2.6.3.1 RNA-Seq dataset HPVN - HPV(+) vs HPV(-) TIL rich tumour comparison

The quality of raw SE read data in FASTQ files was assessed and reads of low quality were trimmed or removed. SE reads were then mapped to the human genome (hg19) using TopHat (version 2.0.9) (Trapnell et al., 2009) and, following the removal of multi-mapping reads, converted to gene-specific read counts for 23,368 annotated genes using HTSeq-count (version 0.5.4) (Anders et al., 2014). Non-specific filtering of count data was performed using the Bioconductor package EdgeR (version 3.4.2) (Nikolayeva and Robinson, 2014, Robinson et al., 2010) such that genes with less than 2 read counts per million in 25% of tumour samples were excluded from further analysis. The remaining 14,528 genes were subject to variance stabilizing

transformation of trimmed mean of M-values (TMM) data normalization (Robinson and Oshlack, 2010) to account for differences in library size from sample to sample. Bioinformatics was performed by Dr Jeongmin Woo at Southampton University (Bioinformatics core).

2.6.3.2 Identification of differentially expressed genes (DEG)

Differentially expressed genes (DEGs) between HPV(+) and HPV(-) groups ($TIL^{mod/Hi}$) were identified with an FDR (false discovery rate) adjusted *p*-value <0.05 (i.e., *q*-value <0.05) and a fold change of >2 or <-2 (log2) using EdgeR (Nikolayeva and Robinson, 2014). Fold change was calculated in EdgeR as the log2 of geometric mean of intensities; a positive and a negative fold change represents genes that were expressed to a greater or lesser extent, respectively, in HPV(+) versus HPV(-) tumours. *q*-values were obtained from differential expression test in EdgeR using the generalized linear model likelihood ratio test and adjusted for multiple testing using the Benjamini and Hochberg method to control the FDR. This package models the negative binomial distribution and implements general linear models to identify DEGs. EdgeR was also used to identify DEGs while adjusting for covariates associated with varying proportions of lymphocyte subsets in each tumour sample as reflected in the expression of CD19 (B-cells) and CD4 and CD8A (T-cells) e.g. R-script used in EdgeR for the covariate adjustment was: design<model.matrix(~adjustv_CD19+ adjust_CD4+adjustv_CD8+Group). Differential testing using EdgeR was performed by Dr Jeongmin Woo at Southampton University (Bioinformatics core).

2.6.3.3 Data visualisation using Heatmaps

Qlucore Omics software (3.2) was used to visualize the normalized RNA-Seq data using Principle component analysis and hierarchical clustering. These were performed on data using the setting: mean=0, variance=1 normalization in the Qlucore Omics software (3.2). Where the clustering is agglomerative, average linkage criterion was used to determine the distance between sets of samples and hierarchical clustering is based on Euclidean distances (sqrt(1-r)^2), where r is the Pearson correlation coefficient. RNA-Seq data were visualised in heat maps where each row represents normalized gene expression values for a given gene; each column represents the gene expression for a given tumour: red shading denotes greater gene expression; blue shading denotes lower gene expression.

2.6.3.4 Gene ontology and pathway analysis

GO terms associated with biological processes and biological pathways that were significantly over-represented for DEGs (*q*-value <0.05) were identified with ConsensusPathDB (Kamburov et al., 2011) (CPDB, release 30) using the hypergeometric test. ConsensusPathDB GO and pathway analyses were performed for genes that were expressed (i) to a greater extent and (ii) to a lesser

extent in HPV(+) compared to HPV(-) TIL rich tumours. REVIGO was used to visualise non-redundant GO terms, it uses a clustering algorithm that relies on semantic similarity measure, enabling large overlapping gene ontology terms to be reduced in number (http://revigo.irb.hr/) (Supek et al., 2011).

2.6.3.5 Validation of findings in The Cancer Genome Atlas (TCGA) HNSCC data

HNSCC RNA-Seq data (TCGA HNSC HiSeqV2 2015-02-24) was obtained from The Cancer Genome Atlas (TCGA) Genome Data Analysis Centre (GDAC) Firehose website (http://gdac.broadinstitute.org/runs/stddata_2015_11_01/data/HNSC/20151101/), the RNA-Seq methodology and processing have been described by TCGA (Cancer Genome Atlas, 2015). As HPV-driven cancers typically arise in the oropharynx, tonsil and base of tongue, these anatomical sites were evaluated from TCGA. Unsupervised clustering and visualisation in heatmaps (as previously described) of 46 HPV16(+) and 26 HPV(-) anatomically matched tumours from the oropharynx, tonsil and base of tongue was performed using the differentially expressed gene lists generated from our own analysis.

2.6.3.6 Data repository for HPVN (HPV(+) vs HPV(-) TIL rich comparison)

RNA-Seq data has been deposited in the Gene Expression Omnibus (GEO) at the National Center for Biotechnology Information (NCBI) under accession number GSE72536.

2.6.1 RT-qPCR on B cells from HNSCC tumours

2.6.1.1 Real-time quantitative reverse transcription (RT) PCR

Tumour infiltrating B cells were isolated and sorted using the standard operating procedure outlined in Appendix A.2 (SOP no. 384 - Cell dissociation and immune cell sorting), samples details are shown in Appendix A.1.3. RNA was extracted from the B cells using the Direct-zol™ RNA MiniPrep system (ZYMO Research Co., Irvine, USA). Reverse transcription (RT) was performed on 1.5ng of RNA using SuperScript® III First-Strand Synthesis System (Invitrogen, Fisher Scientific UK Ltd.). Real-time quantitative reverse transcription PCR (RT-qPCR) assays were performed using Taqman® probes for golgi-associated, gamma adaptin ear containing, ARF binding protein 2 (*GGA2*), ADAM metallopeptidase domain 28 (*ADAM28*), *CD200*, Spi-B transcriptional factor (*SPIB*), stromal antigen 3 (*STAG3*), Vascular Cell Adhesion Molecule 1 (*VCAM1),* Inducible T-Cell Co-Stimulator Ligand *(ICOSLG)* and B-Cell CLL/Lymphoma 2 (*BCL2).* qPCR was performed for the selected genes using TaqMan Gene Expression Assays (Life Technologies Ltd.), according to the manufacturer's instructions (see Table 12 and Table 13): *GGA2* (Human Hs00370910_m1), *ADAM28* (Human Hs00248020_m1), *STAG3* (Human Hs00429370_m1), *CD200* (Human

Hs01033303_m1), *SPIB* (Human Hs00162150_m1), *ICOSLG* (Hs00323621), *BCL2* (Hs01048932_g1) and *VCAM1* (Hs01003372_m1). Analysis of RT-qPCR data was performed using the comparative Ct method (2^-(ΔΔCt) method) using an internal control (*Actin*) and defined as a normalized relative gene expression compared to the control gene (Livak and Schmittgen, 2001). RT-qPCR was reported in accordance with Minimum Information for Publication of Quantitative Real-Time PCR Experiments (MIQE) (Bustin et al., 2009).

Gene expression (RT-qPCR) of *STAG3* and *CD200* was validated in the original tumour RNA, HPV(+) n=8 and HPV(-) n=8 patient tumour samples. B-cells (CD20+/ CD19+) isolated from an independent cohort of HPV(+) tumours (n=6) using a BD FACSAria™ sorter (BD Biosciences, Oxford, UK) were assessed for gene expression of *ADAM28, BCL2, CD200, GGA2, ICOSLG, SPIB, STAG3,* and *VCAM1* (as described previously).

SuperScript® III First-Strand Synthesis System (Invitrogen			
Primer mix	1x (ul)	Condition	
RNA	1 (1.5ng)	65°C	5min
OligoDT	0.25	4°C	1min
Hexamers	0.25		
10mM dNTPs	0.5		
H_2O	3ul		

RT-mix	1x (ul)	Condition	
10x RT buffer	1	25°C	10min
35mM $MgCl_2$	2	50°C	50min
0.1M DTT	1	85°C	5min
RNase OUT	0.5	4°C	-
Superscript III RT polymerase	0.5		
Final reaction mix = 5ul Primer mix + 5ul RT-mix			

Table 12. cDNA synthesis from B cell RNA.
cDNA synthesis reaction conditions using SuperScript® III First-Strand Synthesis System Invitrogen.

TaqMan Gene Expression Assays (Life Technologies)				
qPCR mix	1x (ul)	Condition		
		Defined by ABI instrument 6000 HT		
Taqman reaction mix	5	50°C	2min	Hold
TaqMan Gene Expression Assays	0.5	95°C	10min	
H_2O	2.5	95°C	15sec	40 cycles
cDNA	2ul (1/5 dilution)	60°C	1min	

Table 13. qPCR reaction mix and conditions.
Reaction conditions for the TaqMan Gene Expression Assays, Life Technologies.

2.7 Isolation and analysis of purified immune cells form HNSCC and NSCLC

2.7.1 Tissue dissociation protocol

Immune cells from tumour and non-involved lung were isolated using a combination of mechanical and enzymatic dissociation, the full standard operating procedure is shown in appendix A.2 (SOP no. 384 - Cell dissociation and immune cell sorting). The tissue was cut into small fragments (1-2mm) using a scalpel. Fragments were then incubated at 37^{o}C for 15 minutes in an orbital shaker with 1-2mL RPMI 1640 medium (Gibco, Fisher Scientific UK Ltd., Loughborough, UK) containing 0.15 WU/mL Liberase DL (Roche Diagnostics Ltd., Burgess Hill, UK) and 800 units/mL DNase I (Sigma-Aldrich Co. Ltd., Gillingham, UK). The cell lysate was then passed through a 70µm filter with ice-cold RPMI 1640 medium and centrifuged at 385g for 7 minutes. Cells were re-suspended RBC lysis solution (Qiagen Ltd., Manchester, UK) at a 1:10 ratio (cell pellet volume to RBC lysis buffer) and incubated for 10 minutes at 4^{o}C, followed by centrifugation at 385g for 7 mins. The cells were resuspended in MACS buffer (1xPBS containing 2mM EDTA (0.5M stock Ambion) (pH 8.0) and 0.5% BSA (Molecular grade, Sigma) and counted using a Neubauer haemocytometer (Celeromics, Cambridge, UK), the volume was adjusted to give a concentration of 10e6 cells/mL.

2.7.2 Cryopreservation of tumour and tumour cell suspensions

Cryopreserved tumour and lung samples were stored as dispersed tumour (DST) and lung (DSL) and also as 2mm pieces of tissue (n=4-10 pieces), CT (tumour) and CL (non-involved lung). The DST/ DSL samples processed as above (section 2.6.1) were then cryopreserved in freezing media (50% complete RMPI (Fisherscientific), 40% FCS and 10% DMSO (both Sigma). Cryopreserved DST and DSL samples were thawed, washed twice with pre-warmed (37°C) and room temperature MACS buffer and prepared for staining as above. The CT and CL samples were cryopreserved in 90% foetal calf serum (Gibco) and 10% DMSO. Thawed CT/ CL samples were washed with pre-warmed complete RMPI followed by the tissue dispersion protocol (section 2.6.1) prior to antibody staining.

2.7.3 Surface antibody staining procedure for flow cytometry

Following tumour dissociation the cells were resuspend in 1xPBS and incubated with a fixable viability dye for 30 mins at 4°C (ThermoFisher Live/Dead™ Fixable Aqua; L34965). The cells were then washed and resuspended in MACS buffer and incubated with 20µL FcR block (Miltenyi Biotec

Ltd., Bisley, UK) per 100µL of cell suspension for 10 mins. The Immune cells were then stained with a cocktail of fluorescently conjugated antibodies (Table 14, Table 15 and Table 16) at 4^{o}C for 30 minutes. The unbound antibody was removed by washing the cell suspension with 2ml of MACS buffer and centrifugation at 385g for 7 minutes. The cells were then resuspended in MACS buffer at 10e6 cells/mL ready for analysis on a BD FACSAria™ and BD FACSCanto II(Becton Dickinson, BD).

2.7.4 Intracellular staining protocol

Flow cytometry-based intracellular protein validation was completed by processing samples as described (section 2.6.1). The samples were resuspend in 1xPBS and incubated with a fixable viability dye for 30 mins at 4°C (ThermoFisher Live/Dead™ Fixable Aqua; L34965). The cells were then washed and resuspended in MACS buffer. FcR blocking was performed for 10 minutes at 4°C with 20ul per 100ul of sample with FcR block (Miltenyi, PN-130-059-901), then stained with the surface antibodies (Table 17) for 30 mins at 4°C, followed by 2ml MACS buffer washing at 1500rpm for 7 minutes. The True-Nuclear™ Transcription factor buffer set (Biolegend, PN-424401) was used for intracellular Ki67 staining. 1mL of the True-Nuclear™ 1X Fix Concentrate was added to each tube, and incubated at room temperature in the dark for 45 minutes. Without washing, 2mL of the True-Nuclear™ 1X Perm Buffer was added to each tube and centrifuged tubes at 385g at room temperature for 5 minutes, and the supernatant discarded. This step was then repeated, followed by resuspending the cell pellet in 100µL of the True-Nuclear™ 1X Perm Buffer. FcR blocking was repeated for 10 minutes at 4°C with 20ul per 100ul of sample with FcR block (Miltenyi), followed by staining with the intracellular antibody (Table 17), this was incubated in the dark at room temperature for 30 minutes. The cells were then washed with 2mL of the True-Nuclear™ 1X Perm Buffer and centrifuged at 385g at room temperature for 5 minutes (this step was then repeated). Cells were then resuspended in 0.5mL cell MACS and acquired on the BD FACSCanto II (Becton Dickinson, BD).

2.7.4.1 Flow cytometry analysis

All FACS data was analysed in FlowJo (v10.4.1), and geometric-mean fluorescence intensity and population percentage data were exported and visualized in Graphpad Prism (v8.0). For tSNE and co-expression analysis of flow cytometry data, each sample was reduced to exactly 10,000 randomly selected live and singlet-gated, CD14-CD19-CD20-CD4-CD45+CD3+CD8+ T cells. A tSNE plot was constructed using 1,000 permutations and default settings in FlowJo (v10.4.1), z-score expression was mean centred.

2.7.5 Flow cytometry panels

2.7.5.1 Initial FACS Panel

Antibody	Clone	Supplier	Vol (ul)/ 1e6 cells in 100ul
anti-CD45 FITC	HI30	Biolegend	2.5
anti-CD4 PE	RPA-T4	Biolegend	2
anti-CD8 PerCP-Cy5.5	SK1	Biolegend	2
anti-CD3 PE-Cy7	SK7	Biolegend	1.5
Viability - DAPI	5ug/ml	Sigma	2.5

Table 14. Initial Flow cytometry immune cell sorting panel.
FACS panel for sorting the CD4+ and CD8+ immune cells.

2.7.5.2 Sorting FACS Panel

Antibody	Clone	Supplier	Vol (ul)/ 1e6 cells in 100ul
anti-CD45 FITC	HI30	Biolegend	2.5
anti-CD4 PE	RPA-T4	Biolegend	2
anti-CD8 PerCP-Cy5.5	SK1	Biolegend	2
anti-CD3 PE-Cy7	SK7	Biolegend	1.5
anti-CD19 PerCP-Cy5.5	HIB19	Biolegend	2
anti-CD20 PerCP-Cy5.5	2H7	Biolegend	2
anti-HLA-DR APC	L243	Biolegend	2
anti-CD14 APC-H7	MϕP9	Biolegend	3
Viability - DAPI	5ug/ml	Sigma	2.5

Table 15. Flow cytometry immune cell sorting panel.
Extended FACS panel for sorting CD8+, CD4+, B cells and Macrophages.

2.7.5.3 Analysis of Tissue resident memory cells by FACS

Antibody	Clone	Supplier	Vol (ul)/ 1e6 cells in 100ul
anti-CD45 FITC	HI30	Biolegend	2.5
anti-CD4 PE	RPA-T4	Biolegend	2
anti-CD8 PerCP-Cy5.5	SK1	Biolegend	2
anti-CD3 APC-Cy7	SK7	Biolegend	1.5
anti-CD19 PerCP-Cy5.5	HIB19	Biolegend	2
anti-CD20 PerCP-Cy5.5	2H7	Biolegend	2
anti-PD1 PE-Cy7	eBioJ105	eBiosciences	2
anti-CD103 APC	Ber-ACT8	Biolegend	2
anti-41BB Pacific blue	4B4-1	Biolegend	2
Viability - LIVE/DEAD Fixable Aqua dead cell stain kit	-	Life Technologies	1.5

Table 16. Tissue resident memory cell FACS panel.
Evaluation of T_{RM} cells in NSCLC, non-involved lung and peripheral blood by FACS.

2.7.5.4 Intracellular analysis of Ki67 and Tissue resident memory markers by FACS

Antibody	Clone	Supplier	Vol (ul)/ 1e6 cells in 100ul
anti-CD45 FITC	HI30	Biolegend	2.5
anti-Ki67 PE	Ki67	Biolegend	2
anti-CD8 PerCP-Cy5.5	SK1	Biolegend	2
anti-CD3 APC-Cy7	SK7	Biolegend	1.5
anti-CD19 PerCP-Cy5.5	HIB19	Biolegend	2
anti-CD20 PerCP-Cy5.5	2H7	Biolegend	2
anti-PD1 PE-Cy7	eBioJ105	eBiosciences	2
anti-CD103 APC	Ber-ACT8	Biolegend	2
anti-41BB Pacific blue	4B4-1	Biolegend	2
Viability - LIVE/DEAD Fixable Aqua dead cell stain kit	-	Life Technologies	1.5

Table 17. FACS panel for the Intracellular assessment of Ki67 in Tissue resident memory cells.
Evaluation of Ki67 in T_{RM} cells in NSCLC, non-involved lung and peripheral blood by FACS.

2.7.6 Immune cell sorting for bulk immune cell transcriptomic analysis

After the correct setup of the BD FACSAria™ (BD Biosciences) using a standardised approach, detailed information in Appendices A.2. Briefly, the setup ensured accudrop beads (BD) were at 98>% purity during setup up of the drop delay, the 4-way sort positions were set using the collection tubes followed by sorting onto the lids for the fine tuning of the positions to ensure that the cells would be sorted into the centre of the Trizol LS (Ambion®, Fisher Scientific UK Ltd.). Once acquiring data, the population of interest (e.g. CD8+ T cells) were sorted directly into ice-cold Trizol LS reagent (Ambion®, Fisher Scientific UK Ltd.) to yield 1,000 to 50,000 immune cells at a flow rate of <2000 events/second on a BD FACSAria™ (BD Biosciences). The time from arrival of the tumour in the laboratory to processed was <3 hours. This process has been written in into an SOP and applied to the AMG319 and HARE-40 clinical trials (SOP no. = 384 see appendix A.2 Cell dissociation and immune cell sorting).

2.7.7 Cohort of Isolated immune cells from HNSCC and NSCLC

A cohort of purified immune cells (CD8+/ CD4+ T cells, B cells and macrophages/ activated monocytes) were collected from n=60 HNSCC (n=60/60 by O.Wood) and n=68 NSCLC (Appendix A.1.4 and A.1.5). The collection in NSCLC was a collaborative effort, please see acknowledgements (n=24/68 by O.Wood). Microscaled CD8+ T cell RNA-Seq was performed on n=41 HNSCC cases and n=43 NSCLC cases at the La Jolla Institute for allergy and disease.

2.7.8 Histological characterisation of NSCLC and HNSCC

2.7.8.1 CD8+ T cell and CD20+ B cell immune density

Immunohistochemistry was performed on formalin-fixed paraffin-embedded tumour sections with anti-CD8α (C8/144B; Dako; 1:100), anti-CD103 (EPR4166(2); Abcam; 1:500) and CD20 (CD20 (L26; Novocastra; 1:100). TILs were quantified using a Zeiss AxioCam MRc5 microscope (Zeiss) and Zeiss Axiovision software (version 4.8.1.0; Zeiss). An average of ten high-powered fields (×400) across representative areas of each tumour were counted to account for intratumoural heterogeneity; these were averaged to generate an intratumoural TIL score. Tumours were categorised by generating Tertiles based on the median quantification of cells, the top 1/3 (TIL^{Hi}) or bottom 1/3 (TIL^{Lo}).

2.8 Transcriptomic evaluation of purified immune cells by RNA-Seq

2.8.1 CD8 RNA-SEQ

Total RNA was purified using a miRNAeasy micro kit (Qiagen) and was quantified. Purified total RNA was amplified according to the smart-seq2 protocol. cDNA was purified using AMPure XP beads (1:1.1 ratio, Beckman Coulter). From this step, 1 ng of cDNA was used to prepare a standard Nextera XT sequencing library (Nextera XT DNA sample preparation kit and index kit, Illumina). Samples were sequenced using HiSeq2500 (Illumina) to obtain 50-bp single-end reads. Quality-control steps were included to determine total RNA quality and quantity, optimal number of PCR pre-amplification cycles, and cDNA fragment size. The RNA sequencing was carried out in collaboration with the La Jolla Institute for Allergy & Immunology following the methods outlined in (Rosales et al., 2018).

2.8.2 CD8 RNA-Seq analysis

RNA-Seq data were mapped against the hg19 reference using TopHat57 (v1.4.1.,–library-type fr-secondstrand-C). Sequencing read coverage per gene was counted using HTSeq-count (-m union -s yes -t exon -i gene_id, http://www-huber.embl.de/users/anders/HTSeq/). This bioinformatics was carried out at La Jolla Institute for Allergy & Immunology bioinformatics core.
To identify differentially expressed genes between patient groups, the Bioconductor package DESeq2 was used in the r statistical environment (RStudio). This performs a negative binomial tests for paired and unpaired comparisons, genes were considered differentially expressed when the DESeq2 analysis resulted in a Benjamini-Hochberg–adjusted P value of <0.05 (FDR <0.05). The Qlucore Omics Explorer 3.2 software package was used for visualization and representation (heat maps, principal component analysis) of RNA-Seq. Unsupervised hierarchical clustering of samples based on the expression of genes from the DESeq2 comparisons were visualised as Heatmaps.

2.8.2.1 Data repository for HNSCC and NSCLC CD8+ T cell RNA-Seq

CD8+ T cell RNA-Seq data are deposited into the Gene Expression Omnibus at the National Center for Biotechnology Information (NCBI) under accession code GSE90730.

2.8.1 Immune cell sorting for 10x single-cell transcriptomic analysis (10X Genomics)

For single cell analysis of B cells from tumour, non-involved lung and lymph node, CD19+ and CD20+ B cells were sorted and mixed into 50% ice cold PBS, 50% FBS (Sigma) on a BD Fusion cell sorter. The antibody fluorochrome combinations and clones are shown in Table 18.

The patients (n=12) used for 10x genomics were divided into n=6 $CD8^{Hi}$ and n=6 $CD8^{Lo}$ tumours, a total of n=6 were used for the background lung (n=3 from both $CD8^{Hi}$ and n=6 $CD8^{Lo}$ patients)(Appendix A.1.9). The cell sorting was undertaken at the La Jolla, Institute for Allergy and Immunology (Clarke et al., 2019).

Antibody	Clone	Supplier	Vol (ul)/ 1e6 cells in 100ul
anti-CD45 APC700	HI30	Biolegend	2.5
anti-CD56 BV570	HCD56	Biolegend	4
anti-CD8 PerCP-Cy5.5	SK1	Biolegend	2
anti-CD3 APC-Cy7	SK7	Biolegend	1.5
anti-CD19 BV-421	HIB19	Biolegend	2
anti-CD20 BV-421	2H7	Biolegend	2
anti-CXCR5 BB515	RF8B2	BD biosciences	4
anti-CD103 PE-Cy7	Ber-ACT8	Biolegend	2
anti-CD25 PE	BC96	Biolegend	4
anti-CD127 APC	A019D5	Biolegend	4
Live/ Dead PI	/	Sigma	

Table 18. Sorting panel for 10x genomics single cell experiment.
FACS panel for immune cell isolation.

2.8.2 Single cell RNA-Seq of B cells

Single cell analysis was performed at the La Jolla, Institute for Allergy and Immunology on the same day as the cell sorting. Single-cell RNA-seq using 10X genomics platform was performed using Chromium™ Single Cell 3' v2 Reagent Kits following the manufacturer's protocol; 11 and 12 cycles were used for cDNA amplification and library preparation respectively. Batch effects were minimised by sorting groups of 6 donors per day, cells were pooled for 10X sequencing library preparation (See appendix A.1.9). Barcoded RNA was collected and processed following manufacturer recommendations. Libraries were sequenced on a HiSeq2500 and HiSeq4000 (Illumina) to obtain 100- and 32-bp paired-end reads using the following read length: read 1, 26 cycles; read 2, 98 cycles; and i7 index, 8 cycles (Clarke et al., 2019).

2.8.3 Analysis of single cell RNA-Seq data

Single-cell RNA-seq data were mapped against the human hg38 reference genome using TopHat (v1.4.1., --library-type fr-secondstrand -C) and Gencode version 19 (GRCh37.p13) as gene model reference for alignment. Sequencing read coverage per gene was counted using HTSeq-count (-m

union -s yes -t exon -i gene_id, http://wwwhuber.embl.de/users/anders/HTSeq. Multiple sequencing runs were merged using cellranger count function in cell ranger, then merging multiple cell types with cell ranger aggr (v2.0.2). The merged data was transferred to the R statistical environment for analysis using the package Seurat (Macosko et al., 2015; Patil et al., 2018) (v3.0). RNA-Seq mapping and formation of a standardised data frame for analysis was performed by Ciro Ramirez Suastegui, La Jolla Institute for Allergy and disease.
The Seurat V3.0 package was used for subsequent single cell analysis. Only cells expressing more than 200 genes and genes expressed in at least 3 cells were included in the analysis. The data was then log-normalized and scaled per cell, followed by variable genes detection. Transcriptomic data from each cell was then further normalized by the number of UMI-detected and mitochondrial genes. A principal component analysis was then run on variable genes, and the first 12 principal components (PCs) were selected for further analyses based on the significance and standard deviation of PCs, as determined by the Jackstraw plot and "elbow plot" in Seurat. Cells were clustered using the FindClusters function in Seurat with default settings, resolution = 0.6 and 12 PCs. Differential expression between clusters was determined by analysing cluster specific differences using MAST (model-based analysis of single cell transcriptomics; $q \leq 0.05$ and FC ≤ 0.25) (Finak et al., 2015). Cluster specific genes were evaluated with gene ontology analysis using ToppGene Suite to determine gene families and biological processes that were significantly ($q<0.05$) over-represented (Chen et al., 2009), followed by redundancy reduction with REVIGO (Supek et al., 2011). Genes of interest were visualised in the single cell RNA-Seq data with a combination of T-distributed Stochastic Neighbour Embedding (t-SNE) and violin plots. Gene expression was overlaid on to tSNE plots, grey = low; navy blue = high expression. Violin plots were used to display the median and distribution of the gene expression in the different clusters.

2.9 Retrospective analysis of NSCLC and HNSCC survival data

2.9.1 NSCLC and HNSCC patient cohorts

A cohort of NSCLC (*n* = 689) were used for retrospective analysis of survival according to density of CD8+ or CD103+ TILs (IHC). A cohort of previously reported HNSCC cases (n = 123) (Ward et al., 2015) were again used to assess overall survival according to density of CD8+ or CD103+ TILs. Clinical details for the NSCLC cohort were collated by Dr Serena Chee, The HNSCC cohort were collated by Dr Matthew Ward. Tissue microarray construction and IHC was performed by research histology, Southampton Hospital.

2.9.1.1 CD8+ and CD103+ Immune cell survival analysis

Tumour-tissue microarrays from patients with NSCLC were stained with anti-CD8α (C8/144B; Dako; 1:100) or anti-CD103 (EPR4166(2); Abcam; 1:500) and were viewed under low-powered magnification (2.5× objective) to determine the density of CD8+ or CD103+ cells, a score of High, moderate or low was assigned and independently checked by an accredited pathologist (Prof. G.J.Thomas).

2.9.2 Survival data and analysis

The primary endpoint was overall survival, and survival time was measured from the date of diagnosis until the date of death or date last seen alive. All Kaplan–Meier plots were made in Graphpad Prism (v8.0), with log-rank tests to determine significance of overall survival (p-values) (Ward et al., 2015). Patients were excluded from analysis if survival was <30 days, to exclude the possibility of surgery-related mortality.

Chapter 3: Immunological and transcriptomic analysis of tumours

3.1 RNA isolation from solid tumours

3.1.1 Introduction

Analysis of tumour RNA by transcriptomics has been applied extensively to a large range of tumours using both micro arrays and RNA-Seq (Cancer Genome Atlas, 2015) (Thurlow et al., 2010). This gives insight into mutational load, transcriptomic variation and immune infiltrate levels (Ottensmeier et al., 2016a). Isolation of RNA from tumour tissue is a well-established technique and was used in the analysis of multiple cancers by the CGA (Cancer Genome Atlas, 2015), with the RIN (RNA integrity number) allowing the level of RNA degradation to be estimated in a sample prior to further analysis (Schroeder et al., 2006). A set of mouse TC1 tumours were used to establish a protocol and look at RNA yield and RNA integrity prior to testing on human tumour samples.

3.1.2 Objectives

Establish a work flow for the isolation of RNA from cryopreserved samples. In order to ensure no degradation had occurred during the procedure, two key factors needed to be taken into account: the yield of RNA from small cryosections and also the integrity of the RNA. This will be carried out on a test cohort of mouse TC1 tumours prior to use on human tumour HNSCC samples.

3.1.3 Results

3.1.3.1 Optimising RNA isolation from frozen tumour tissue.

3.1.3.1.1 Mouse TC1 RNA

In order to establish a robust RNA isolation procedure, the TC1 cell line was injected sub-cutaneously (5e5 TC-1 cells) into C57/BL6 mice; the resulting tumours were harvested at day 15-30. The RNA yield was assessed by taking 10µm cryosections from a 3mm x 3mm tumour (TC1) ranging from 5 to 30 cryosections (n=5, 10, 20 and 30), these were done in triplicate on the same tumour followed by RNA extraction (Qiagen RNeasy) and enumeration using the nanodrop to give ng/µl. The yield of RNA increases with the number of cryosections as would be expected. The

linear relationship continues up to 30 cryosections, which does not overload the RNA extraction column (data shown in Appendix A.3). The yield of the column will plateaux once the binding capacity has been reached. Using this information, 20-30 cryosections were set for optimal RNA yield with an adjustment of +/- 5 to 10 sections depending on the tumour size. The average yield of RNA from 22 TC1 tumours was 378ng/µl (total elution volume = 30 µl) (data shown in Appendix A.3) with an average RIN of 8.9, showing a robust procedure that yields excellent RNA quality, with the exception of 1 case that was likely due to experimental error/ tissue handling.

3.1.3.1.2 Isolation of RNA from HNSCC tumour samples

The optimised procedure was applied to HNSCC tumours where a consistent yield of RNA was obtained from tumours ranging in size from 2mm x 2mm to 7mm x 7mm, with an average of 221 ng/µl and 219 ng/µl for dataset TRTS (tumour replicates across time and space cohort) and HPVN (HPV positive compared to HPV negative cohort) respectively (Figure 9A). The average RIN for cohort TRTS= 8.9 with an average RIN = 7.4 for cohort HPVN (Figure 9B). From the HPVN (HPV positive compared to HPV negative) cohort 11/51 were removed due to poor RIN, 7 of those were obtained from Liverpool (total Liverpool samples = 10), indicating a potential storage or shipping issue for those tumours. The TRTS cohort (tumour replicates across time and space) produced better RNA quality with an average RIN of 8.4 and similar RNA yields, 7 samples were removed prior to sequencing in cohort TRTS (7/47).

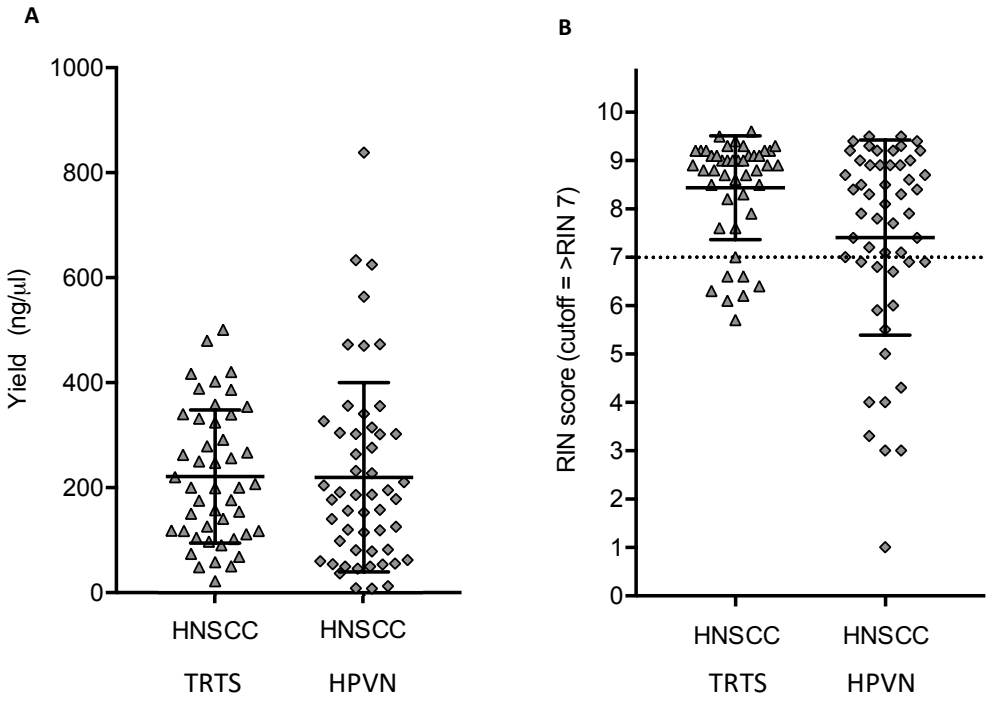


Figure 9. HNSCC cohorts RNA yield and integrity.
(A) HNSCC cohort TRTS (1) and HPVN (2) RNA yield. The average RNA yield was 221 ng/µl and 219 ng/µl for respectively. (B) HNSCC cohort TRTS (1) and HPVN (2) RIN assessment. An average RIN of 8.9 was seen in the TRTS cohort and an average of 7.4 in the HPVN.

3.1.3.2 Discussion

Successful isolation of high-quality RNA from 10µm cryosections was achieved on TC1 tumours, yielding excellent RIN scores with only one sample being removed due to RNA degradation, which was mostly likely due to sample handling and/ or the freezing protocol. Isolation of RNA from HNSCC tumours using the same methodology yielded similar results with robust RIN scores indicating intact RNA. However, the HNSCC patient samples displayed a higher number of cases with poor RIN scores as a result of RNA degradation. In particular, a group of samples from Liverpool where 7/10 of the samples displayed very poor RIN values, a review of the shipping log highlighted a potential problem with the dry ice volume upon arrival.

3.1.4 Conclusion

The optimisation of tumour RNA isolation from mouse TC1 tumours was achieved with a high average yield and RIN score. The same protocol was then applied to HNSCC tumours with the RNA yield being comparable to that of the TC1 mouse tumours. However, the RIN score for the HNSCC was more variable but still gave high average RIN scores. This meant the protocol was effective for routine extraction of RNA for high resolution transcriptomic analysis using RNA-Seq.

3.2 Analysis of HNSCC tumour replicates across time and space

3.2.1 Introduction

Morphological heterogeneity in cancer that is visible at the microscopic level is used by pathologists in routine diagnostics to inform treatment decisions. Recent data investigating tumours at much higher resolution using transcriptomics show that tumour cell clones undergo marked diversification over space and time as a result of mutational divergence and selective pressure (e.g. immunological attack) (Alizadeh et al., 2015, Dunn et al., 2002, Hensley et al., 2016, Jamal-Hanjani et al., 2015, Rooney et al., 2015, Dunne et al., 2016, Gyanchandani et al., 2016). Immune attack can vary according to its location in the cancer tissue: when quantifying the density of tumour-infiltrating immune cells by microscopy, the spatial distribution of immune cells with respect to tumour cells needs to be taken into account (Bindea et al., 2013). Given that immunotherapeutics only benefit a minority of patients (Seiwert et al., 2016), heterogeneity of immune cell distribution might contribute to treatment failure; and in that case small biopsy samples might not accurately represent the tumour microenvironment and immune status of a patient's tumour. In clinical practice however, it is difficult to sample multiple tumour areas at any time other than at surgical resection. Valuable prognostic information can be gained from simple enumeration of T-cell infiltrates in both human papillomavirus expressing [HPV(+)] and HPV independent [HPV(-)] head and neck squamous cell carcinoma (HNSCC) (Ward et al., 2015, Ward et al., 2014b).

Transcriptomic analysis of primary HNSCC using RNA-sequencing (RNA-Seq) allows further characterization of global molecular signatures (Cancer Genome Atlas, 2015) and detailed assessment of immune infiltrates (Wood et al., 2016a). Such evaluation of gene expression provides unprecedented insight into biological processes that occur in the tumour tissue. Like morphological assessments, molecular data are shaped by tumour-driving mutations, tissue heterogeneity, disease progression or selective pressures from treatment (McBryan et al., 2015, Wyatt et al., 2014, Zhang et al., 2014). Technical issues may further affect interpretation: in breast cancer, significant differences were identified between diagnostic tissue core biopsies when compared to tumour excisions (Pearce et al., 2016). Surgical ischemia was identified as a significant confounding factor in transcriptomic analysis.

Knowledge of variability across tumours is critical for the understanding and effective evaluation of immunotherapy interventions using checkpoint inhibitors (e.g. anti-CTLA4, anti-PD1 and anti-PDL-1 antibodies) (Garon et al., 2015, Hodi et al., 2010). The mechanisms that underpin immunotherapeutic success as opposed to treatment failure, remain poorly understood but must be linked to the pre-existing immune infiltrates (Gentles et al., 2015, Rooney et al., 2015, Tumeh

et al., 2014). To make sense of treatment failure and of any changes detected in longitudinal assessments, understanding of natural variability is critical, as is the understanding of how results from different assays (e.g. histological and transcriptomic assessment) correlate with each other. Clinical trials that test the efficacy of therapeutic agents are increasingly being used in "window of opportunity" Studies. These studies aim to investigate compounds in treatment-naïve patients prior to standard treatment - patients consent to receiving a new treatment strategy between diagnosis and delivery of their standard care. This has the benefit of testing therapeutic response in a patient group that has not been exposed to other anti-cancer treatments, hence avoiding possible confounding effects from any previous treatment (Glimelius and Lahn, 2011). An example of this type of study Is the phase II neoadjuvant trial AMG319 (AMG319 in HPV positive and negative HNSCC: NCT02540928), in which the drug AM319 is being studied in relation to the patient's own immune system in HNSCC. Patients are given a course of the drug between diagnostic biopsy and their normal care (surgical resection), enumeration of CD8 cells on FFPE tumour samples and analysis of RNA-Seq on snap frozen tissue are being used as trial endpoints to assess the drug.

3.2.2 Objectives

To evaluate the variability in the distribution of immune infiltrates in HNSCC across space and time using transcriptomic analysis (RNA-Seq) and CD8 IHC on tumour replicates at diagnostic biopsy and surgical resection. The stability of a tumour over this time scale is important when considering "window-of-opportunity" studies, in which intervention with immunotherapeutic agents such as AMG319 can be used in treatment naïve patients. The outcome will be to assess if tumour biopsies taken for "window-of-opportunity" studies are stable at an immunological level over time and space. A stable immune and tumour profile will enable a higher level of confidence in observable differences post immunotherapy treatment.

3.2.3 Results

3.2.3.1 Patient summary of TRTS cohort (tumour replicates across time and space)

A total of 16 patient tumour samples were collected (n=44 replicates), the full analysis included 14 patients and n=37 tumour replicates. Single timepoint replicates (n=15) were obtained from 6 patients, 8 patients had replicates (n=22) from two timepoints, diagnosis and resection, with an average of 26 days between them; cohort details shown in Appendix A1.1.

3.2.3.2 Principle component analysis (PCA) of tumour replicates for all gene features and 4000 genes.

The RNA-Seq data was first visualized using Principle component analysis, Figure 10 A displaying the 4000 most variable genes and Figure 10 B all genes (n=18979). When displaying all genes and also the 4000 most variable genes, replicates from the same patient group together, showing globally similar gene expression profiles. PCA of 4000 genes begins to separate the cases into different clusters, which is not as apparent when visualising all genes. Specifically, case 1, 7 and 10 begin to group together in addition to cases 5, 6 and 11. This highlights that each case likely has a distinct transcriptional profile, while also sharing global similarities to other tumours.

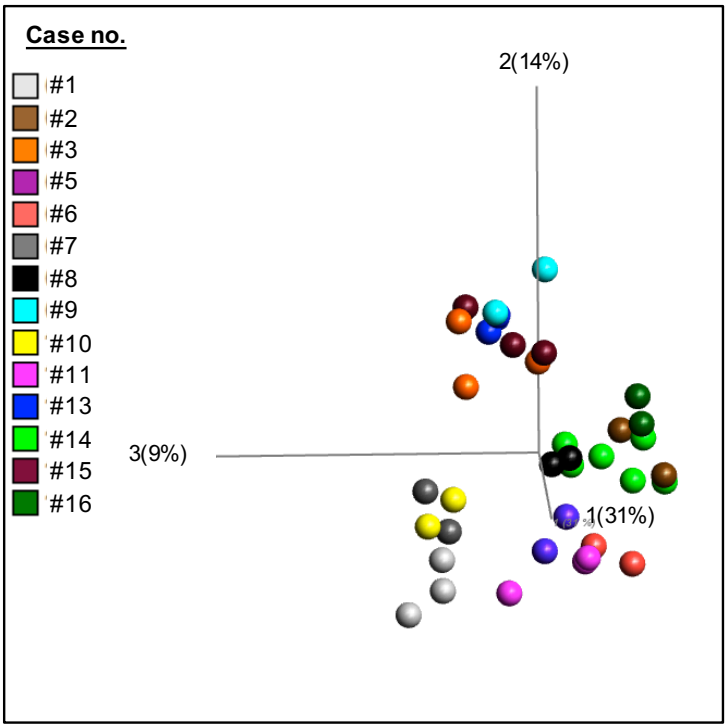
3.2.3.3 Tumour replicates from the same patient cluster hierarchically.

A decreasing number of gene features were visualised across the cohort by variance filtering (removes genes that have similar expression across the cohort) the RNA-Seq data. All genes (n=18979), 4000, 1000, 100 and 20 genes were displayed to asses patient replicate clustering at different gene numbers. These were hierarchically clustered and displayed in heatmaps Figure 11 shows all genes and 4000 genes; with Appendix B.1.1 showing 1000, 100 and 20 genes). The 37 replicates from 14 patients evaluated at either diagnosis or resection cluster well across the decreasing gene features (Table 19 and Appendix B.1.1). The varying number of gene features highlights that each patient has a distinct tumour gene expression that is conserved across the replicates, some patients share similarities, for example cases 9 and 13, where others display a more unique expression profile (case 6). As the filtering by variance reduces the gene number to 20 genes, gender-specific gene differences are identified, females show differential expression of XIST and males express DDX3Y and RPS4Y1. In order to rationalise the number of genes for downstream analysis, 4000 genes were selected, this number of genes was then used for further correlation analysis and Euclidean distance assessment.

No. of genes	All (18979)	4000	1000	100	20
No. Clustering	9/14	9/14	13/14	13/14	13/14

Table 19. Number of hierarchically clustered patients for each gene feature.
Tumour replicates cluster well across a variable number of gene features.

A



B

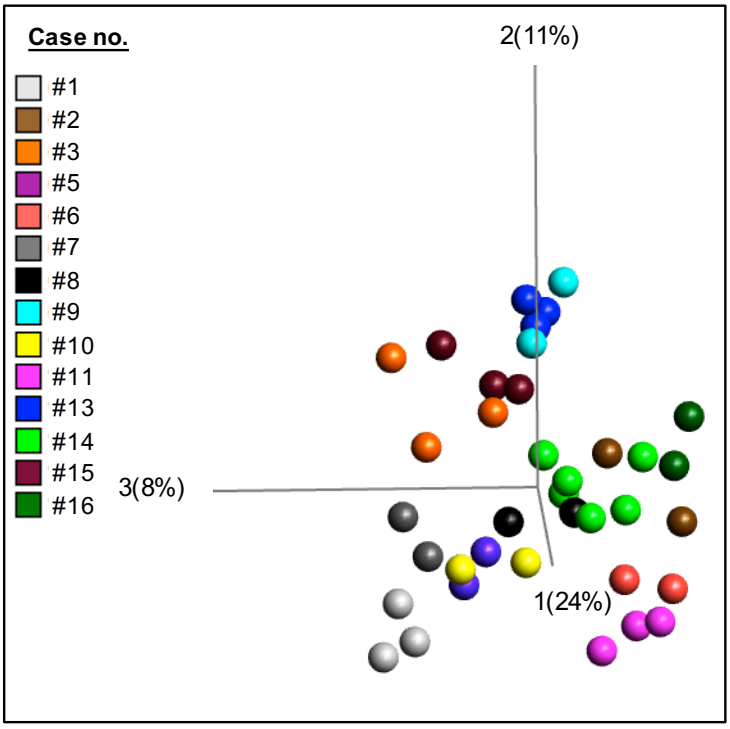


Figure 10. Analysis of tumour replicates using Principle Component Analysis (PCA).

RNA-Seq analysis showing normalized read counts for the 4000 (variance filtered) most variable genes across the tumour replicates. Patient tumour replicates are colour coded and displayed on the principle component analysis (PCA) plot. The axis represents the 1st, 2nd and 3rd principle component, with the associated % variance. (A) PCA Displaying the 4000 most variable genes across tumour replicates and (B) PCA of all genes. Both plots show the similarities between many of the tumour replicates (e.g. case #13 where the replicates are close together). Conversely, case #3 shows that the replicates are further apart and less closely related.

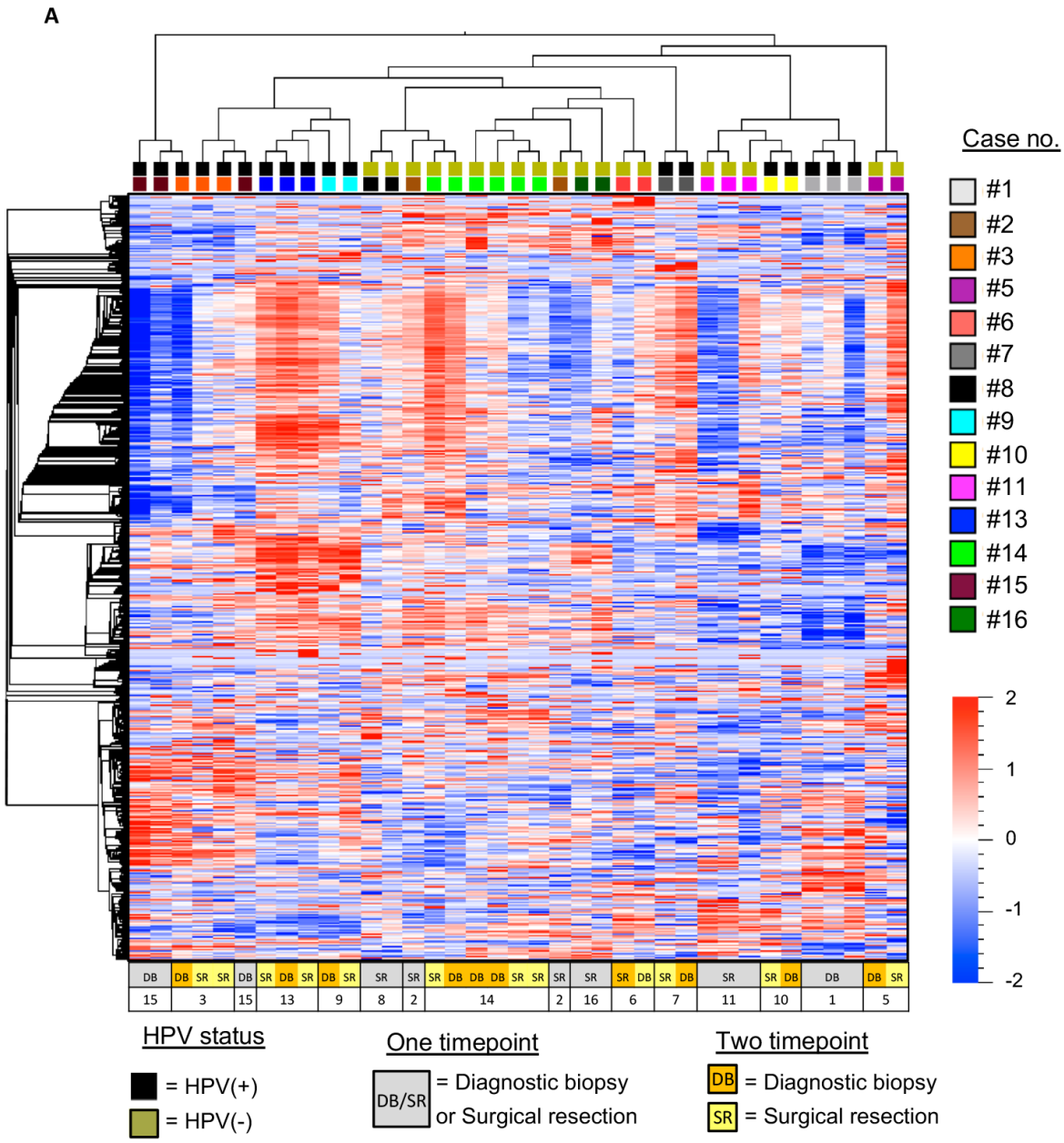


Figure 11. Hierarchical clustering of cases by decreasing gene features using variance filtering.

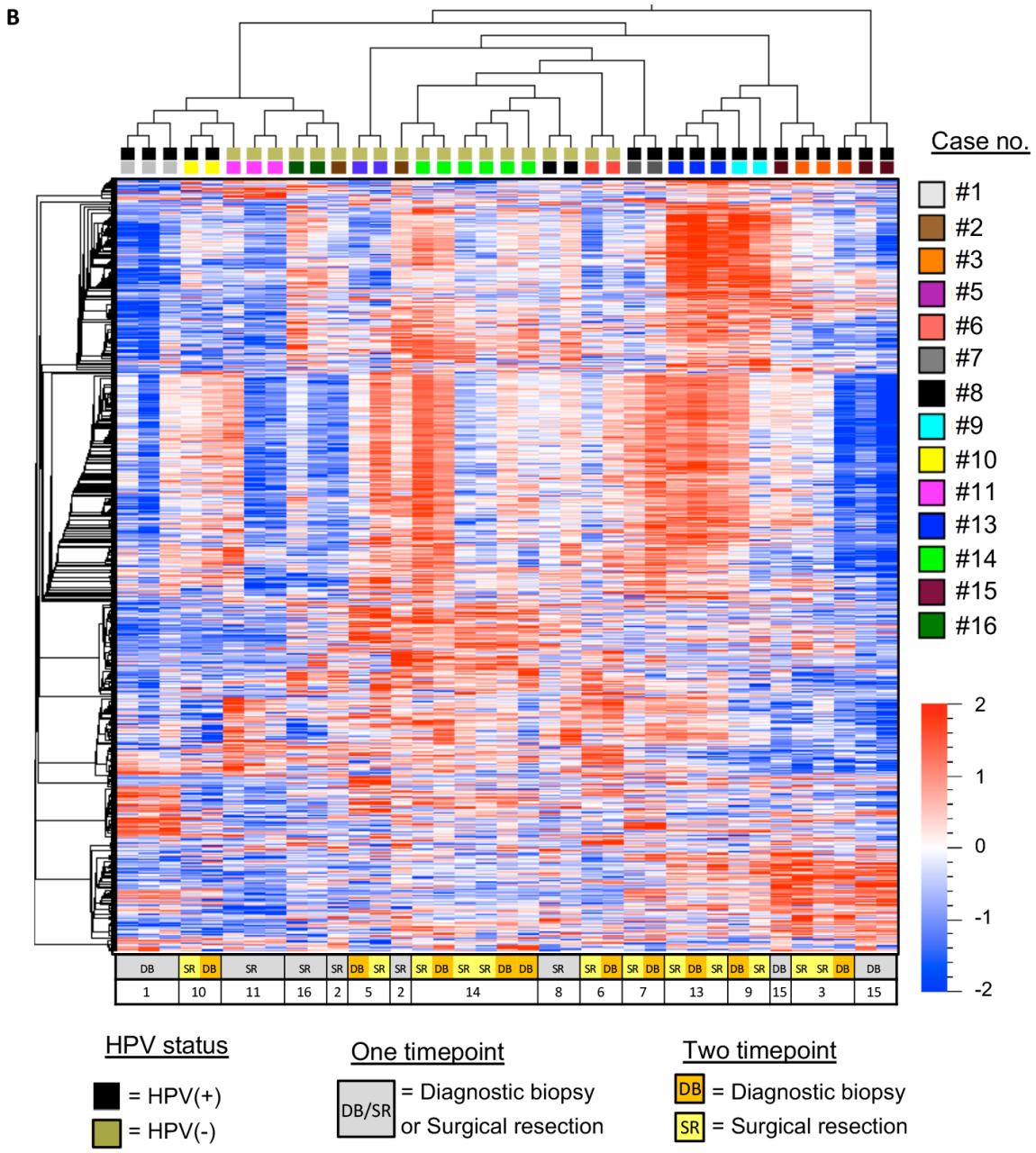


Figure 11. Hierarchical clustering of cases by decreasing gene features using variance filtering.
A heatmap of RNA-Seq data displaying all genes (A), and the 4000 most variable genes (variance filtered) (B). Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of tumour replicates shows clustering of related replicates in many cases (e.g. case #5). Conversely, it also shows some cases where the replicates cluster further apart (e.g. case #2). Data is shown as row wise z-scores of normalized read counts; patient tumour replicates are colour coded; HPV(+)=black and HPV(-)=beige; diagnostic biopsy (DB) and surgical resection (SR) replicates are annotated below.

3.2.3.4 Euclidean distance analysis of hierarchical clustering

In order to quantify the level of variation between replicates, the Euclidean distance was assessed for each replicate (Figure 12 A); this is a numerical value determined from the distance metric used in the hierarchical clustering. The graph displays each case and the median intrapatient (circle) distance compared to the median interpatient distance (diamond). The length of the line

connecting circles and diamonds is a measure of difference when replicates from one case are compared to all other samples. A long line indicates large differences (e.g. case #13), conversely a short line highlights cases where this difference is small (e.g. case #5). The minimum and maximum distance for each replicate is shown in Figure 12 B, here the intrapatient replicates can overlap with another replicate in some cases (e.g. case 7) where the minimum distance is lower than the median value for that replicate. This indicates that although replicates on the whole are very consistent, they do share similarities with other cases, this was also observed in the Figure 11 heatmaps. The individual Euclidean distances are also presented as a scatter plot in Figure 12 C demonstrating that on the whole, tumour samples from the same patient (intrapatient) are significantly closer to each other than to those from other patients (interpatient) ($p=<0.0001$).

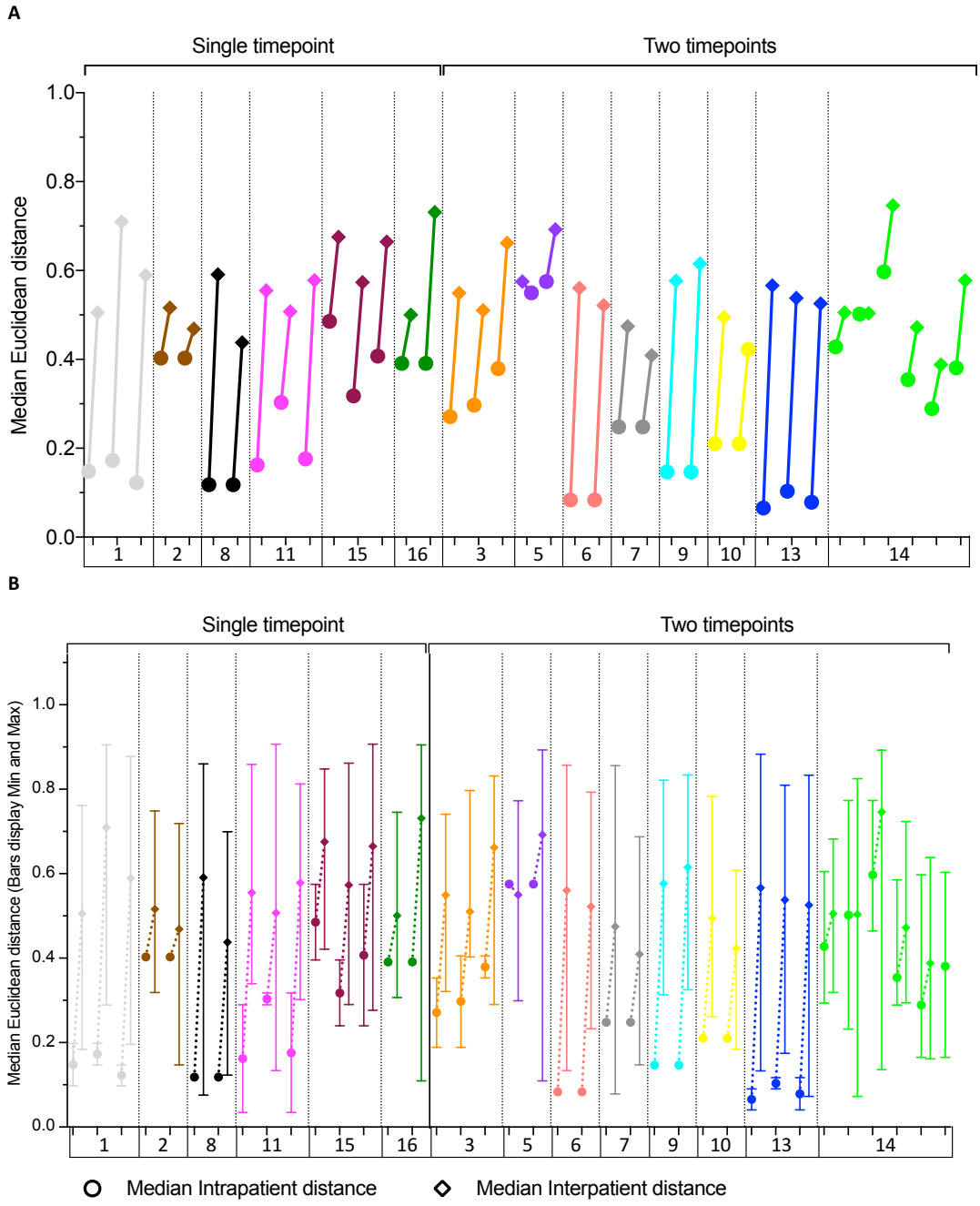


Figure 12. Comparison of Euclidean distance from hierarchical clustering of tumour replicates.
The distances between paired tumour replicates in the hierarchical tree was calculated and used to assess how closely related intrapatient replicates were, compared to interpatient replicates. The median Euclidean distance (sqrt(1-pearson)2) for intrapatient replicates (circles) was plotted, as was the median Euclidean distance for interpatient replicates (diamonds). The smaller the distance and the shorter the line the more related the intra and interpatient samples are, conversely a larger distance or longer line identify lower homology between intra and interpatient samples. (A) Displays the relationship between the intra and interpatient median Euclidean distances for each replicate in the cohort respectively. The length of the connecting line between circles and diamonds is a measure of the difference between an individual replicate and all samples from all other patients. (B) Comparison of minimum and maximum Euclidean distance from hierarchical clustering of tumour replicates. The minimum, maximum and median Euclidean distance for intrapatient replicates was plotted alongside the median Euclidean distance for interpatient replicates. Overall intra patient tumour replicates are closer to each other than the interpatient replicates.

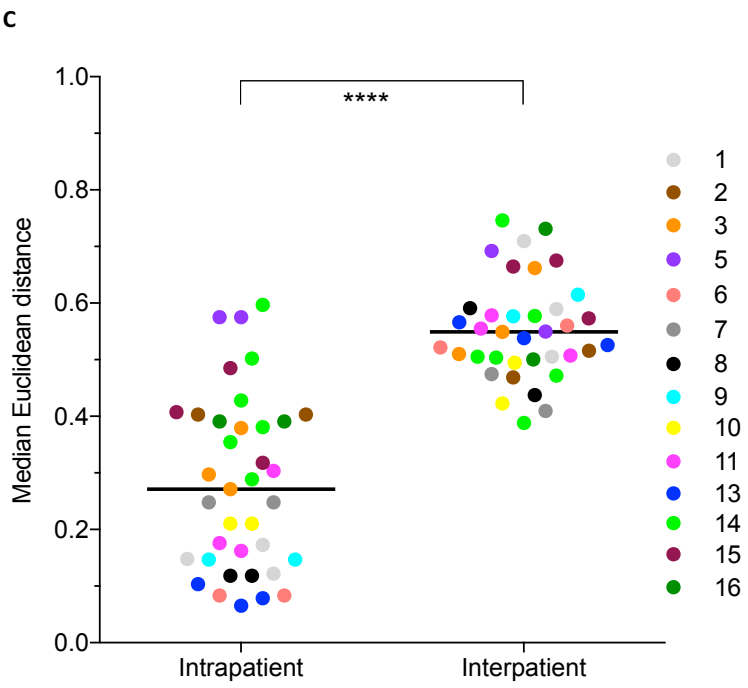
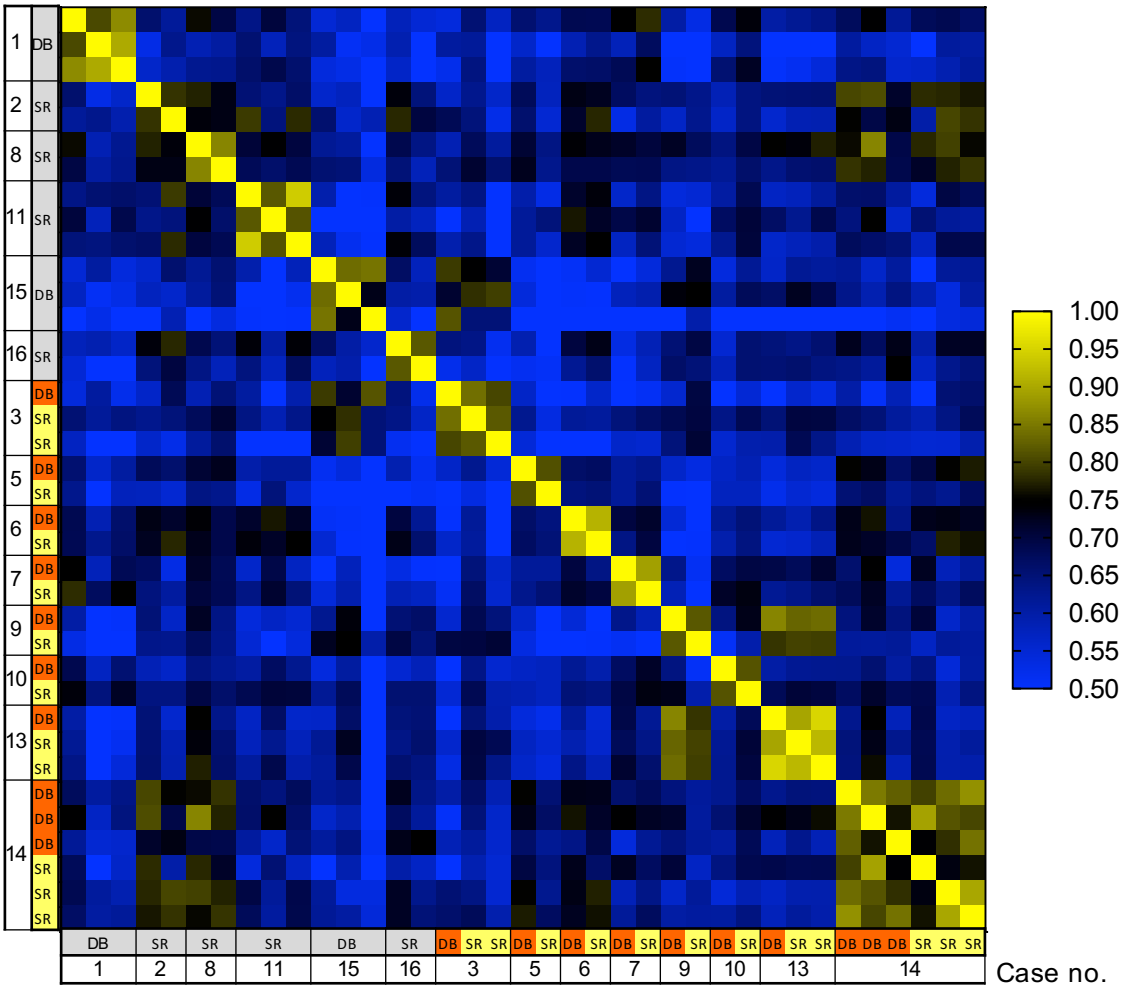


Figure 12. Comparison of Euclidean distance from hierarchical clustering of tumour replicates. (C) Represents a scatterplot of the median intrapatient distances compared to the median interpatient distances. There is a significant difference between intrapatient and interpatient distances (p=<0.0001, Wilcoxon test).

3.2.3.5 Replicates from the same patient are highly correlated.

Correlation (Spearman) analysis of the top 4000 most variable genes was carried out across all samples and tumour replicates (Figure 13 A). The r values in the correlation matrix range from r=1 (highly correlated, yellow) to r=0.5 (less correlated, values <0.5 are shown as blue). It shows that intrapatient tumour replicates are more correlated to each other than interpatient replicates. Case #14 shows a variable correlation between samples, but there is a higher correlation between its own replicates than to samples from other cases. The global median correlation coefficient for intrapatient replicates was r=0.82 compared to an r=0.63 for interpatient replicates (Figure 13 A). When cases are displayed in a correlation matrix, cases #13 and #9 have similar gene expression; cases #2, #8 and #14 also display similar correlation values indicating similarities at the global transcription level, this was also visible in Figure 11 (heatmaps). The data again suggest that each patient's tumor has a distinct transcriptomic landscape. Correlation analysis of all genes across all samples with the inclusion of the less variable genes increases the correlation between samples (r scale=0.9 to 1); again, replicates from the same patient are most similar to each other (Figure 13 C and D).

A

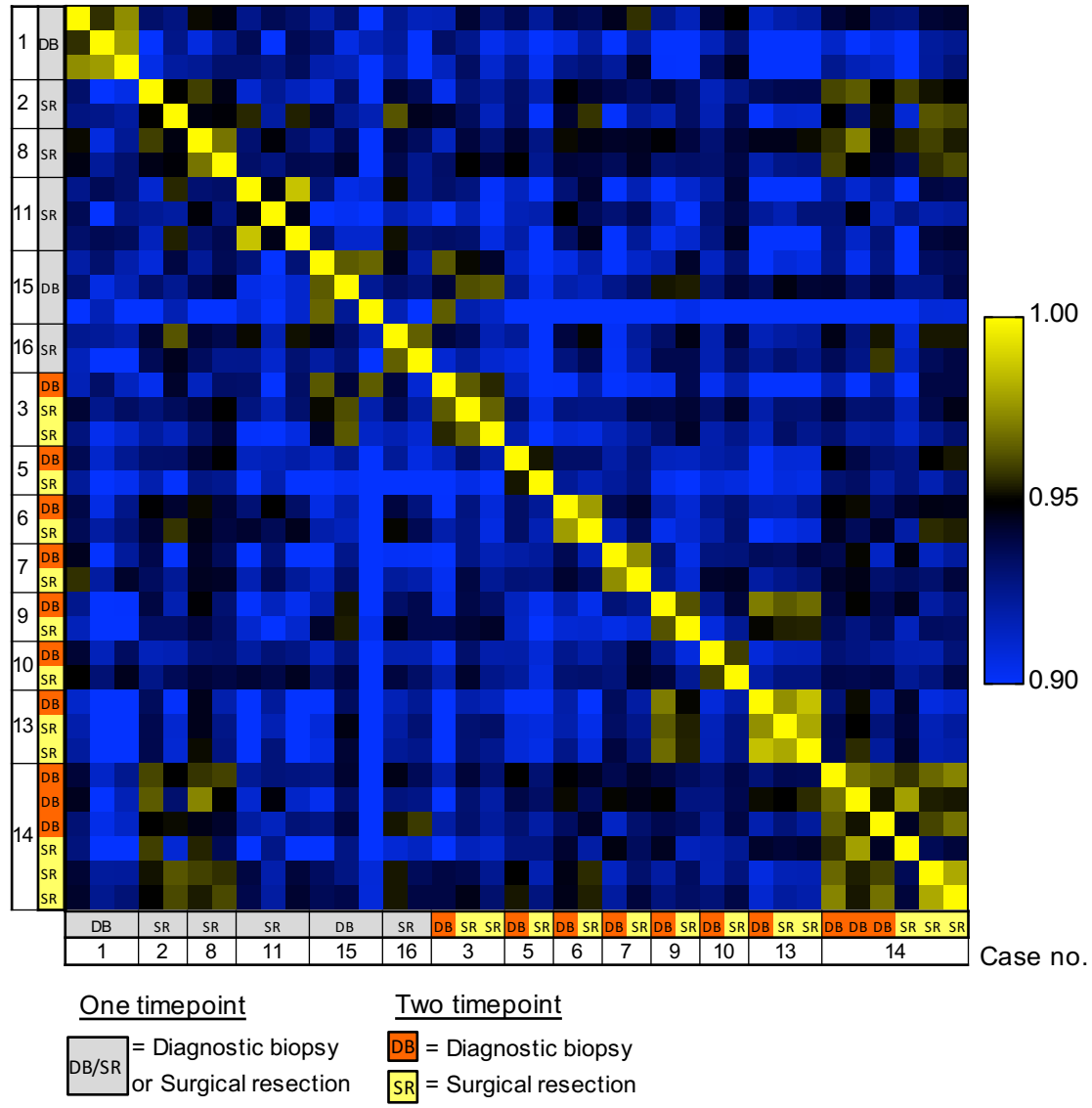


B

Spearman correlation **4000** genes		
r	Intrapatient	Interpatient
Median	**0.82**	**0.63**
Min	0.76	0.26
Max	0.93	0.86

Figure 13. Correlation analysis of tumour replicates.

C



D

Spearman correlation **all** genes		
r	Intrapatient	Interpatient
Median	**0.96**	**0.92**
Min	0.91	0.82
Max	0.98	0.97

Figure 13. Correlation analysis of tumour replicates.

Single timepoint (sampling across space) and two timepoint (sampling across time between diagnosis and resection) tumour replicates were assessed using a correlation matrix of gene expression (Spearman correlation of top 4000 variance filtered genes and then all genes); each tumour replicate's gene expression was correlated to itself and to each other sample. (A) and (B) 4000 genes show Intrapatient tumour replicates were more correlated with a median correlation of r=0.82 compared to an interpatient median correlation r=0.63. (C) and (D) displays correlation of all genes (n=18979), where Intrapatient tumour replicates were more correlated with a median correlation of r=0.96, compared to an interpatient median correlation r=0.92.

3.2.3.6 Immunological gene expression in tumour replicates from the same patient is consistent.

The prognostic importance of T-cell infiltration in HNSCC is well-documented (Ward et al., 2015, Ward et al., 2014b, Cancer Genome Atlas, 2015). Gene expression for immune cell lineage markers, effector function, exhaustion and regulatory genes were therefore represented in a heatmap (Figure 14 A). Samples were ordered by case number, then cases with replicates at a single timepoint were grouped, as were cases with replicates from two timepoints. Visual consistency in the immunological signature between replicates from the same patient including the markers of immune effector function (IFNG/ GZMA (Rooney et al., 2015)) and targets of cancer immunotherapy (PDL1/ CTLA4 (Ahmad et al., 2015)) was observed. Cases #1, #8, #5, #6 and #7 contain a consistent and low relative level of immune-gene expression, while cases #16, #9 and #13 have consistently high immune-gene expression. However, in some cases, replicates #2, #3, #6, #10 and one sample from case #14, the level of immune-gene expression is more variable. This variability is reflected in the hierarchical clustering of the tumour replicates for immune genes only, 9/14 cases cluster (Figure 14 B). The hierarchically clustered heatmap also shows a clear grouping of the tumours by the level of immune gene expression, the sample replicates could easily be divided into immune low, moderate and high. In addition to this, the RPKM of CD3E, GZMA, IFNG, CTLA4 and PDL1 (CD274) are displayed as dot plots as an alternative way of visualising the data and to see the spread in expression across the tumour replicates. (Appendix B.1.2). The dot plots complement the data in the heatmaps and show that the majority of the tumour replicates have very consistent expression of these genes.

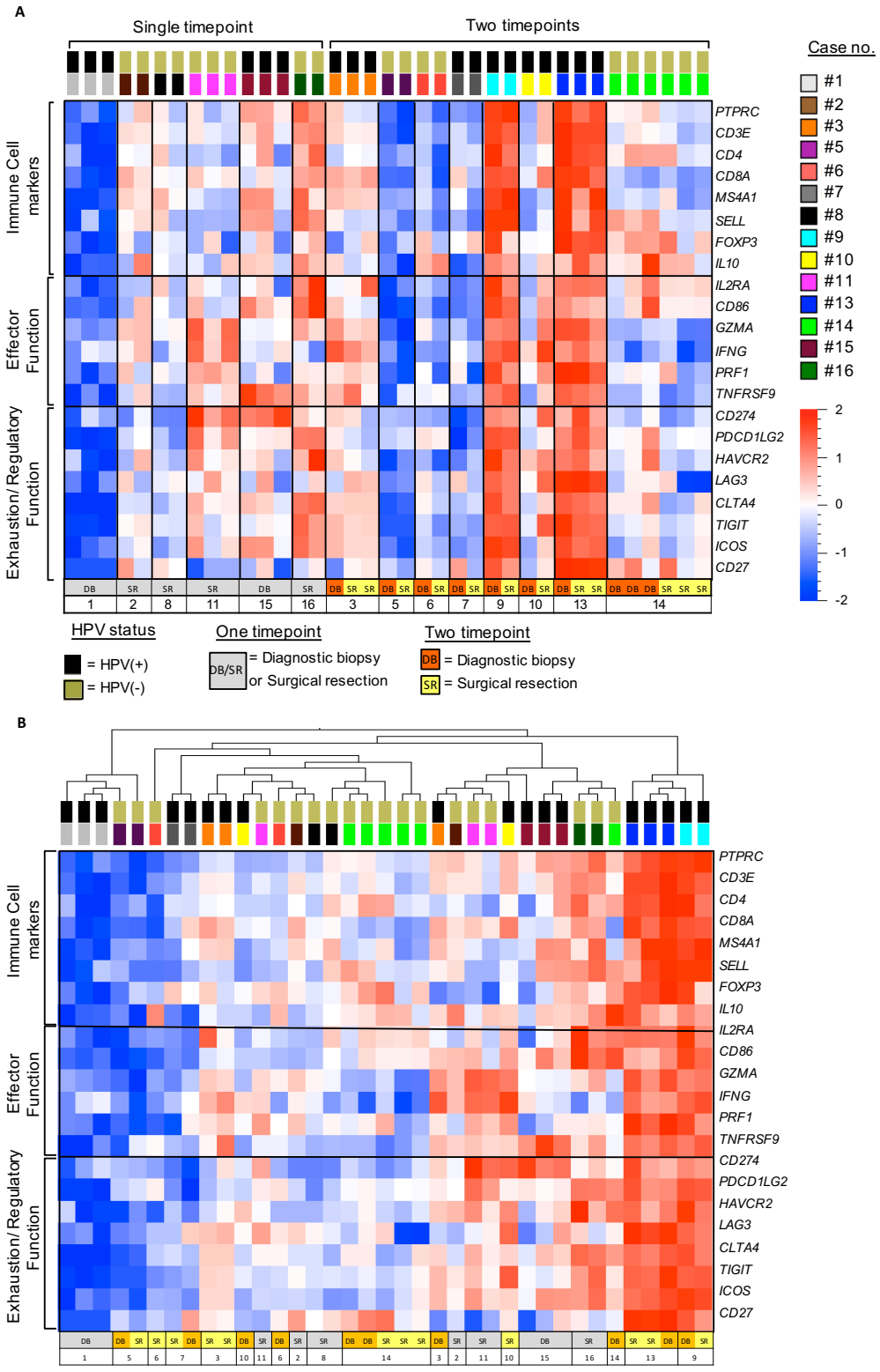


Figure 14. Gene expression of immune markers is consistent in tumour replicates from the same patient.

RNA-Seq analysis showing row wise z-scores of normalized read counts for genes associated with immune lineage markers, cytotoxic function, exhaustion and regulatory function. (A) Samples were grouped by patient and then by whether replicates had been collected at a single timepoint or at two timepoints. Expression profiles are consistent across tumour replicates in most cases. (B) Samples were hierarchically clustered (distance measure = Pearson's correlation metric; clustering = average linkage method). Tumour replicates cluster largely by patient displaying similar immune gene expression profiles.

3.2.3.7 Correlation of CD8+ numbers by IHC and gene expression.

The gene expression for CD8A was compared to the assessment of CD8 cell counts by standard IHC. This was done using tumour tissue sections that were immediately adjacent to those used for RNA-Seq. Material was available for 11/14 cases (insufficient material remained after RNA isolation for n=3). Gene expression for CD8A is presented as dot plots for cases grouped by whether we had replicates from a single timepoint or two timepoints (Figure 15 A). CD8+ cell counts (average of 10 HPF) evaluated by IHC are shown in Figure 15 B. Visually, the cases with high CD8A transcript count appear to be equally rich in CD8+ T-cells. This was quantified by assessing the Spearman correlation coefficient of the CD8A gene expression to CD8+ T-cell IHC count with an r = 0.82 (Figure 15 C), representative images are shown in Figure 15 D.

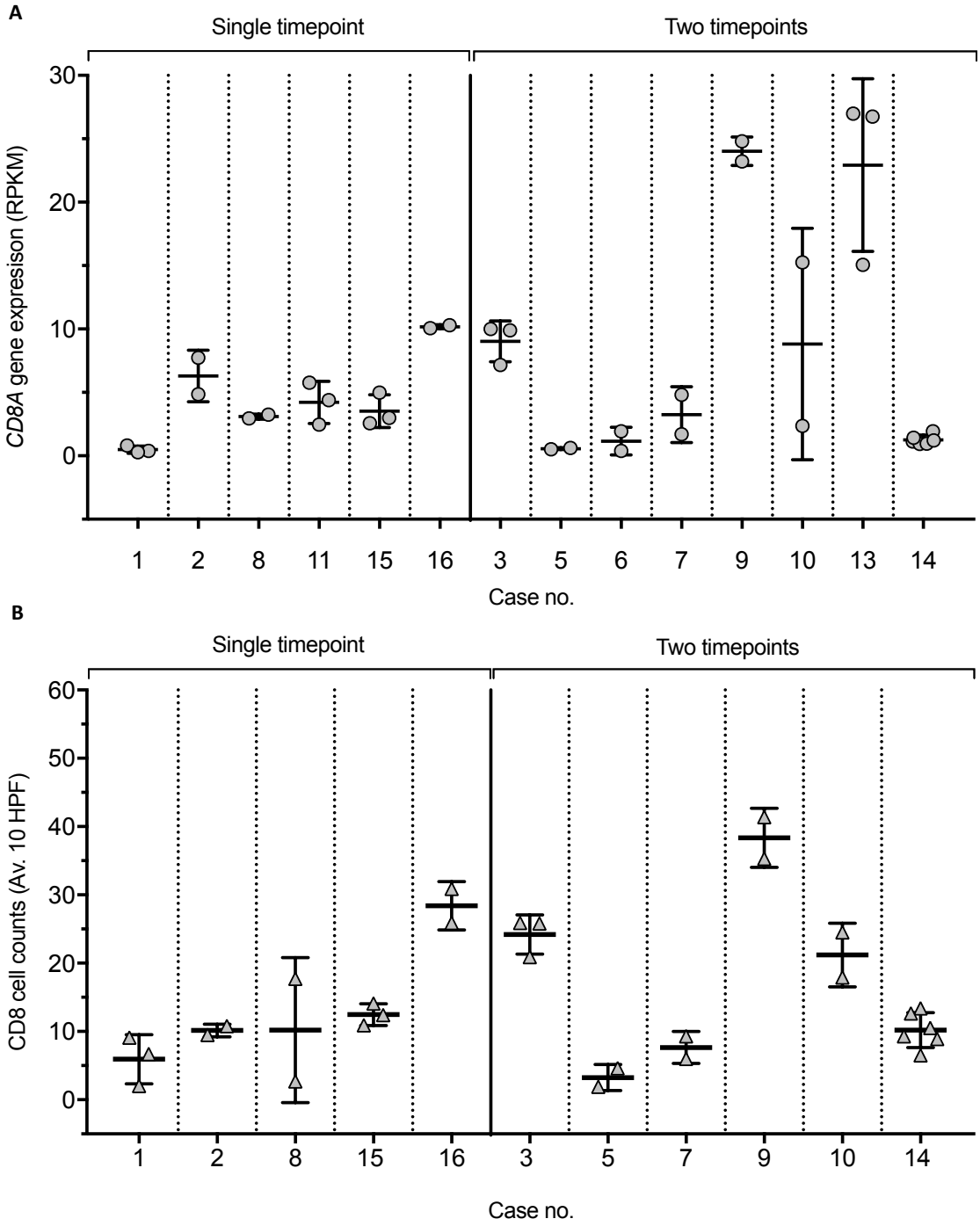


Figure 15. Comparison of CD8A gene expression with CD8 immunohistochemistry (IHC) counts across time and space.

(A) The gene expression of CD8A (RPKM – Reads per kilobase per million mapped reads) is shown for the cases grouped by single timepoint and two timepoint tumour replicates. (B) CD8 IHC counts shown as the mean across 10 high power fields (HPF) on frozen tissue sections taken from the adjacent material used in the RNA-Seq analysis, tumour replicates are arranged as described above.

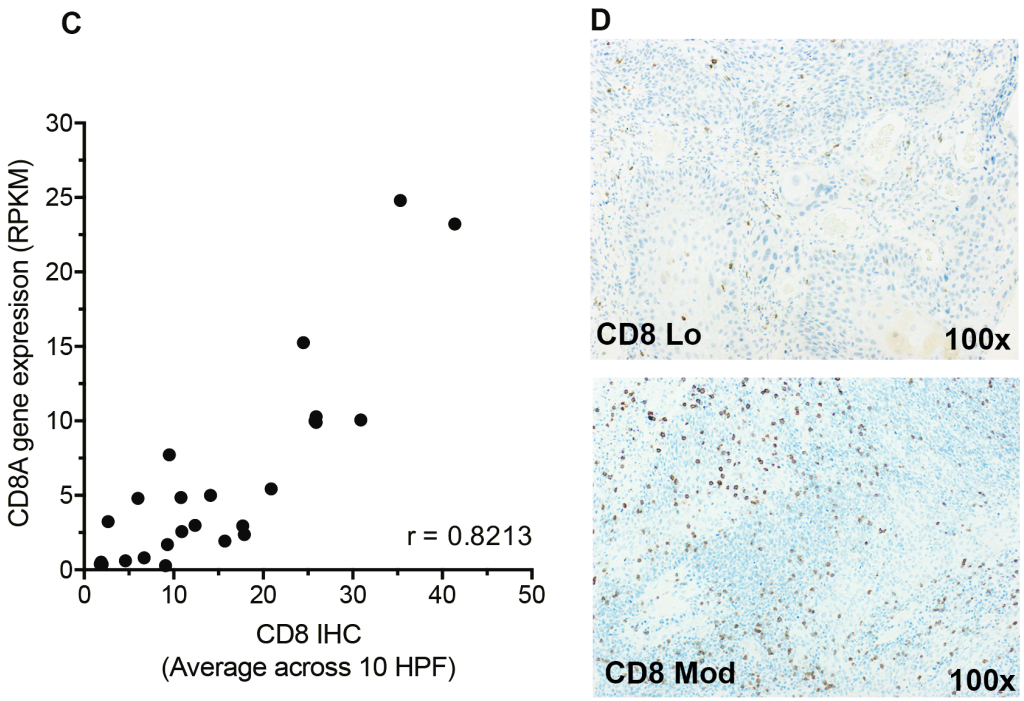


Figure 15. Comparison of CD8A gene expression with CD8 immunohistochemistry (IHC) counts across time and space.

(C) Spearman correlation analysis of CD8A gene expression and CD8 immunohistochemistry (IHC), Spearman analysis shows a high level of correlation between CD8 IHC and CD8A gene expression with an r = 0.82. (D) Representative images of CD8 IHC at 100x magnification, a $CD8^{Low}$ and $CD8^{Mod}$ are shown. The CD8A gene expression and CD8 IHC show similar pattern of immune density and are consistent across time and space with a high level of correlation. (IHC slides prepared by research histology – Monette Lopez and Maria Machado).

3.2.4 Discussion

In solid cancers, morphological heterogeneity is commonly observed and highlights the issues of how representative of the whole cancer small biopsy samples are. This is important clinically as biopsy samples are used to make treatment decisions and failure to appreciate heterogeneity could contribute to treatment failure. In addition, if tumour samples are used to gain detailed information into molecular events resulting from therapeutic intervention, understanding the degree of transcriptomic heterogeneity becomes important. This is especially true for window of opportunity studies that look to monitor treatment effects in treatment naïve patients. Current high-resolution studies have identified that beyond morphological features, genetic features can vary within an individual's cancer (Jamal-Hanjani et al., 2015, Dunne et al., 2016, Gyanchandani et al., 2016). The analysis of tumour and immunological heterogeneity at high resolution using RNA sequencing in multiple tumour biopsy samples from individual patients was carried out to try and address this issue. The focus was to understand gene expression that reflects adaptive immune attack, as boosting this is the key for success of immunotherapy. In HNSCC it has been shown that TIL density is tightly linked to outcome, in both HPV(+) and HPV(-) cancers (Ward et al., 2014b,

Wood et al., 2016a). Here RNA-Seq was used to quantify the immune signatures and global tumour gene expression profiles from tumour biopsies, separated in space (more than one sample at the same time) and time (samples taken at diagnosis and compared to samples from surgical resection). Molecular immune cell quantification was then correlated to data generated by the current 'gold standard', immunohistochemical assessment of HNSCC patients, which is known to be an important prognostic indicator (King et al., 2014, Ward et al., 2015, Ward et al., 2014b, Wood et al., 2016a).

The gene expression data generated using RNA-Seq for the patient cohort were assessed by PCA, hierarchical clustering and correlation analysis of the top 4000 most variable genes, as well as evaluation of specific immune-genes. Hierarchical clustering and Euclidean distance measures showed that multiple samples from the same patient were significantly more similar to each other than to those from other patients, this was also mirrored in the correlation analysis. The data suggest a single sample is able to capture the key immune characteristics of the patient's primary HNSCC, irrespective of whether patient samples were taken at the same time or at different timepoints. Tumour biopsies clustered by patient and also displayed similar expression levels of immune genes. In contrast, there is considerable variability between patients (Ottensmeier et al., 2016b, Ward et al., 2014b, Wood et al., 2016a). The similarity of tumour replicates was also maintained in both HPV(+) and HPV(-) HNSCC. This opens up the possibility of using gene expression analysis on biopsies to inform on treatment options (e.g. PD1 expression), it also justifies using biopsies before and after novel treatment regimens (e.g. PI3Kδ inhibitor) to measure change. Although the current study is limited to the window between the diagnostic biopsy and the surgical resection.

The study here has shown that a consistent immune signature in an individual tumour across both location and time was observed, at least within the window between diagnosis and resection (av. 26 days); additionally, the data demonstrates that tumours can be grouped by the immune-gene expression profiles. The evaluation of gene expression of key markers for immune attack (IFNG/ GZMA (Rooney et al., 2015)) and targets of cancer immunotherapy (PDL1/ CTLA4 (Ahmad et al., 2015)) further confirm a stable immune signature in the individual patient. The gene expression of CD8A correlated strongly (r=0.82) with the CD8 T-cell count assessed by manual counting of cells using IHC, this again adds weight to the use of gene expression as a tool for evaluating tumours. However, intrapatient tumour replicates were not a perfect match with each other and in some instances (cases #2, #3, #10, #11, #14 and #15) the tumour samples did not cluster perfectly by patient: In case #10, gene expression differences were observed at the immune-gene level which were also reflected in the IHC of CD8. Conversely, case #2 displayed a stable CD8 and immune cell marker signature but variability in the expression of effector function genes. A larger cohort of

samples taken at biopsy and resection would enable more confidence in the transcriptomic stability, in both spatially and temporally distinct tumour replicates. Cases #2, #6, #10 and #14 were laryngeal tumours, which have been highlighted as having a higher level of mutational heterogeneity compared to other sites in HNSCC (Ledgerwood et al., 2016). The heterogeneous genomic landscape that has been identified by other groups (Alizadeh et al., 2015, McGranahan et al., 2016) would need to be assessed on the tumour replicates in order to allow the assessment of genomic variability and its impact on the transcriptome.

3.2.5 Conclusions

The data supports that in HNSCC the global tumour and adaptive immune signatures are stable across space and time between replicate samples from the same tumour. This has a number of implications, it suggests that immunological heterogeneity is not likely to be a key reason for immunotherapy failure for the primary cancer in HNSCC; paired analysis of primary and metastatic disease would be needed to evaluate this question further for patients with later stage disease. It also implies that it is important to examine in other tumours whether a consistent immunological 'fingerprint' is present even in cancers with known genetic heterogeneity. Finally, it supports that from a background of transcriptomic stability in untreated patients, RNA sequencing may be useful for the detection of change resulting from treatment effects carried out during window of opportunity clinical trials.

3.3 Stratification of patients by immunotyping tumours with gene expression of key immunological markers.

3.3.1 Introduction

The characterisation of tumours at increasing levels of detail has allowed a greater understanding of tumour cell biology and its interaction with the host. The cancer genome atlas (CGA) presents a vast repository of data that can be interrogated from a variety of perspectives. One of the uses of such data is the analysis of the immune landscape in different cancers (Thorsson et al., 2018). In the study by Thorsson et al, the mutational characteristics and immunological features were determined and shown to have prognostic importance across 33 cancer types (10,000 tumours). The immunologically active (inflammatory and IFNγ dominant) tumours displayed an improved prognosis over the immunologically dormant (immunologically quiet and lymphocyte depleted). Key immunological features (e.g. cell markers CD4 and CD8A; functional markers GZMB and IFNG) have been used in a multitude of studies to predict for survival (Bindea et al., 2013, Rooney et al., 2015, Galon et al., 2006).

This highlights that a core of immunological analytes can be used to stratify patients into different outcome groups. An alternative dimension that can be assessed uses the addition of markers that are predictive of poor prognosis such as stromal cells, cancer associated fibroblasts (CAFs) and glycolytic stress markers. Elevated levels of CAFs and changes in glycolytic activity in tumours have been linked to a poorer prognosis (Hanley et al., 2018, Ottensmeier et al., 2016b). The question of whether a limited number of analytes may be able to stratify a patient cohort is important, especially in light of the number of immunotherapy trials that are on-going (Zhang and Chen, 2018). Stratification of patients based on a subset of genes, provides a way of selecting patients who are most likely to benefit from a particular immunotherapy. IHC is the most widely accepted method of determining immunological features e.g. CD8, PD1, PDL1, and this fits in with routine histological classifications of tumours. Alternative approaches that use one or a combination of techniques such as qPCR, RNA-Seq and flow cytometry, may allow more rapid immunological stratification of patients and the targeted selection of immunomodulatory therapies (Fehrenbacher et al., 2016), (Dobbin et al., 2016) (Hartmann et al., 2018) (Takahashi et al., 2019). The advantage of qPCR over RNA-Seq is that it allows for a lower number of samples to be run per batch, as well as enabling a faster turnaround with the assay taking less than 24 hours to complete from samples to result. The Fluidigm Biomark HD platform provides a flexible microfluidics approach to qPCR that enables reproducible large-scale experiments, the 24 sample x 192 analytes allows 4,680 qPCR reactions to be completed simultaneously. The immuno-

oncology panel provides information on 165 genes across 24 samples (Fehrenbacher et al., 2016). Other platforms such as the Nanostring and EdgeSeq also offer immuno-oncology panels and can work from FFPE and whole tissue (Fehrenbacher et al., 2016), (Dobbin et al., 2016) (Hartmann et al., 2018) (Takahashi et al., 2019). New clinical trials in the design phase are aiming to deliver personalised and targeted therapies based on information from the patient's tumour. This presents a variety of challenges that have an impact on the ability to deliver rapid clinically relevant data without interfering with the normal routine diagnostic procedures.

3.3.2 Objectives

To determine the immunological finger print in HNSCC (head and neck cancer) and NSCLC (lung cancer) RNA-Seq data using a set of key immunological and tumour microenvironment genes. To evaluate a rapid qPCR gene expression assay for tumour samples, that covers the key immune targets observed in the RNA-Seq data.

3.3.3 Results

3.3.3.1 Immunological and tumour microenvironment features in Head and neck (HNSCC) and Lung cancer (NSCLC).

Immunological and tumour microenvironment gene markers (see Table Appendix B.5.3) were visualised as heatmaps (Figure 16, Figure 17 and Figure 18) using RNA-Seq data from our own HNSCC cohort, CGA HNSCC and CGA NSCLC cohorts (Cancer Genome Atlas Research, 2014) (Cancer Genome Atlas, 2015) (Campbell et al., 2016). The data reveals an immunological signature across the different cohorts, ranging from immune "cold" to immune "hot" tumours, this is effectively captured using the selected Immunological gene markers. The samples from our own HNSCC data cluster mostly by immune density as determined by H and E TIL status (TIL^{Lo}, ^{Mod} and ^{Hi}), which matches the immune gene expression. An exception to this, is one TIL^{Lo} tumour that clusters with the TIL^{Mod} and ^{Hi} cases (Figure 16). The HNSCC and NSCLC CGA RNA-Seq data (Figure 17 and Figure 18) can again be clearly grouped by immune density, the data has been manually annotated into TIL^{Hi}, ^{Mod} and ^{Lo} (green = Lo; pink = Mod and yellow = Hi) based on the immunological gene expression. Specific targets of immunotherapy like CD8+ T cells , PD1 (PDCD1) and CTLA4 can also be grouped into Hi and Lo for the purposes of stratifying patients. The expression of tumour microenvironment genes (CAF and glycolysis markers) was higher in the immune "cold" tumours (TIL^{Lo}). However, the relationship was not as clear as the immune genes, with groups of samples displaying higher CAF expression in the immune "hot". The glycolytic gene markers also showed a pattern of lower expression in a group of immune "cold" tumours.

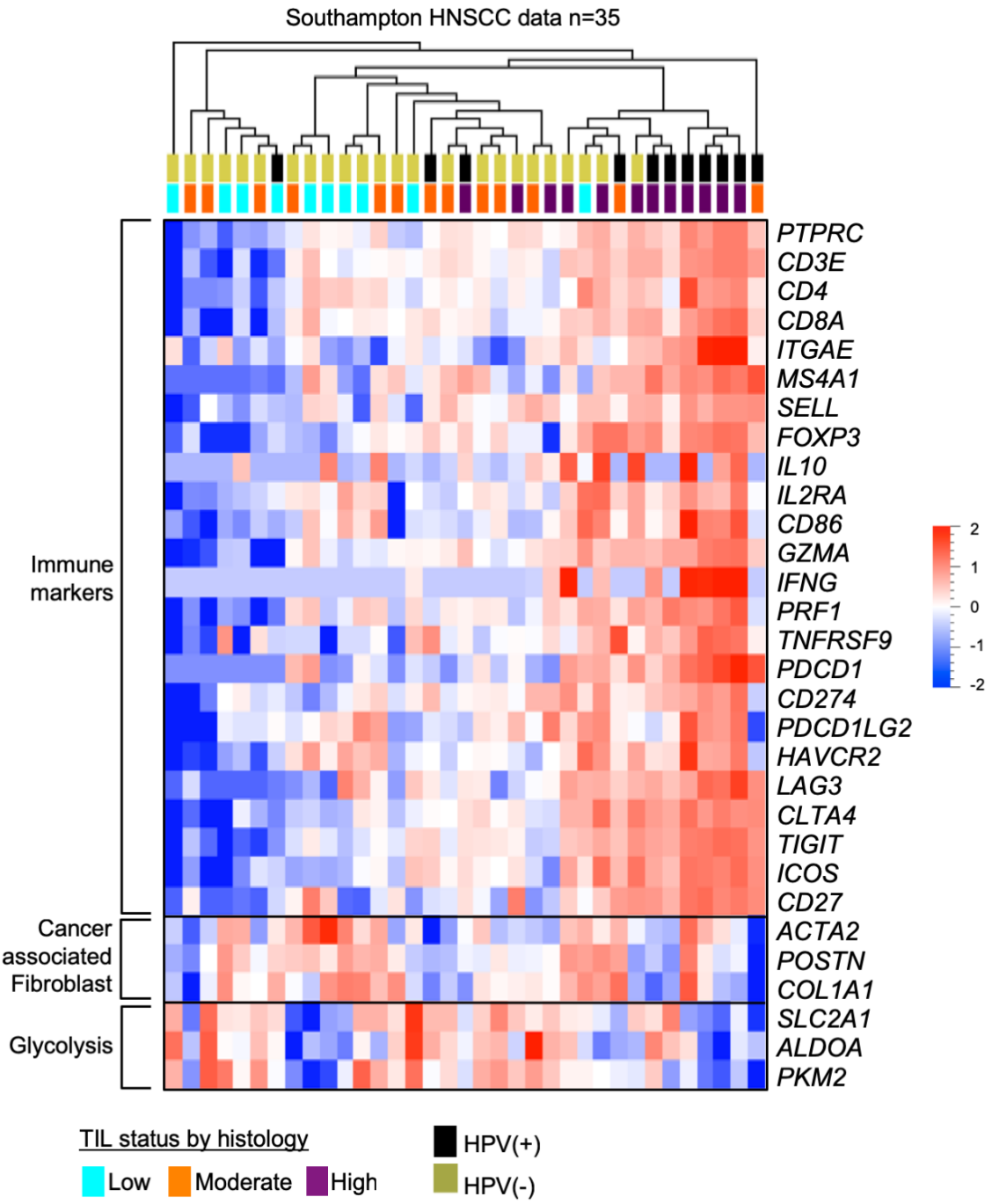


Figure 16. Gene expression of immune markers in TIL stratified HNSCC.

A heatmap to illustrate the gene expression profiles for immune and tumour microenvironment markers in HNSCC tumours. The immune gene markers were visualised in our own RNA-Seq cohort n=35, where tumours are stratified by TIL (low, moderate and high) using histological assessment (haematoxylin and eosin) and are hierarchically clustered by TIL in most cases. Visualisation of the same genes in the HSNCC CGA (cancer genome atlas) data n=518 where a clear TIL high and low gene expression is seen allowing TIL categorisation based on gene expression.

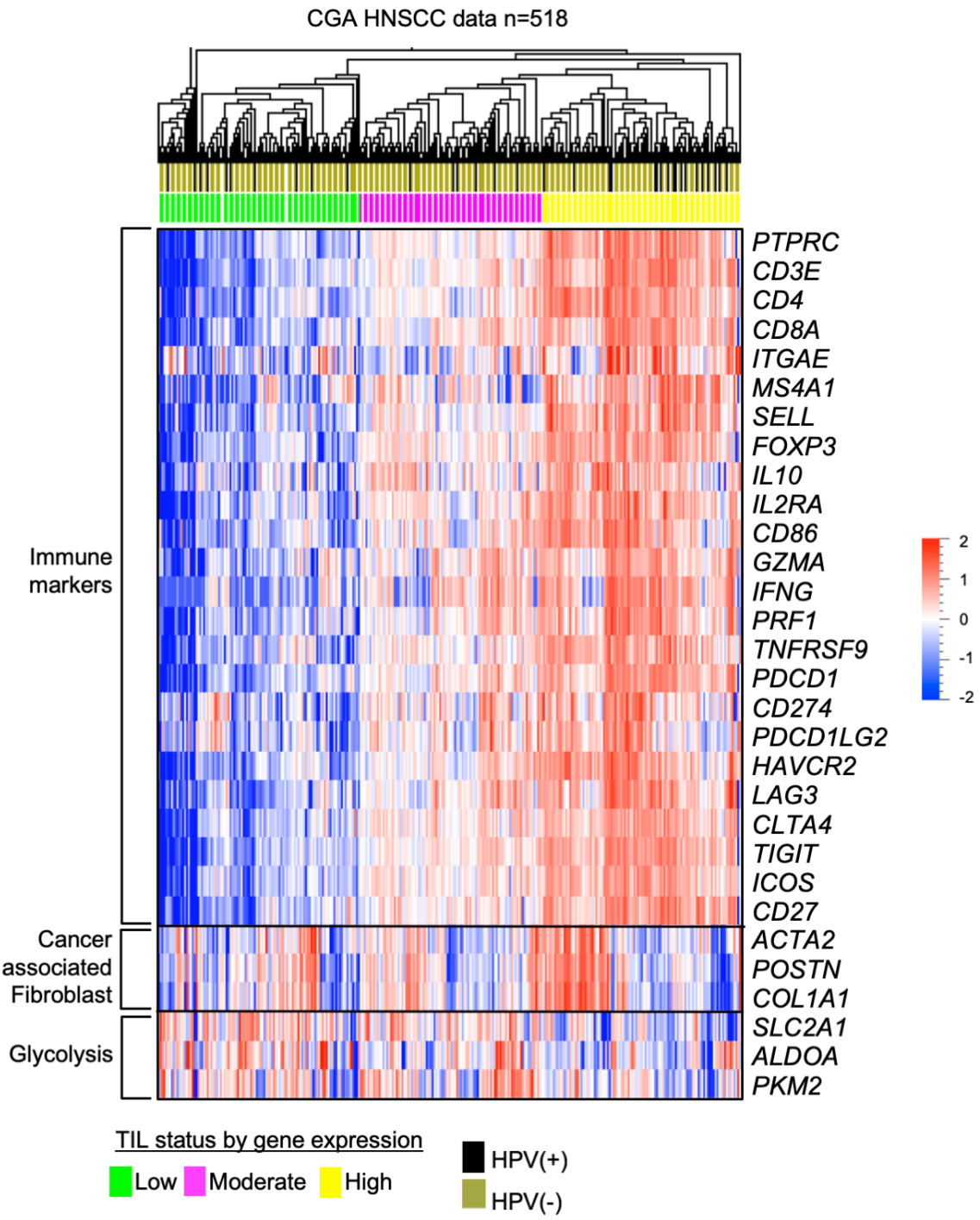


Figure 17. Gene expression of immune markers in HNSCC CGA RNA-Seq data.

Heatmap visualisation of genes associated with immune and tumour microenvironment markers in the HSNCC CGA (cancer genome atlas) data n=518. A clear TIL high and low gene expression pattern is seen allowing TIL categorisation based on gene expression in whole tumour samples (data sourced from (Cancer Genome Atlas Research, 2014)

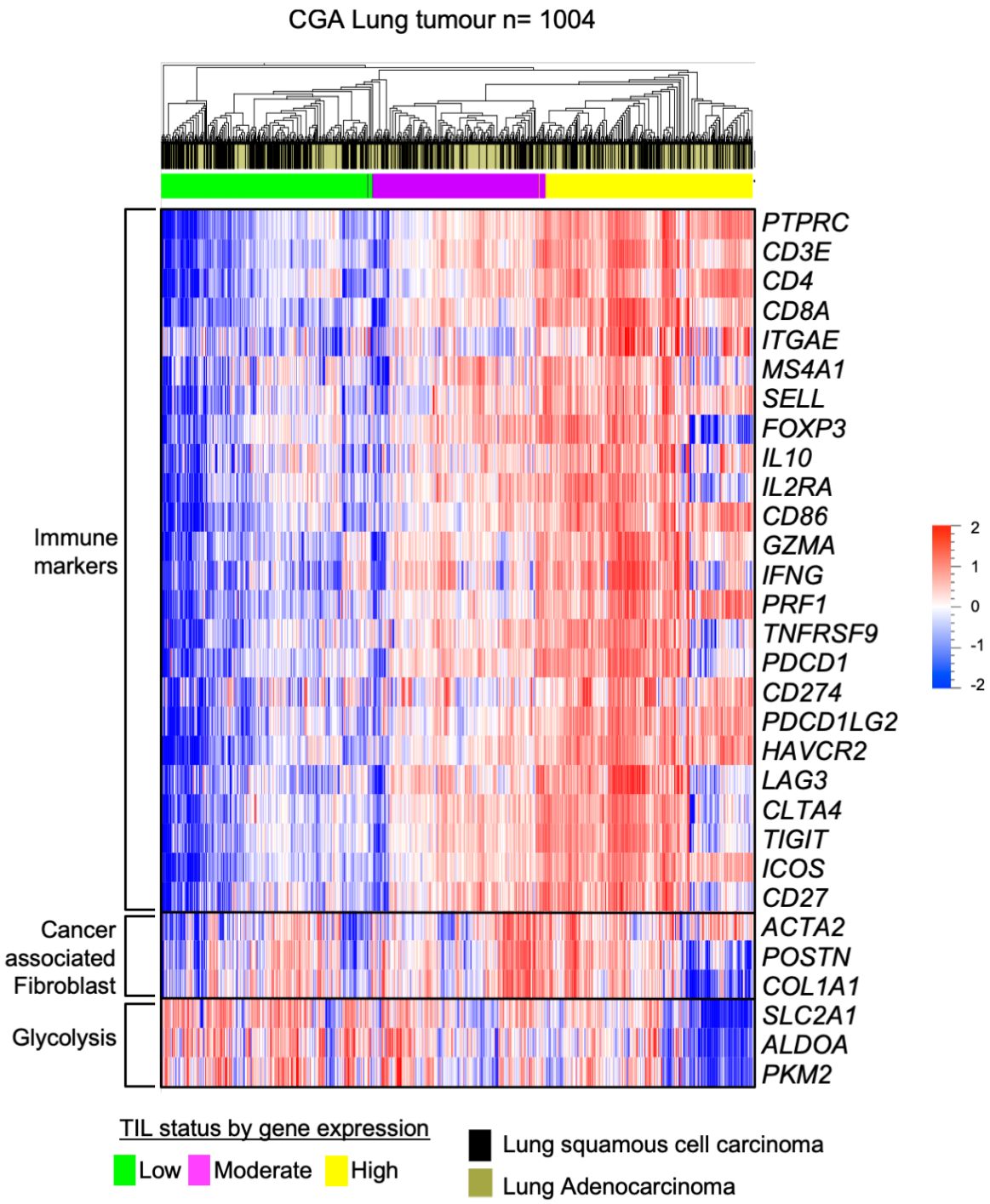


Figure 18. Gene expression of immune markers in Lung NSCLC CGA RNA-Seq data.

Heatmap visualisation of genes associated with immune and tumour microenvironment markers in the Lung CGA (cancer genome atlas) data n=1004, again a clear grouping of cases by immune gene expression can be observed (data sourced from (Cancer Genome Atlas, 2015) (Campbell et al., 2016).

3.3.3.2 Rapid stratification of patients by gene expression profiling and translation to clinical practice.

The use of gene expression profiling for specific immune gene signatures divides patient cohorts into immune Lo, Mod and Hi as previously identified. This could be highly beneficial information for patients eligible for immunotherapy. The potential stratification process is outlined in Figure 19, where patient samples are received from surgery and are processed as per the laboratory procedures with two key samples for rapid analytics, RNAlater and cryopreserved tumour tissue. The aim of this would be to deliver rapid (<5days) immunological data on key immunotherapy targets such as PD1 and CTLA4.

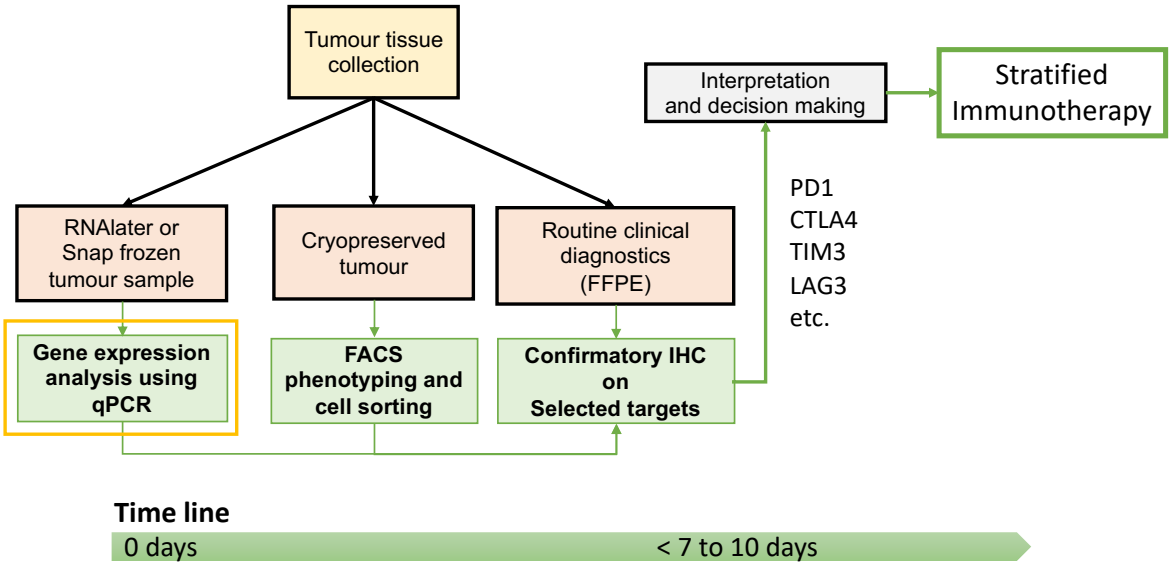


Figure 19. Patient stratification work flow.
Potential work flow that could be used to deliver immunological readouts in a time frame that would allow rapid selection of targeted therapies. qPCR would be used as an initial rapid immune readout, followed by flow cytometry and routine immunohistochemistry to complement the normal diagnostic reporting procedures.

The immuno-oncology panel (Fluidigm) provides information on 165 genes across 24 samples. This has been used on a subset of samples that were analysed using RNA-seq and cover a range of immune densities, firstly to assess applicability of the assay to deliver rapid immune information and secondly the correlation of the two methods. The genes from the Immuno-oncology panels A (91 genes) and B (74 genes) were visualised as a heatmap using our HNSCC cohort (n=35 samples) RNA-Seq (HPVN dataset) (See Appendix B.5.4 A-C). The samples are ordered by the first principle component (highest variance) and show a hierarchical clustering of genes linked to immune density (top half). The bottom half represent the oncology gene markers with heterogeneous expression across the 35 samples.

A subset of these samples (TIL^{Lo} n=3; TIL^{Mod} n=4 and TIL^{Hi} n= 3) were selected to test the qPCR assay (immune-oncology panel, Fluidigm) and determine a range of qPCR values for immune "hot" and "cold" tumours. The qPCR assay was performed in duplicate and included a positive control for all genes as well as a PBMC control RNA, the data was analysed in relation to two endogenous control genes B2M and ACTB. The relative gene expression for each gene (ΔΔct) was calculated relative to the positive control (pc), in order to obtain a positive value the ΔΔct was subtracted from the number of thermal cycles (n=40) ($\Delta\Delta Ct(GOI) = 40-((Ct[GOI]_{sample} - meanCt[REF]_{sample}) - (Ct[GOI]_{pc} - meanCt[REF]_{pc}))$). The relative qPCR values for the 165 genes ranged from 23 to 40 and were visualised as a heatmap (See Appendix B.5.4 B) using the mean = 0 and variance = 1 (z-score). The duplicate data points are visually comparable, the positive control shows high expression across all analytes. The Immune density across the samples wasn't apparent due to skewing of the z-scores from the high expression observed in the PBMC (control RNA) and positive control, this led to the apparent low expression of all genes in the test cohort. The PBMC and positive control samples were removed and the data was re-visualised in Appendix B.5.4 C, this allowed the range of gene expression within the tumours to become more apparent with the immune "hot" and "cold" tumours being more visible across the cohort.
The number of genes was then reduced to 14 key immune genes covering immunotherapy targets and cell markers (Figure 20).

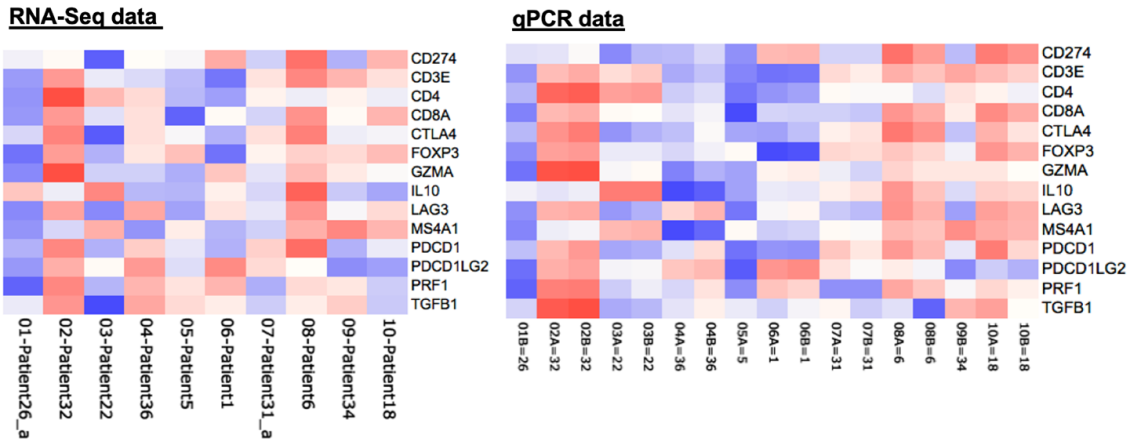


Figure 20. Comparison of Immune gene expression in the Southampton HNSCC RNA-Seq data and qPCR data (Bookmark HD).
The cohort of 10 samples with existing RNA-Seq data were compared to the qPCR platform, it demonstrates a high level of consistency between the RNA-Seq and the IO (Immuno-oncology) qPCR panel (Fluidigm) on whole tumour RNA. Heatmap data is shown for 14 key immune genes and cell markers across the two methods.

The immune gene expression profiles can easily be determined and are consistent between the RNA-Seq and the qPCR assays with the immune rich tumours displaying higher expression of the key immune cell markers (CD3E, CD8A, CD4). In order to visually compare the expression profiles in tumour samples between the two methods they were aligned side by side (Figure 21 A), the two methods yield excellent concordance with very few samples and markers not showing comparable data e.g Patient 18 = IL10. This was quantified further by assessing Spearman correlation coefficients on the 12 genes to see whether a significant correlation was observed, the correlation coefficients range from 0.299 to 0.891 and are all significantly ($p=<0.05$) positively correlated (Figure 21 B).

A key part of delivering gene expression data for stratification of patients will be defining reference ranges for each gene of interest, an example of this is shown in Figure 22. Where even in a small subset of 10 HNSCC samples high and low expression in CD8A, PDCD1 and CTLA4 can be observed and possible cut-offs can be assigned to define expression ranges.

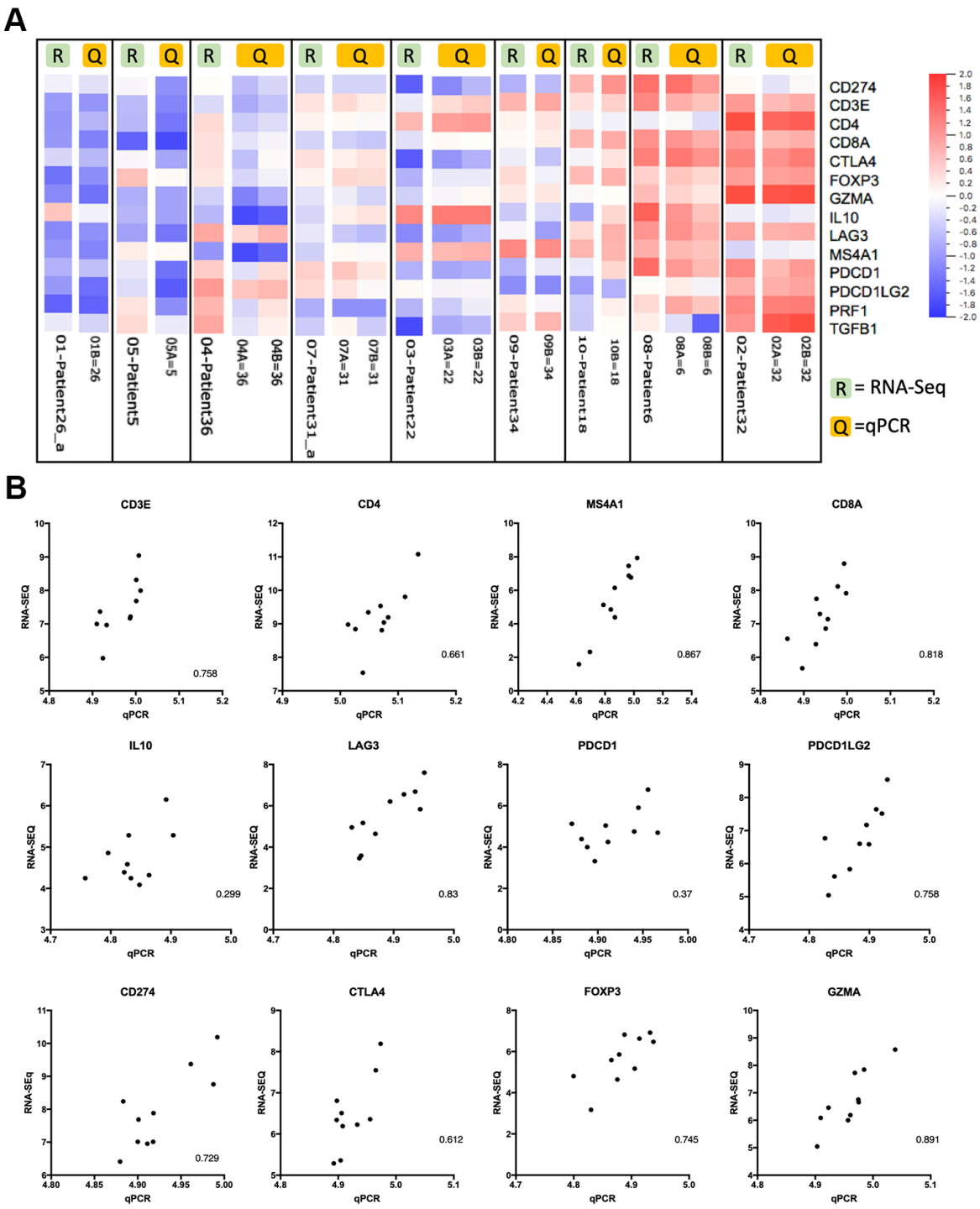


Figure 21.Comparison of immune gene expression between RNA-Seq and qPCR using the IO panel (Fluidigm).

(A) Displays heatmap data of 14 key immune genes and cell markers side by side for the two methods. A high level of visual consistency can be seen between RNA-Seq and qPCR for the expression of these genes. (B) Quantification of this using spearman correlation coefficients demonstrated significant positive correlations ranging from r = 0.299 to 0.891 for these targets.

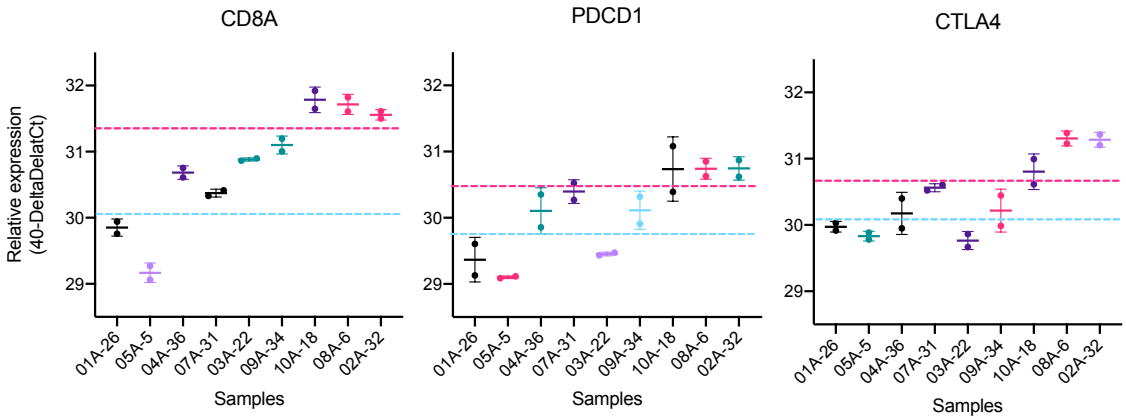


Figure 22. Hypothetical reference ranges for the allocation of immune gene density using qPCR. The relative gene expression for CD8A, PD1 (PDCD1) and CTLA4 presented as dot plots indicate that even within a small cohort of samples categorisation into high and low can be achieved. It demonstrates that each target would need to be validated and relative expression cut-offs would be target specific.

3.3.4 Discussion

Transcriptomic and genomic profiling of tumours across different subtypes (cancer genome atlas) has led to a better understanding of tumour biology. A pan-cancer analysis of immune subtypes in different cancers revealed that particular immune subtypes are linked to a favourable outcome (Thorsson et al., 2018). The data presented here focuses on a selected number of immune genes that capture the overall immune density of the tumours. The immune gene expression profiles in HNSCC (HPV[+] and [-]) and NSCLC (Adenocarcinoma and Squamous cell carcinoma) were able to group the cohorts into immune “hot” and “cold”. The ability to group patients by immune density is attractive clinically as it allows stratification of patients into better survival groups, as well as enabling a targeted approach to immunotherapy where immune “cold” PD1 low tumours could be targeted with alternative treatments such as a NOX4 inhibitor that aim to modulate the tumour stroma (Hanley et al., 2018). More effective patient selection has the potential to transform the effectiveness of immunotherapies. By choosing patients that are likely to respond (immune ‘hot’), the response rates could massively be improved (currently 20-30%). It would also spare patients that are unlikely to respond from the adverse of effects immunotherapy.

In this evaluation transcriptomics was able to accurately capture the immune contexture of tumours using RNA-Seq, however there are a variety of issues that present themselves when trying to deliver this approach within a clinically relevant time frame. Batching of samples and batch effects in RNA-Seq data is a known issue when trying to compare different experiments (Leek et al., 2012). The need to normalise the data is also problematic, especially if the sample cohort doesn’t accurately capture the immunological extremes (immune “hot” and “cold”). The

time for an RNA-Seq run to complete is highly variable depending on the system being used and/ or throughput required, often needing multiple sequencing libraries to be run simultaneously. An alternative approach to this is a more targeted approach using qPCR and Taqman® assays for specific gene targets, this enables higher sample turnover at lower costs and improved batching. A pilot study was carried on 10 HNSCC patients with a range of immune densities using an immuno-oncology panel. The aim was to investigate if the qPCR and RNA-Seq data could both stratify the samples by expression of immune genes. The data generated from the qPCR assay displayed an overall good level of similarity between RNA-Seq and qPCR, a significant positive correlation (Spearman) was observed for 12 immune markers. The data was also able to capture the overall immune density of the samples as previously observed in the RNA-Seq data (TIL^{Hi}, ^{Mod} and ^{Lo}). In order for the qPCR assay to be useful clinically an extended set of well curated samples covering a range of tumour subtypes would need to be tested, in addition to this a reference range from Hi to Lo would need to be made, similar to the approach used in the Mammatyper (Hartmann et al., 2018). The application of qPCR enables extremely rapid detection of immune targets on very small biopsies, it opens up the use of needle aspirates and core biopsies for the management and monitoring of later stage of disease (in-operable). This ultimately enables treatment decisions to be adapted over time, and in response to changes within the tumour.

The data here displays a good level of immune variability with an immune Lo and Hi signature present, obviously without the extremes it would be difficult to determine an accurate reference range for specific targets.. Each specific immunotherapy target would need to be validated and an appropriate reference range set. The current analysis used a positive control as a reference point to generate the ΔΔCt(GOI), future analysis would need to use an RNA sample of know expression as reference point for each experiment to minimise run to run variability (Agilent and Clonetech qPCR standards). The caveat with all whole tumour analysis is that the cell of origin for a particular signal is not known. It has however been shown that the overall gene expression data maps well onto clinical response for PD1 (Thommen et al., 2018). Even within the limited number of samples tested here the samples can be stratified into high and low expression of key immunological genes (PDCD1 and CTLA4).

3.3.5 Conclusions

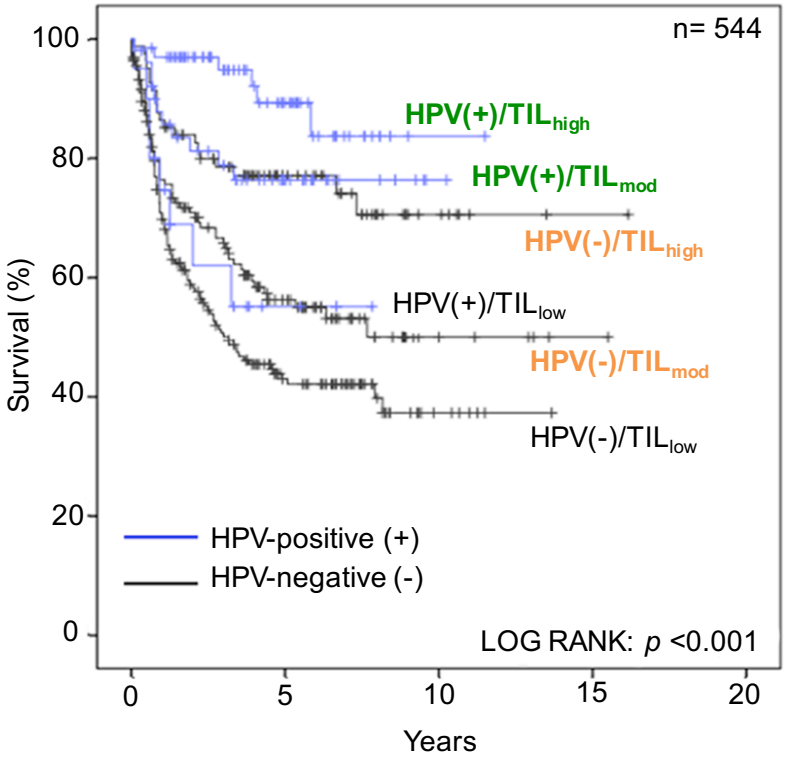
A select number of immune genes were able to accurately capture the immune context of different tumours, this provides an interesting avenue for patient stratification in the clinical setting. Where therapeutic options can be tailored to specific patients based on detailed immunological information, provided it can be delivered in a timely fashion. A pilot study using

qPCR shows that it can be used in a clinically relevant time frame (results in <24 hours) and is comparable to whole tumour RNA-Seq. Further work and validation on larger cohorts would need to be performed to establish an effective work flow and define reference ranges for the immune genes.

3.4 Transcriptomic analysis of TIL rich HNSCC

3.4.1 Introduction

The worldwide prevalence of HPV-associated HNSCC is approximately 30% (Chaturvedi et al., 2011, Attner et al., 2010, Ward et al., 2014b), of which the majority are caused by HPV16. The anatomical location linked to HPV-driven HNSCC is the oropharynx, which includes the base of tongue and tonsil (Attner et al., 2010). HPV positive (HPV(+)) patients have a significantly better prognosis than HPV negative (HPV(-)) patients, with the 3- and 5-year survival at 84% and 62% for HPV(+) patients compared to 57% and 26% for HPV(-) patients, respectively (Ang et al., 2010).
A high number of tumour-infiltrating lymphocytes (TILs) is linked to good prognosis in many solid tumours (Galon et al., 2006). More recent analyses of The Cancer Genome Atlas (TCGA) data demonstrate that the effect is mediated by CD8+ GZMA+ PRF1+ T-cells (Rooney et al., 2015, Cancer Genome Atlas, 2015). In HPV(+) disease, the persistent viral oncoproteins E6 and E7 cause the malignant phenotype, while the immunological visibility of these two proteins contributes to the infiltration of the tumour by T-cells (Heusinkveld et al., 2012). It has previously been demonstrated that in HPV(+) HNSCC, TIL density correlates with outcome (Ward et al., 2014b).
In contrast HPV(-) tumours are considered a separate disease entity and are driven by heterogeneous genetic events, they also display poorer survival relative to HPV(+) tumours (Cancer Genome Atlas, 2015, Stransky et al., 2011).
Differential gene expression profiling comparing HPV(+) and HPV(-) tumours using microarray, RNA-Sequencing (RNA-Seq) and RT-PCR, have led to an improved understanding of the events associated with cellular transformation and oncogenesis (Jung et al., 2010, Pyeon et al., 2007, Slebos et al., 2006, Russell et al., 2013). Thurlow et al. used spectral clustering and gene ontology (GO) analysis to identify discrete gene expression patterns that linked to patient outcome, these involved the genes of adaptive and innate immunity. However, the underlying biology of TILs had not been addressed despite their link to survival in both HPV(+) and (-) HNSCC (Chung et al., 2004, Thurlow et al., 2010, Ward et al., 2014b, Cancer Genome Atlas, 2015, Jung et al., 2010, Pyeon et al., 2007, Slebos et al., 2006).
A previous cohort of HNSCC cases that focused on HPV(+) disease (Ward et al., 2014b) was re-assessed with the inclusion of the HPV(-) cases (total n=544), as in the HPV(+) tumours, TIL status stratified for outcome in HPV(-) HNSCC (Figure 23). Furthermore, HPV(+) $TIL^{Hi/mod}$ patients had significantly better survival compared to HPV(-) $TIL^{Hi/mod}$ patients (log rank $p<0.001$). The survival difference between HPV(+) and HPV(-) TIL rich tumours is interesting as it could arise from qualitative difference in the immune response or be due to quantitative differences in immune cells.



HPV(+)/TIL$_{high}$	61	26	1	
HPV(+)/TIL$_{mod}$	51	21	1	
HPV(+)/TIL$_{low}$	20	4		
HPV(-)/TIL$_{high}$	81	33	7	1
HPV(-)/TIL$_{mod}$	134	47	6	1
HPV(-)/TIL$_{low}$	192	49	6	

Figure 23. Kaplan-Meier curves for HNSCC mortality stratified according to HPV status and TIL density.

Survival of a retrospective cohort of HNSCC patients (n=544) with respect to HPV status and the density of immune cell infiltrate. TIL density predicts for outcome in both the HPV(+) and HPV(-) patients; log-rank test, p<0.001. Adapted from (Ward et al., 2014b).

3.4.2 Objectives

To interrogate the survival differences between HPV(+) TIL$^{Hi/mod}$ and HPV(-) TIL$^{Hi/mod}$ tumours in a cohort of consecutive HNSCC surgically resected cases, with a particular focus on the immune gene expression profiles. This will be done using high resolution transcriptomic analysis (RNA-Seq), accompanied by histological assessment to accurately determine immune cell density. The presence of HPV driving the tumour may generate a virus-specific immune response, or just lead to an increase in immune cell density compared to that of HPV independent tumours, which may account for the survival differences.

3.4.3 Results

3.4.3.1 Clinical information for the HNSCC cohort (HPV(+)TIL$^{Hi/mod}$ compared to HPV(-) TIL$^{Hi/mod}$)

In order to investigate the survival difference between HPV(+) and (-) tumours the transcriptome was evaluated in a prospective cohort (n=39), patient demographics, tumour characteristics and tumour sampling/ processing information for the HPV(+) and HPV(-) patient cohort are shown in Appendix A.1.2. The focus for this analysis was to understand immune cells in the HPV(+) and (-) setting, we focussed on TIL rich (TIL$^{Hi/mod}$) tumours and excluded the 16 TIL^{low} cases from the current analysis. The cohort consisted of 23 TIL$^{Hi/mod}$ cases, HPV(+) n=10 and HPV(-) n=13. The clinical and histological descriptors were distributed as expected (Appendix A.1.2); HPV(+) tumours were located in the oropharynx of non-smoking patients, HPV(-) tumours were located in the larynx (n=4), oral cavity (n=5) and oropharynx (n=4). The HPV(-) cases consisted of n=7 smokers and n=6 non-smokers; n=2 samples from the oropharynx were from smokers. The clinical classification of HPV status was determined by routine IHC against p16 and mapped appropriately to the expression of E6 and E7, as detected by RNA-Seq.

3.4.3.2 Differentially expressed genes between HPV(+) and HPV(-) tumours (TIL$^{Hi/Mod}$)

Differential gene analysis was performed to assess the differences between HPV(+) and HPV(-) tumours (TIL$^{high/mod}$), EdgeR was used and identified 1634 genes as significantly differentially expressed (q-value <0.05) between the two groups; these were visualised in a heatmap (Figure 24 A). Of these genes, 894 were expressed to a greater extent and 740 to a lesser extent in HPV(+) compared to HPV(-) tumours. These gene expression differences segregated HPV(+) and HPV(-) tumours in all except one HPV(-) subject (patient 21), whose tumour clustered within the HPV(+) cohort (Figure 24 A). A review of the histological data revealed this patient had a basaloid SCC, which is a rare and clinically distinct form of HNSCC. The same genes (1634) were also visualised in the larger CGA HNSCC dataset (Cancer genome atlas: TCGA HNSC HiSeqV2 2015-02-24(Cancer Genome Atlas, 2015); data source outlined in methods section 2.7) (Figure 24 B). The same separation of HNSCC into HPV(+) and (-) was observed in that dataset, however due to the limited clinical annotations (TIL status not available) the same differential gene analysis could not be carried out in EdgeR. The relative expression of key genes linked to HPV were also visualised in box and whisker plots (Figure 25), gene expression of the HPV-associated genes CDKN2A (p16), *E6* and *E7* were as expected between HPV(+) and HPV(-) tumours. We also found differences in the expression of genes associated with 'immune cell markers' between HPV(+) and HPV(-) tumours (Figure 25); expression of these genes was much greater in HPV(+) tumours. Similarly, the expression of GZMA, IFNG and CDNK2A and key genes that link to T-cell activation and

exhaustion, such as CTLA4, PD1 and HAVCR2 (encoding TIM3), were all increased in HPV(+) compared to HPV(-) tumour.

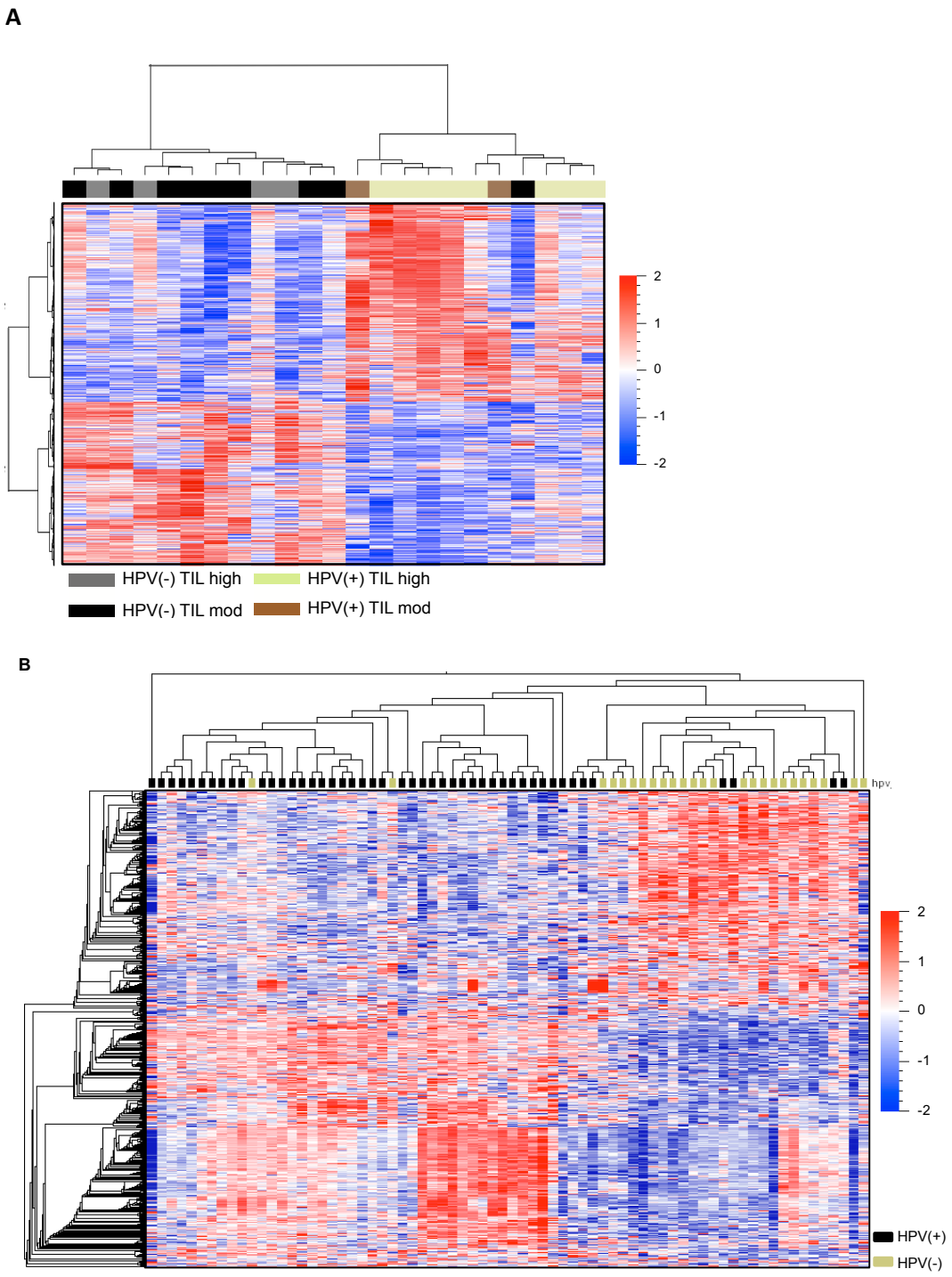


Figure 24. Differentially expressed genes between HPV(+) and HPV(-) immune rich tumours. (A) A heatmap to illustrate the DEGs between HPV(+) and HPV(-) tumours; showing row-wise z-scores of normalized read counts; the scale of z-score is shown: red shading denotes greater gene expression, blue shading denotes lower gene expression. Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of genes and tumours based on their expression profile is reflected in the dendrograms to the left and the top of the heatmap, respectively. (B) Expression of the same DEGs in the CGA HNSCC data, HPV(+) and (-) cohorts cluster separately and show similar expression of the identified genes.

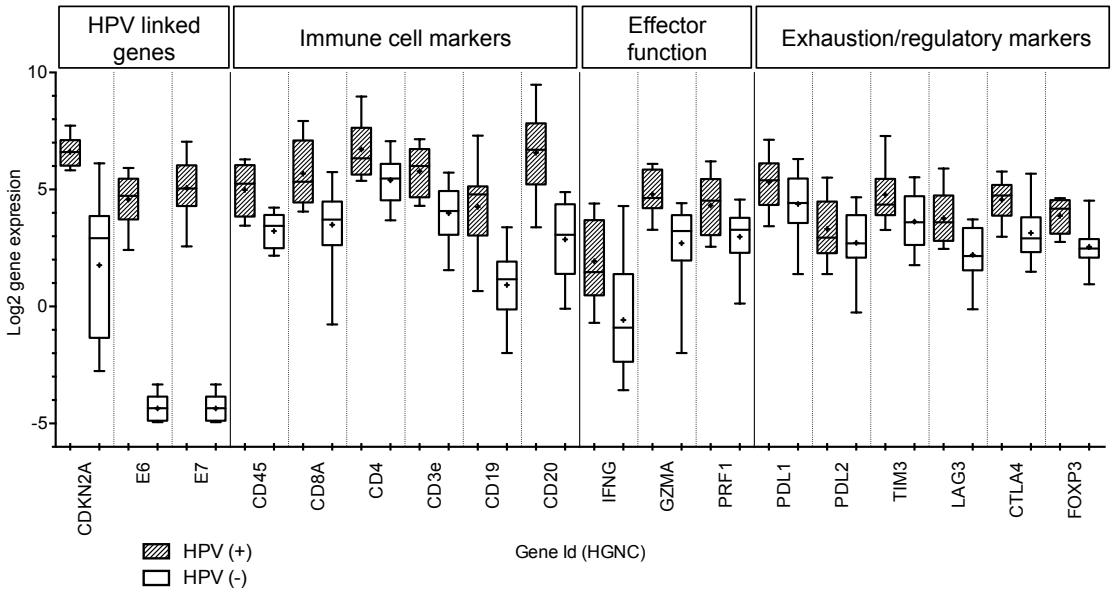


Figure 25. Expression of key genes in HPV(+) and (-) tumours

The expression of key genes associated with HPV, immune cell markers, immune effector function and immune exhaustion/regulation are displayed, for both HPV(+) and HPV(-) tumours box plots show the median and min/max with the + representing the mean. A greater expression of immune associated genes is observed in HPV(+) tumours (gene expression data from normalized transcript counts).

3.4.3.2.1 Gene ontology (GO) and pathway analysis:

GO and pathway analysis was performed to understand the biological significance of the 1,634 DEGs. GO terms that were significantly over-represented for DEGs were identified using CPDB (Kamburov et al., 2011); GO analysis was performed independently for those genes expressed to a greater extent and for genes expressed to a lesser extent in HPV(+) compared to HPV(-) tumours. Full details of the GO terms, including the specific genes, number and percentage of DEGs associated with each GO term, are presented in Appendices B2.1 to B2.5. The GO terms were visualised using REVIGO scatterplots (Appendix B2.1), where a 2D cluster representation shows the most significant terms after redundancy reduction. The data reveal that those genes with greater expression in the HPV(+) cohort were predominantly associated with an immune reaction (e.g., adaptive immune response, GO:0002250; lymphocyte activation, GO:0046649; positive regulation of immune system process, GO:0002684; B-cell activation, GO:0042113), whereas those expressed to a lesser extent were associated with cellular processes involved in tissue development, (GO:0009888), keratinization (GO:0031424) and cell differentiation (GO:0030154). Specifically, there was greater expression of genes associated with the adaptive immune system, including T-cells (CD4+ and CD8+) and B-cell receptor signalling pathways, as well as NK-cell-mediated cytotoxicity. Those genes expressed to a lesser extent (Appendix B2.2) represented different biological processes, including extracellular matrix organisation, collagen formation,

beta1 integrin cell surface interactions and alpha6 beta4 integrin-ligand interactions. The biological pathways (e.g., *KEGG*) over-represented for DEGs mirrored the results of GO analysis displaying an enriched number of genes linked to immunological signalling pathways in greater expressed genes and pathways linked to extracellular matrix organisation and collagen formation in genes expressed to a lesser extent Appendix B2.5.

3.4.3.3 Quantification of TILs and evolution of the analysis:

The GO and pathway analysis indicated that the most differentially expressed biological processes were related to the immune system and specifically to B and T-cells in the genes expressed to a greater extent in HPV(+) tumours. However, it was not clear whether these B- and T-cell immune-related terms identified by GO and pathway analysis were simply the result of numerical differences in lymphocytes between HPV(+) and HPV(-) tumours or the result of differences on a per cell basis in the transcriptional signature. In order to address this question, the tumour-infiltrating lymphocytes were quantified using IHC for CD8, CD20, CD4 and CD3 (representative images at 40x magnification shown in Figure 26) followed by manual counting of 10 high-power fields (Figure 27 A) and gene RNA transcript levels (log2 normalised) for lymphocyte cell surface markers (Figure 27 B). A significant difference in cell number was observed between HPV(+) and HPV(-) tumours for the cell markers CD4, CD20 and CD3 ($P \leq 0.05$) but not CD8 (ns) by IHC. Gene RNA transcript levels were significantly different for all cell markers (CD4, CD20, CD3E and CD8A). Spearman correlation coefficients between IHC and gene expression for the cell markers CD8 ($r = 0.76$), CD3 ($r = 0.52$), CD4 ($r = 0.34$) and CD19 ($r = 0.47$) are shown in Figure 27 C to F. Both analyses demonstrated that the numbers of B-cells and CD4+ and CD8+ T-cells were higher in HPV(+) tumours compared to HPV(-) tumours. As a result of this difference the number of immune cells as determined by the gene expression of CD4, CD8A and CD20 within the tumour from the RNA-Seq data were used to correct for the immune cell differences using EdgeR (function in EdgeR used for correction of batch effects and confounding variables). The immune cell correction was undertaken to try and establish if an immunological phenotypic difference between the tumours still existed, and whether it could account for the difference in survival, or whether it was merely due to numerical differences in immune cells.

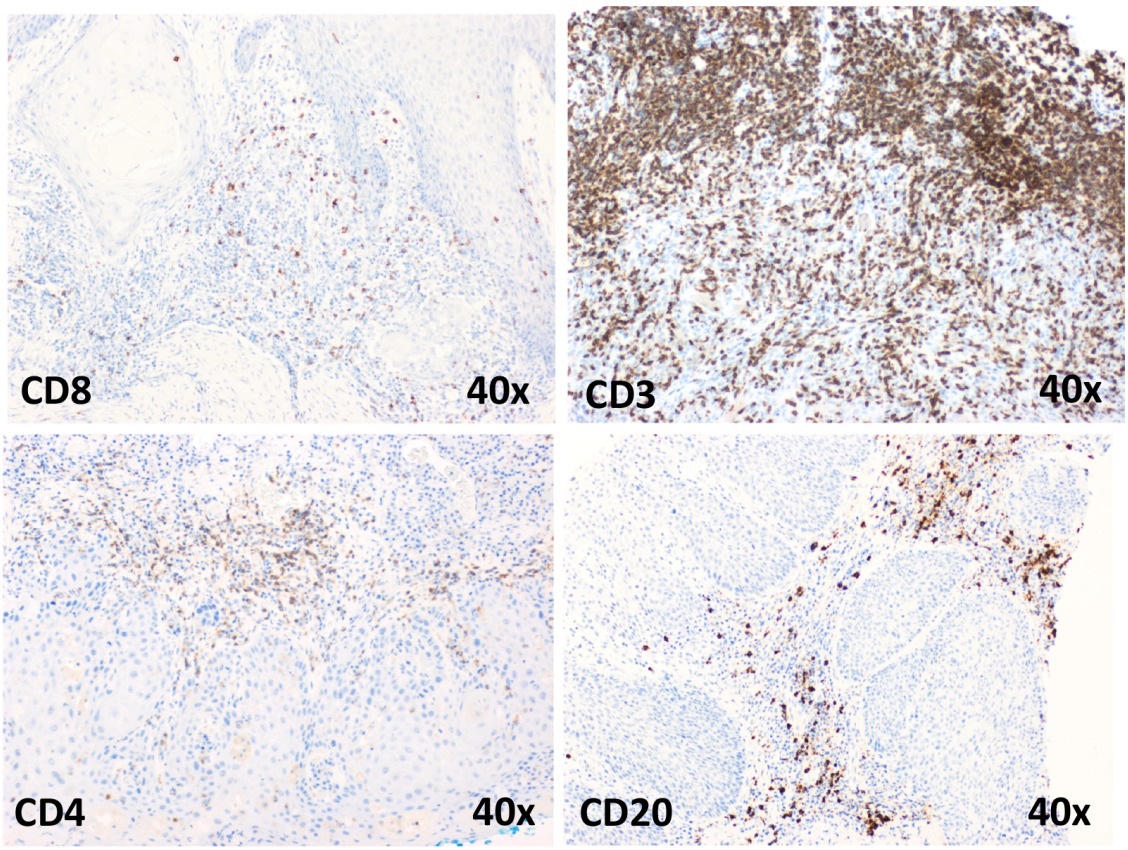


Figure 26. Representative images for tumour infiltrating CD8+ (TIL^{LMod}), CD4+ (TIL^{Mod}), CD3+ (TIL^{Hi}) and CD20+ (TIL^{Mod}) cells at 40x magnification.
The different tumours display highly variable immune infiltrate levels (IHC slides prepared by research histology – Monette Lopez and Maria Machado).

3.4.3.1 Analysis of gene expression data following correction for numerical differences in TILs

The global gene expression data were corrected for TIL number using the gene expression of CD19 (pan B-cell marker), CD4 and CD8A in each sample as a covariate (using EdgeR). When correcting both HPV(+) and HPV(-) cohorts in this way, genes co-ordinately expressed in lymphocyte subsets were no longer differentially expressed; as a result the number of DEGs dropped from 1,634 to 437 (Appendices B3.1). As would be expected, there was a large overlap in DEGs between the initial uncorrected and the TIL corrected data sets (Figure 28); from the immune cell correction, 77 new DEGS were identified in the greater expressed genes and 58 new DEGS in the lesser expressed. The TIL corrected dataset was next subject to GO and pathway analysis.

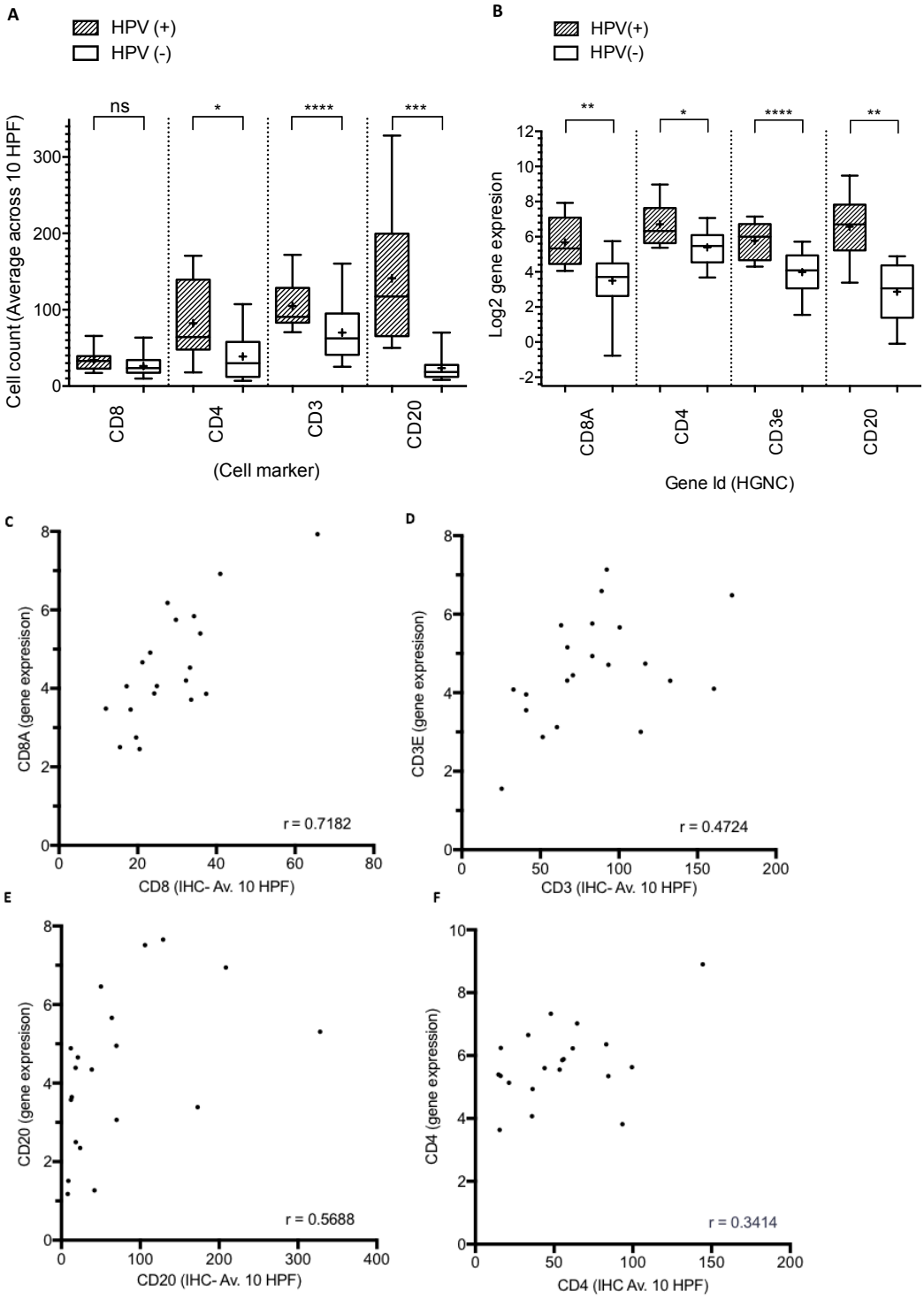


Figure 27. Immune cell subset analysis of HPV(+) and HPV(-) tumours.

(A) The distribution of CD8, CD4, CD3 and CD20-expressing cells in HPV(+) and HPV(-) tumours as detected by IHC; cell counts are given as a mean of 10 high-power fields. (B) Gene expression (normalized transcript counts) of CD8A, CD4, CD3E and CD20 of HPV(+) and HPV(-) tumours displayed as box plots (min/max) with the + representing the mean. Significant differences in TIL density between HPV(+) and HPV(-) tumours were observed both by gene expression profiling and IHC analysis. Spearman correlation analysis between immunohistochemistry and gene expression for Immune cell types. Correlation of IHC and gene expression determined by RNA-Seq for (C) CD8, (D) CD3, (E) CD20 and (F) CD4. A positive correlation was observed for each marker when counting 10 high power fields and correlating it with the level of immune gene transcripts determined by RNA-Seq (two group t-test = Mann-Whitney).

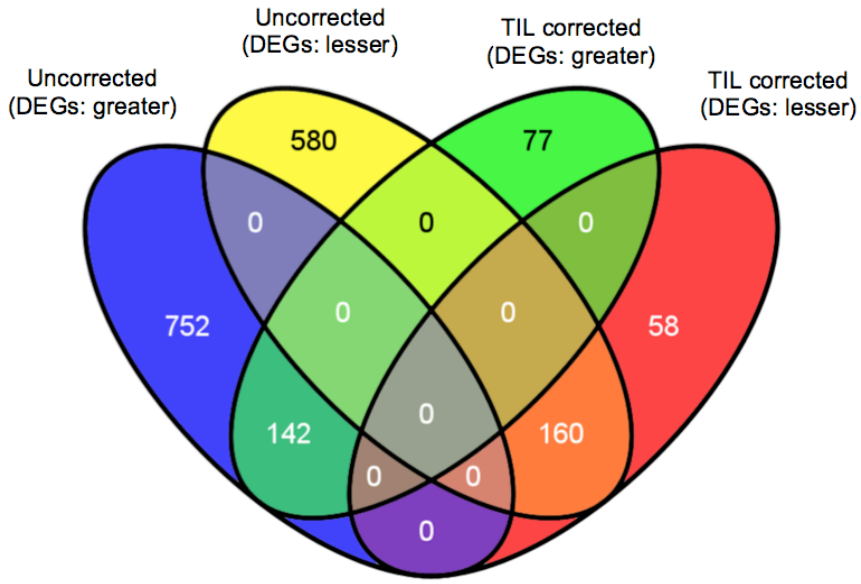


Figure 28. Overlap of DEGs identified from uncorrected and TIL corrected gene expression data. A Venn diagram to illustrate the overlap of DEGs between HPV(+) and HPV(-) tumours identified directly by RNA-Seq (n=1,634) or following correction of the gene expression data to account for numbers of infiltrating immune cells (n=467); genes expressed to a greater or lesser extent in HPV(+) versus HPV(-) tumours. For TIL corrected gene expression data, new DEGs were identified, which were expressed to a greater (n=77) or lesser extent (n=58) in HPV(+) tumours

3.4.3.2 Heatmap visualisation of immune cell corrected DEGs

Unsupervised hierarchical clustering of all 437 TIL-corrected DEGs for our own dataset and TCGA data (the HPV(-) tumours were anatomically matched; arising in the oropharynx, tonsil and base of tongue). In both datasets, tumours cluster according to HPV status; sub-clusters are evident in the larger TCGA dataset (anatomically matched) shown in Figure 29 A and B. The separation on the HPV(+) and (-) tumours is even more pronounced after correction in both ours and the CGA data. The removal of the immune cell confounder has the potential to reveal any immune differences and/ or transcriptomic features of the tumours; to investigate this further GO analysis and pathway analysis was performed.

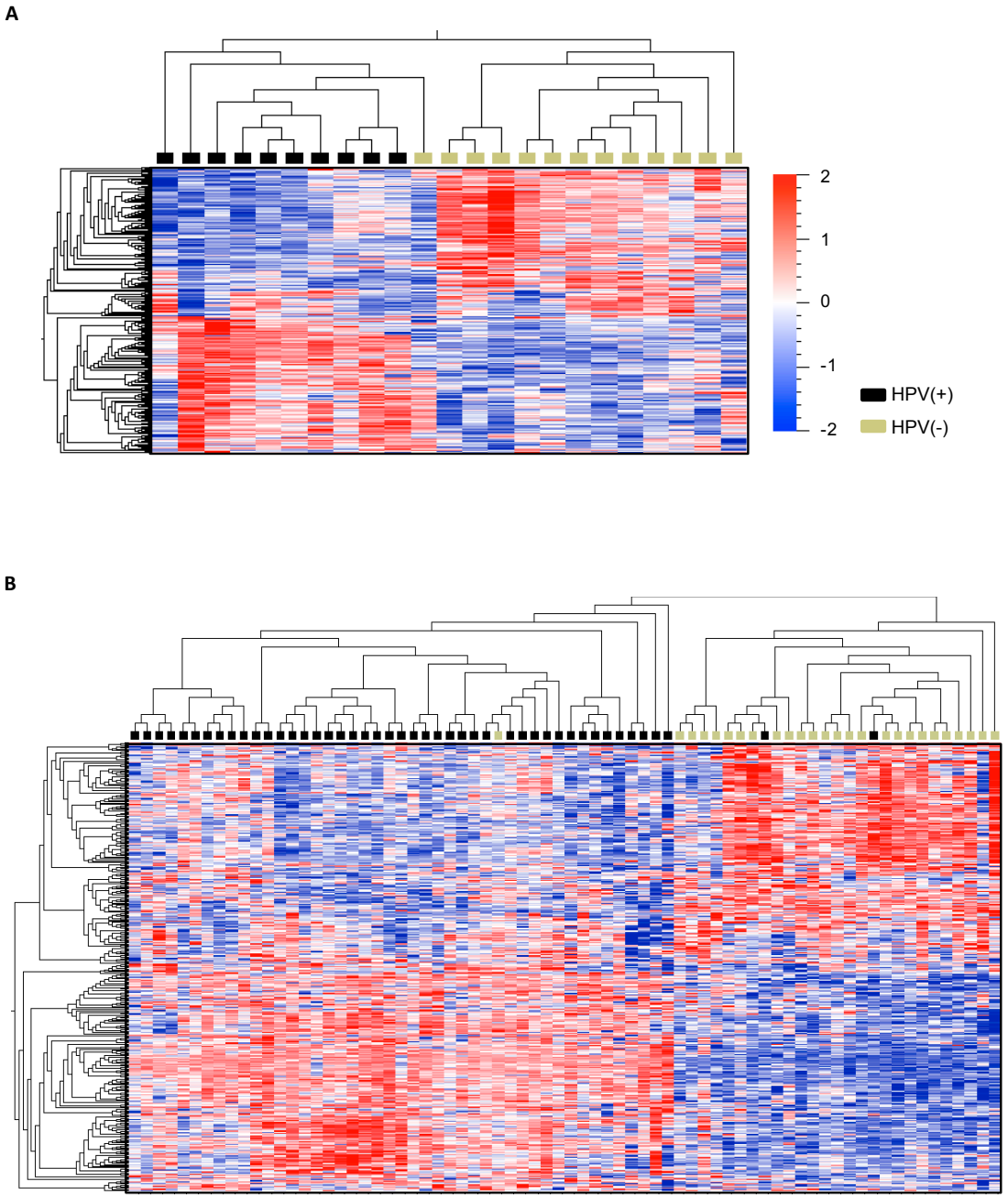


Figure 29. Visualisation of the TIL corrected DEGs between HPV(+) and (-) TIL$^{Hi/Mod}$ tumours.
Heatmaps to illustrate gene expression of the TIL corrected DEGs (n=437) between HPV(+) and (-) tumours showing row wise z-scores of normalized read counts; the scale of z-score is shown: red shading denotes greater gene expression, blue shading denotes lower gene expression. Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of genes and tumours based on their expression profile is reflected in the dendrograms to the left and the top of the heatmap, respectively. (A), a heatmap of the HNSCC dataset (HPV(+) n=10 and HPV(-) n=13). (B), a heatmap of TCGA HNSCC dataset (HPV(+) n=46 and HPV(-) n=26 from anatomically matched tumours arising in the oropharynx, tonsil and base of tongue. In both datasets, tumours cluster according to HPV status; sub-clusters are evident in the larger TCGA dataset (data sourced from (Cancer Genome Atlas Research, 2014)

3.4.3.3 GO and pathway analysis of Immune cell corrected data:

GO and pathway analysis was again performed independently for genes expressed to a greater extent (n=219; Appendices B.3.2) and a lesser extent (n=218; Appendices B.3.2) in HPV(+) compared to HPV(-) tumours. The vast majority of immune and lymphocyte-related terms were no longer over-represented in HPV(+) tumours. This included markers of T-cell effector function (e.g., *IFNG, GZMB* and *PRF1*), which, prior to TIL correction were all over-represented in the HPV(+) tumours. Pathway analysis also confirmed the loss of immune-related signalling pathways (Appendices B.3.3). Immune GO terms that remained following correction for numerical differences in TIL were B-cell activation (GO:0042113), which included BCL2, VCAM1 and ICOSLG, with a greater expression in HPV(+) compared to HVP(-) tumours (Appendices B2). The surviving non-immune GO terms and biological pathways (Appendices B.3.3) over-represented in HPV(+) tumours were associated with cell cycle (GO:0007049), cell phase transition (GO:0044770) and chromosome organisation (GO:0051276).

The GO terms were again visualised in Revigo (Appendix B2.2), showing clearly the loss of immune GO's and instead, gene ontology terms relating to the HPV(+) associated tumour process. The loss of T-cell and the majority of B-cell-related GO terms following TIL correction of gene expression data, indicated that gene expression differences between HPV(+) and HPV(-) tumours largely resulted from numerical differences in these cell types. To determine if any differences in lymphocyte gene expression between the HPV(+) and HPV(-) tumour cohorts were retained following TIL correction, the DEGs identified after correction were overlapped with lymphocyte-specific marker genes (CD19+ B-cell genes and CD8+ and CD4+ T-cell genes, Appendices B.4 (Shoemaker et al., 2012, Abbas et al., 2005, Watkins et al., 2009, Palmer et al., 2006, Grigoryev et al., 2010, Whitney et al., 2003)). The majority of the surviving signals were associated with the following B-cell associated genes: GGA2, SPIB, CD200, ADAM28 as well as STAG3, which was not previously known to be a B-cell associated gene; these genes are discussed in more detail below. A single CD8-associated gene (CD8B) also survived correction (Figure 30).

GO and pathway analysis of DEGs expressed to a lesser extent in HPV(+) compared to HPV(-) tumours were largely unchanged following TIL correction of the data (Revigo Plots and gene ontology shown in Appendix B2.2, B.2.2 and B.2.3). There were 218 lesser expressed genes following lymphocyte correction, of which many were associated with development (skin, epidermis, epithelium, tissue, organ) and keratinization. In addition, lesser expressed genes associated with IL-12 and IL-6 production, the inflammatory response and the response to oxidative stress. Individual DEG's included keratin's (KRT10, 14, 16 and 17), kallikrein-related peptidase 5, 7 and 14 (KLK5, 7 and 14), caspase 14 (CASP14), tumour necrosis factor (ligand) superfamily member 9 (TNFSF9 or CD137L), thrombospondin receptor (CD36) and chemokine (C-C

motif) ligand 19 (CCL19); pathway analysis of lesser expressed genes returned no significantly over-represented pathways (q-value <0.05).

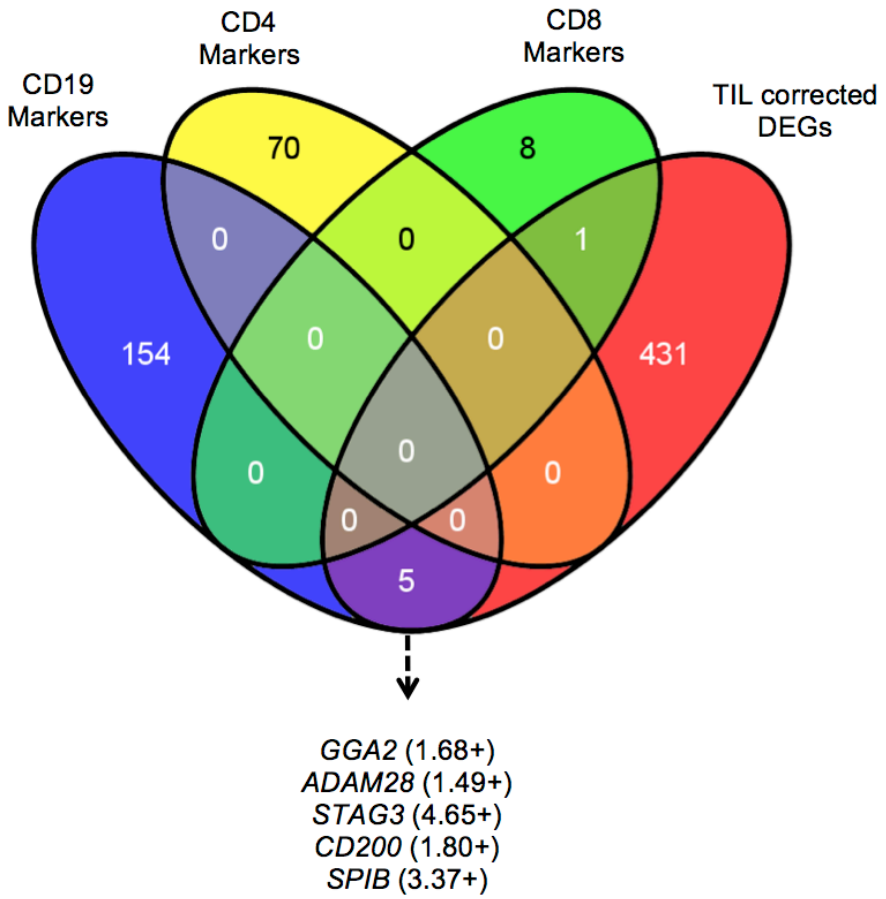


Figure 30. Differentially expressed genes between HPV(+) and HPV(-) tumours identified from TIL corrected data; overlap of DEGs expressed to a greater extent in HPV(+) tumours with published immune cell type gene sets.
A Venn diagram to illustrate the overlap of 437 DEGs expressed to a greater extent in HPV(+) versus HPV(-) tumours identified from TIL corrected gene expression data with immune cell type-specific marker genes as defined by at least two published databases: CD19 markers (n=159), CD4 markers (n=70) and CD8 markers (n=9). An increase in the expression of 5 B-cell-associated genes was observed in HPV(+) compared to HPV(-) tumours: *GGA2, ADAM28, STAG3, CD200* and *SIPB*; fold-change in expression is shown in brackets.

3.4.3.4 Further assessment of the B cell associated genes using TCGA RNA-Seq data

Analysis of the identified 8 B-cell-related genes GGA2, SPIB, CD200, ADAM28, BCL2, VCAM1, ICOSLG and STAG3 was carried out in an independent HNSCC dataset from the CGA (Cancer Genome Atlas, 2015). Analysis of these DEGs in the CGA data allowed potential anatomical bias to be addressed: since HPV(+) tumours predominantly arise in specific anatomical locations (tonsil, base of tongue and oropharynx) that could contribute to the DEGs.

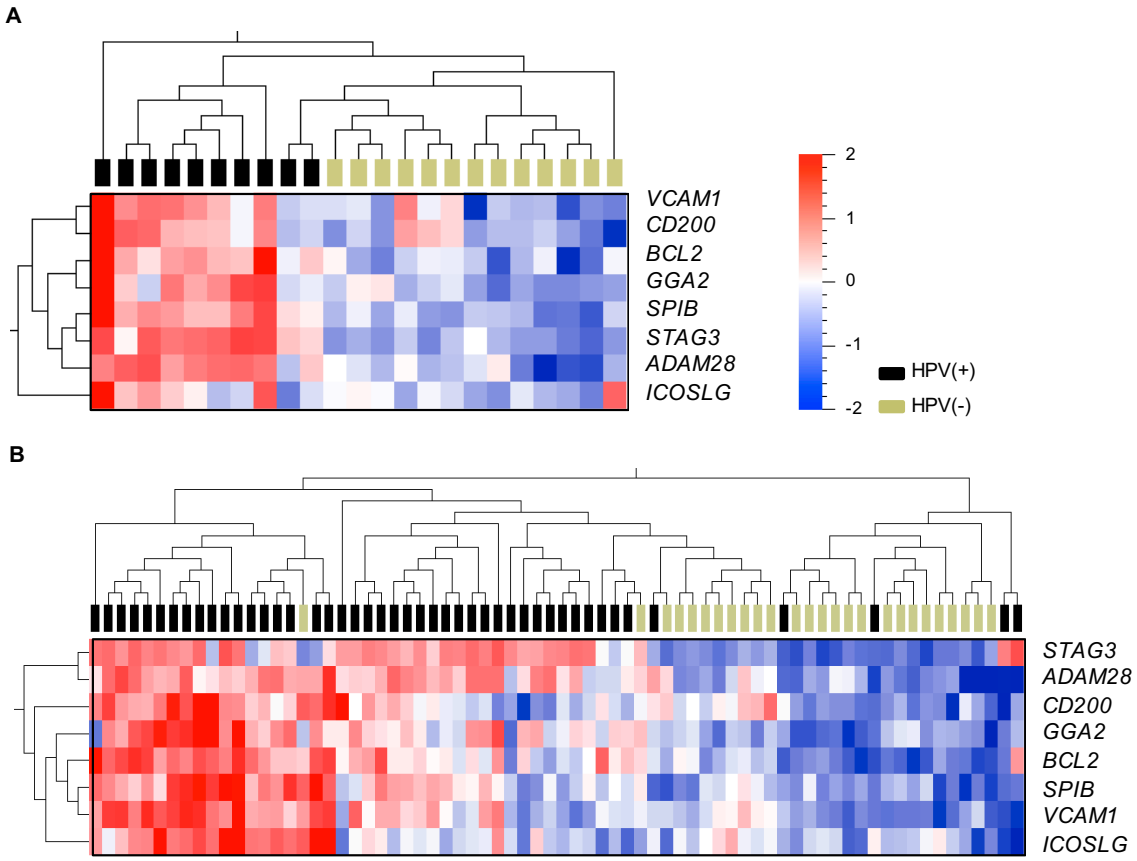


Figure 31. Expression of B-cell-associated genes by RNA-Seq.
Heatmaps to illustrate gene expression of the identified B-cell-associated genes between HPV(+) and (−) tumours: GGA2, SPIB, CD200, STAG3, ADAM28, BCL2, VCAM1 and ICOSLG, showing row wise z-scores of normalized read counts; the scale of z-score is shown: red shading denotes greater gene expression, blue shading denotes lower gene expression. Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of genes and tumours based on their expression profile is reflected in the dendrograms to the left and the top of the heatmap, respectively. (A), a heatmap of our HNSCC dataset (HPV(+) n=10 and HPV(−) n=13). (B), a heatmap of the TCGA HNSCC dataset (HPV(+) n=46 and HPV(−) n=26); publicly available data from anatomically matched tumours arising in the oropharynx, tonsil and base of tongue. In both datasets, tumours cluster according to HPV status, with a greater expression of the identified B-cell-associated genes in HPV(+) tumours (data sourced from (Cancer Genome Atlas Research, 2014)

The Identification of 72 cases (46 HPV(+) and 26 HPV(-)) from the anatomically matched locations (tonsil, base of tongue and oropharynx) were used for visualisation of the 8-gene signature (*GGA2*, *SPIB,* CD200, *ADAM28, BCL2*, *VCAM1, ICOSLG* and *STAG3).* The genes were hierarchically clustered in heatmaps using our data (Figure 31 A) and with the 72 CGA cases (Figure 31 B), the tumours clustered according to HPV status, with a greater gene expression in HPV(+) compared to HPV(-) tumours. These data confirmed that anatomical bias was not the reason for the B-cell-associated differences in gene expression. The larger CGA data however, did reveal a high expressing and moderate expressing sub-cluster within the HPV(+) cohort.

3.4.3.5 Validation of B cell-associated genes by RT-qPCR

In order to confirm these findings, RT-qPCR of CD200 and STAG3 was carried out on the whole tumour RNA samples used for the RNA-Seq (n=8 HPV(+) and n=8 HPV(-). This showed the same trend with HPV(+) tumours compared to HPV(-) tumours having increased expression of STAG3 and CD200 (Figure 32 A; STAG3, ***p<0.001 and CD200 ns, p=0.116). In addition to this, RT-qPCR was also used to investigate the expression of the genes GGA2, SPIB, CD200, STAG3, ADAM28, BCL2, VCAM1 and ICOSLG in B-cells isolated from an independent HPV(+) HNSCC tumour cohort (n=6) (Figure 32 B and A.1.3). Each of the identified B cell genes were confirmed as being expressed in the purified B cells, STAG3 was only expressed at a very low-level relative to the other genes; it is a component of the meiosis specific cohesin complex (Prieto et al., 2001). However, RT-qPCR of whole tumour tissue confirmed its differential expression between HPV(+) and HPV(-) tumours highlighting that the differential signature is likely not to be originating from the B-cell compartment.

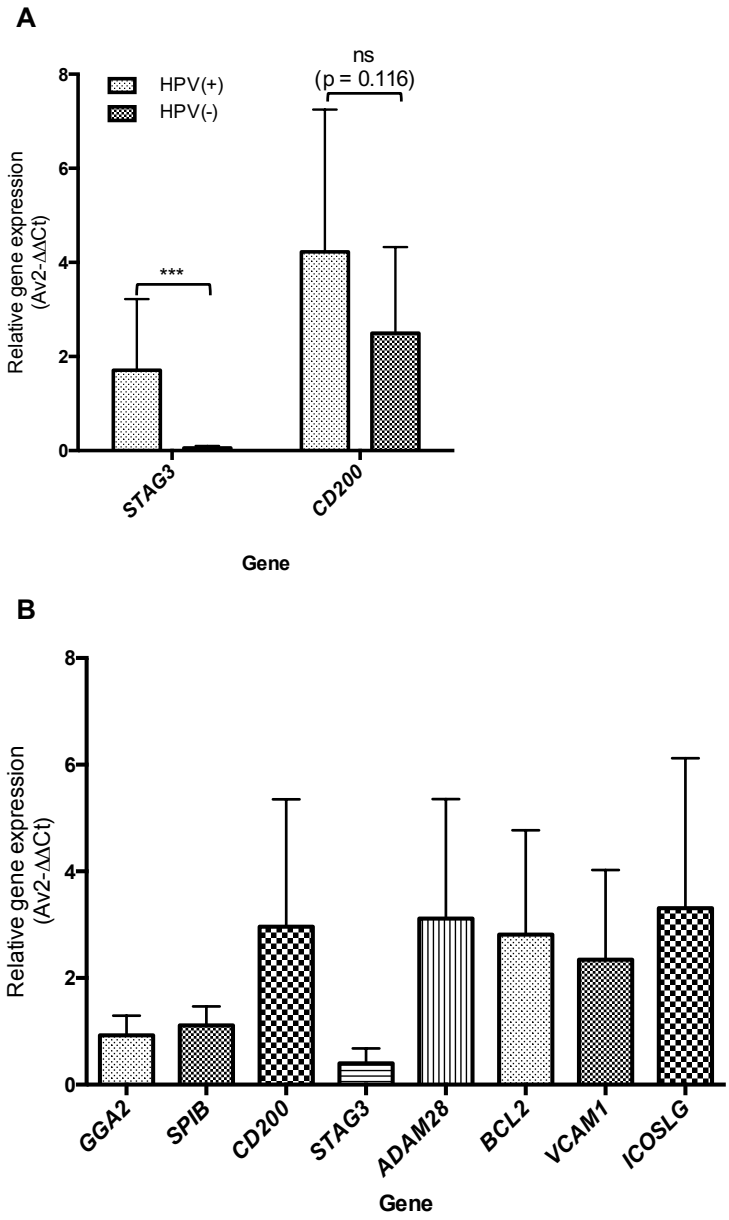


Figure 32. Relative expression of B-cell-associated genes by RT-qPCR.

(A) The average (mean and standard deviation) relative gene expression of CD200 and STAG3 was measured by RT-qPCR* in RNA extracted from the whole tumour, as used for the RNA-Seq analysis. The expression of STAG3 and CD200 was determined for HPV(+) (n=8) and HPV(-) (n=8) tumours. This showed the same trend with HPV(+) tumours compared to HPV(-) tumours having increased expression of STAG3 and CD200 (STAG3, ***p<0.001 and CD200 ns, p=0.116). (B) Displays the average (mean and standard deviation) relative gene expression of B-cell-associated genes measured by RT-qPCR*. The expression of the B-cell-associated genes GGA2, SPIB, CD200, STAG3, ADAM28, BCL2, VCAM1 and ICOSLG was confirmed in B-cells sorted from an independent cohort of HPV(+) tumours (n=6).

*Relative gene expression by RT-qPCR, calculated using the comparative Ct method with Actin as the control gene (2^-(ΔΔCt) method) (23). Asterisks in column labels indicate a significance level of a two-sample t-test comparison of RT-qPCR between HPV(+) and HPV(-) tumours: ns = not significant (value stated) and ***P <0.001).

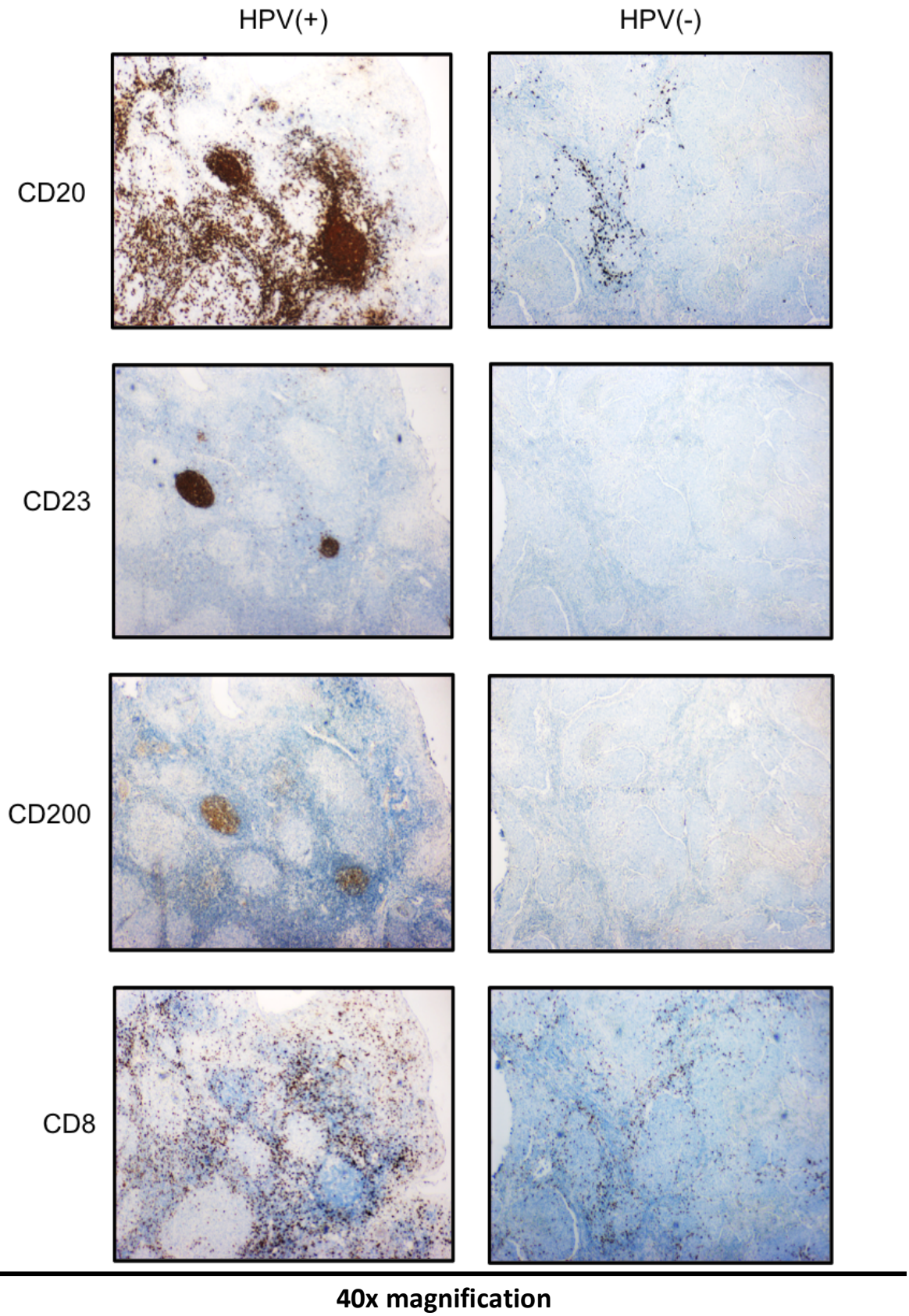


Figure 33. Expression of B-cell markers by IHC

Cell subset analysis by IHC for the B-cell markers CD20, CD23 and CD200, as well as the T-cell marker CD8, was performed on sequential sections for HPV(+) (n=9) and HPV(−) (n=13) tumours; representative data is shown for one tumour from each cohort. Tertiary lymphoid structure formation is apparent in HPV(+) tumours along with dense infiltrate of CD20+ B-cells.

(IHC slides prepared by research histology – Monette Lopez and Maria Machado).

3.4.3.6 IHC assessment of B cells in HPV(+) and (-) tumours

IHC assessment of HPV(+) and HPV(-) tumours identified dense clusters of tumour-infiltrating B-cells in the HPV(+) samples; Figure 33 shows representative histology for one HPV(+) and one HPV(-) tumour. The follicular morphology within HPV(+) tumours was apparent following staining for CD23, a marker of follicular B-cells. Furthermore, IHC confirmed the presence of CD200+ cells within and around tertiary lymphoid follicles, together with a diffuse CD8+ T-cell infiltrate. These features were absent in the HPV(-) tumours although sparse CD8 and B-cell infiltrates were seen.

3.4.4 Discussion

It has previously been shown that HPV(+) HNSCC patients with a dense immune cell infiltrate within the tumour have a better outcome than those with sparse infiltrates (Ward et al., 2014b), this has also been identified in HPV(-) tumours (Wood et al., 2016b, Wood et al., 2016a). Here, fine-resolution transcriptomic analysis was used to evaluate whether it could give insight into the biological difference in the immune infiltrate between patients with a known viral driver (HPV), compared with those patients where virus is absent. Although differential gene expression studies comparing HPV(+) and HPV(-) tumours have been reported (Jung et al., 2010, Pyeon et al., 2007, Slebos et al., 2006, Russell et al., 2013, Cancer Genome Atlas, 2015), the question as to whether the pathogenesis of the cancer is reflected in differences in immune cells themselves remains open.

Here the focus was to analyse $TIL^{Hi/mod}$ cases from a consecutive cohort and, following RNA-Seq analysis, examine the DEGs. Initially, the most striking difference was the immune signature, which was significantly greater in HPV(+) tumours and hence likely reflected an immune responses to viral antigens. In contrast HPV(-) tumours revealed a prominent tissue development/re-organisation gene signature. The molecular data mapped well onto the TIL characterization by IHC, although the variability was higher when TILs were counted manually. This was likely a reflection of the fact that RNA-Seq analysis uses homogenized tumour, which averages the geographical differences within the tissue.

A significant difference in TIL density, determined by RNA-Seq gene transcript levels of CD4 and CD8A (T-cells) and CD19 (B-cells), remained between the HPV(+) and HPV(-) patient cohorts. This was clearly demonstrated in the IHC assessment, HPV(-) $TIL^{Hi/mod}$ patients had a significantly lower expression of B- and T-cell-related genes compared with HPV(+) $TIL^{Hi/mod}$ patients. In order to assess the difference between HPV(+) and HPV(-) TIL enriched tumours, the data were corrected according to the number of immune cells present in the tissue as determined by RNA-Seq gene transcript levels (*CD8A*, *CD4* and *CD19*), with the aim of accounting for the numerical difference in immune cells.

RNA-Seq data that had been corrected for TIL number showed that the vast majority of immune-related DEGs were no longer differentially expressed, suggesting that in both patient cohorts the lymphocytes were qualitatively similar. The T-cell immune signature, present in both HPV(+) and HPV(-) tumours, no longer showed differentially expressed genes. Thurlow *et al.* have previously demonstrated that both HPV(+) and HPV(-) cohorts can mount adaptive immune responses (Thurlow et al., 2010), while the association of cytolytic activity of effector cells and immunoediting of the tumour has also been reported (Rooney et al., 2015). These studies however, did not correct for lymphocyte numbers; the HPV(+) tumours in which no adaptive response was detected, likely represented TIL^{low} patients. It was expected that a viral driver would promote a distinct T-cell-driven TIL signature, but this was not observed in the data. Instead, the data demonstrate that T-cells in $TIL^{Hi/mod}$ tumours of different pathogeneses are transcriptomically similar, at the level of bulk tumour analysis. This has important implications in the treatment of immune rich tumours, as it argues that the type of cancer, in this case HPV(+) and (-) display similar immune responses. This then leaves the tumour and the tumour microenvironment as a key determining factor in the immunological response.

In contrast, there was a distinct B-cell signature between HPV(+) and HPV(-) tumours after correction for TIL numbers. There was a small subset of B-cell associated genes that continued to have greater expression in HPV(+) tumours, including GGA2, SPIB, CD200, ADAM28, BCL2, VCAM1 and ICOSLG, as determined by previously published datasets for B-cell gene signatures and the GO term (B-cell activation, GO:0042113) (Shoemaker et al., 2012, Abbas et al., 2005, Watkins et al., 2009, Palmer et al., 2006, Grigoryev et al., 2010, Whitney et al., 2003). The expression of the differentially expressed B cell-associated genes is not unique to B cells as they can be identified in other cell types. This includes CD200, which has been found to be expressed on 1-2% of basal cell carcinoma cells (Colmont et al., 2013). However, the data on purified B-cell populations from HPV(+) HNSCC confirms the expression of the BCL2, ADAM28, CD200, ICOSLG and SPIB genes in B cells isolated from HNSCC tumours. The DEGs were confirmed in a larger, independent cohort derived from HNSCC CGA data. This showed that the DEGs were not the result of anatomical location bias in HPV(+) tumours, the differences were maintained in anatomically matched HPV(+) and HPV(-) tumours. Additionally, applying the TIL-corrected DEG list to the CGA confirmed the validity of these observations.

Greater expression of BCL2 by HNSCC has been proposed as a predictor of good response to chemotherapy; this is consistent with its expression in our HPV(+) cohort, a group of patients that respond well to treatment, including chemotherapy (Moreno-Galindo et al., 2014, Ward et al., 2014a). During a normal humoral response, ICOSLG is expressed on activated B-cells within germinal centres, which are formed in follicles and are central for an antigen-specific humoral response. Histologically, all of the HPV(+) tumours had very well developed tertiary lymphoid

structures (TLS) (Liu et al., 2014). TLS within solid human tumours have previously been described in breast, cervical and non-small-cell lung carcinoma (Kobayashi et al., 2002, Nelson, 2010). A survival benefit has also been observed in tumours that are positive for TLS, however the formation of a TLS is accompanied by a plethora of immune cells that play a role in antitumor activity and so the specific cell types involved in this survival benefit are not at this time clear (Wirsing et al., 2014). The abundance of TLS in tumours may then be an excellent biomarker of an on-going protective immune response, the data here shows high numbers of TLS in immunogenic (abundant antigens) HPV(+) tumours, these also have dense overall immune infiltrates. Evaluation of B cells and TLS in other tumour types is needed to determine their role immunotherapy responses.

In melanoma, ICOSLG is linked with increasing numbers of regulatory T-cells (T_{reg}) (Martin-Orozco et al., 2010), but has not previously been described in HPV(+) HNSCC. When a stimulatory ICOS antibody was used in combination with anti-CTLA-4, a significant improvement in tumour rejection in both melanoma and colon cancer mouse models was observed, suggesting that HPV(+) tumours expressing high levels of ICOSLG may respond better to anti-CTLA-4 therapy (Fan et al., 2014). *SPIB* is a transcriptional activator that is specific for lymphoid cells and has previously been identified in germinal centres where it was associated with an activated B-cell phenotype (Su et al., 1996).CD200+ B-cells were identified using IHC within and peripheral of the follicular structures. The viral antigens E6 and E7 generate a strong B-cell response in patients with HPV(+) HNSCC (Liang et al., 2012). It is likely that the expression of the B-cell activation marker CD200 (Barclay et al., 2002) is linked to a persistence of HPV-driven, tumour-derived antigens, which in turn stimulate specific B-cells in the germinal centre and the tumour itself. It is also possible that the CD200+ B-cells resident at specific locations in the tumour vs. the follicles, have a distinct phenotype and properties. These cells could be part of an inhibitory pathway: CD200R activation stimulates the differentiation of T-cells to T_{reg} (Holmannova et al., 2012), the numerical increase of T_{reg} in parallel with TIL number has been reported in HPV(+) HNSCC (Badoual et al., 2006). These $CD200^{+}$ B cells can act through CD200R and effector mechanisms may include immunomodulatory cytokines such as IL-10, (Lindner et al., 2013) release of granzyme B (Iwata et al., 2011) or other yet to be identified modes of action. The receptor is expressed on both lymphoid and myeloid cells, hence direct and indirect inhibition of T-cell immunity could be possible (Rygiel and Meyaard, 2012). It must, however, be noted that expression of CD200 is thought to have links to anti-tumour effects by inhibiting activity of tumour-associated myeloid cells via IL-10 (Wang et al., 2010), arguing caution in targeting a molecule with potentially pleiotropic effects.

It is evident that a phenotypic differences exists between the B-cells of HPV(+) and HPV(-) tumours. Data from an HPV-driven mouse tumour model supports a reduction in tumour growth

as a result of the depletion of B-cells via anti-CD20 (Affara et al., 2014). It is tempting to speculate that anti-CD200 could provide a more targeted approach to B-cell manipulation within these tumours rather than the use of anti-CD20, which would globally remove multiple B-cell subsets. In the clinical setting, an anti-CD200 blocking antibody was safe and well tolerated in Phase II testing (NCT00648739), but as yet the clinical effects have not been reported. The B-cell receptor targeting therapies (ibrutinib) have efficacy on $CD200^{+}$ cells within the B-CLL setting and potentially could be used to remove a cell that has a negative role in antitumor immunity (Woyach et al., 2012).

3.4.5 Conclusions

Analysis of HPV(+) and HPV(-) HNSCC, controlled for TIL number using RNA-Seq, identified a number of genes as being expressed both to a greater and lesser extent between tumour types. Correction of the RNA-Seq data to account for numerical differences in immune cells between virally driven and virus-independent HNSCC, was an alternative approach to revealing DEGs. Analysis of the data post correction for immune cells, revealed that differences between the transciptome in T-cells between HPV(+) and HPV(-) HNSCC appears predominantly quantitative, but that a distinct B-cell profile exists in HPV(+) cancers. In order to assess the qualitative differences in B- and T-cells between HPV(+) and HPV(-) tumours the immune cells need to be isolated and analysed separately.

Chapter 4: Investigation of Tumour infiltrating lymphocytes with high resolution analysis of purified immune cell subsets.

The prognostic relationship between intra-tumoural density of immune infiltrates and survival is well documented, with immunomodulatory therapies aiding patient survival. However, these therapies still fail in a large proportion of patients, for reasons that are still unclear. Better knowledge of immune cells within the tumour micro-environment may allow alternative treatments to be implemented to further improve patient survival. One way of understanding the features of the immune response is by using transcriptomic analysis of the isolated immune cells.

4.1 Isolation of immune cells from cancer

4.1.1 Introduction

High resolution transcriptomic analysis of purified immune cell subsets presents an opportunity to understand immune cells in different clinical contexts. The technique of cell sorting has been used extensively in mouse immunology to understand the gene expression patterns of different cell types (Painter et al., 2011, Cohen et al., 2013). Micro-scaled RNA-Seq on sorted immune cells has also been used to assess phenotype and function in the asthma setting (Seumois et al., 2012, Seumois et al., 2014). The application of cell sorting to tumour-infiltrating lymphocytes followed by transcriptomic analysis has the potential to allow a better understanding of immune function and phenotype in different clinical settings. The first stage in this process is dissociation of the tumour using mechanical and enzymatic (collagenase) methods to generate a single cell suspension. Data from a comparison of different collagenase preparations has shown that Liberase DL (Roche) led to less cell death and lower proteolytic cleavage of cell surface markers (Layfield, PhD unpublished data), when compared to more crude collagenase D preparations (Sigma). In order to isolate different immune cell populations, an antibody panel that would delineate the immune cells into CD8+, CD4+, B cells and macrophages (activated monocytes - CD14+ HLA-DR+) was also required. Standardised immune phenotyping is routine and an amended protocol was designed to isolate these cells types (Maecker et al., 2012). The immune density, as determined by flow cytometry, will be assessed in relation to previous studies for HNSCC and NSCLC (Lechner et al., 2017, Lizotte et al., 2016, Stankovic et al., 2018).

4.1.2 Objectives

Isolation of the immune cells CD8+, CD4+ T -cells, activated monocytes / macrophages (activated monocytes - CD14+ HLA-DR+ cells will broadly be termed macrophages) and B cells from freshly resected tumours to gain further insight into their characteristics in these tumours. Initial optimisation of cell isolation methods and generation of a protocol was required for robust and continued collection of immune cells from patient tumours for subsequent analysis. Once a cohort of samples had been acquired, a collaborative work flow was used to undertake the RNA-Seq, using proprietary methodology developed in La Jolla, Institute for Allergy and Immunology (Seumois et al., 2012). The analysis will focus on the assessment of CD8+ T cells and B cells in tumours isolated from HNSCC and NSCLC in relation to immune density and clinical parameters.

4.1.3 Results

4.1.3.1 Optimisation of Immune cell sorting

Initial testing of a FACS panel for immune cell sorting (sorting CD4+ and CD8+) was carried out using PBMC's (Appendix C.1.1). This allowed the flow cytometry gating strategy and optimal sort settings to be established prior to use on HNSCC and NSCLC tumours (Appendix C.1.1). The initial optimisation FACS panel (CD45, CD3, CD4 and CD8) sort gating strategy is shown in Appendix C.1.1, the gating structure uses Singlet exclusion > dead cell exclusion > lymphocyte gating (CD45+) > CD3+ and then CD4+ and CD8+ as the sort gates. Initial testing on tumour samples highlighted that cell aggregates occurred once the tumour was dispersed leading to flow cytometer blockages and a time delay. The flow cytometry buffer was adjusted to contain 2mM EDTA; this chelates Ca^{2+} and alleviates the problem of cell adhesion and formation of aggregates. When isolating CD8+ cells, the purity of the sort was >98%; this panel was then used to sort n=18 HNSCC tumours.

4.1.3.2 Immune cell sorting panel

In order to capture additional immune cell populations and maximise the 4-way sort capabilities of the FACS ARIA, the flow cytometry panel was expanded to include the markers CD14 and HLA-DR to capture activated monocytes (the CD14+ HLA-DR+ cells will broadly be termed macrophages) and also B cells with CD19 and CD20; the gating strategy is shown in Figure 34 A and B. Due to instrument limitations with the number of available fluorophores, CD19 and CD20 were run in the same fluorophore as CD8 (PerCP-Cy5.5), using the CD3 positive/ negative gate to distinguish between them. This protocol was then formalised into an SOP and used to sort

immune cells from HNSCC and NSCLC. It has also been used on two different clinical trials as tertiary endpoint assays (HARE40 and AMG319), please see Appendix A2 for details.

PBMC

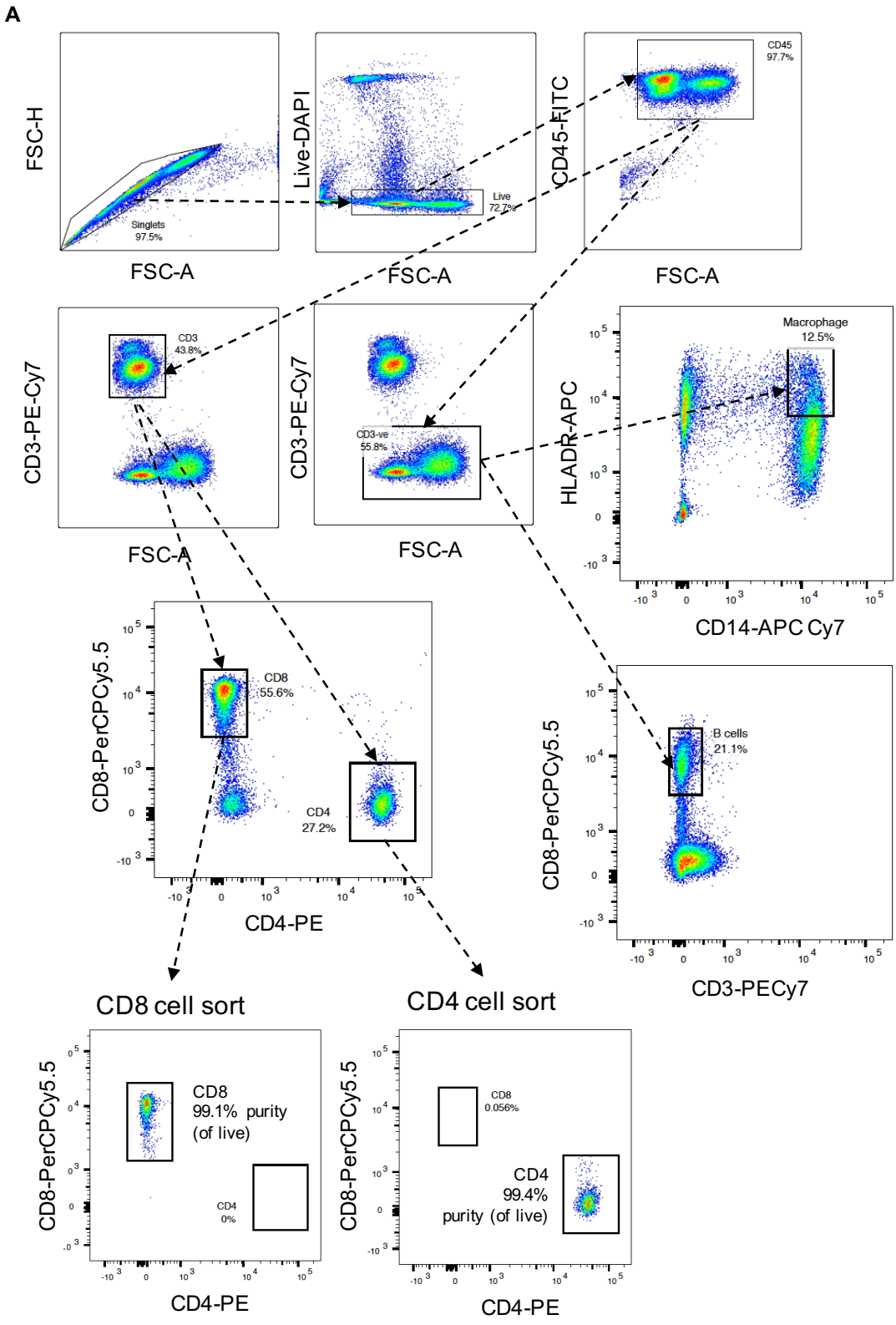


Figure 34. Immune cell sorting panel.

Tumour

B

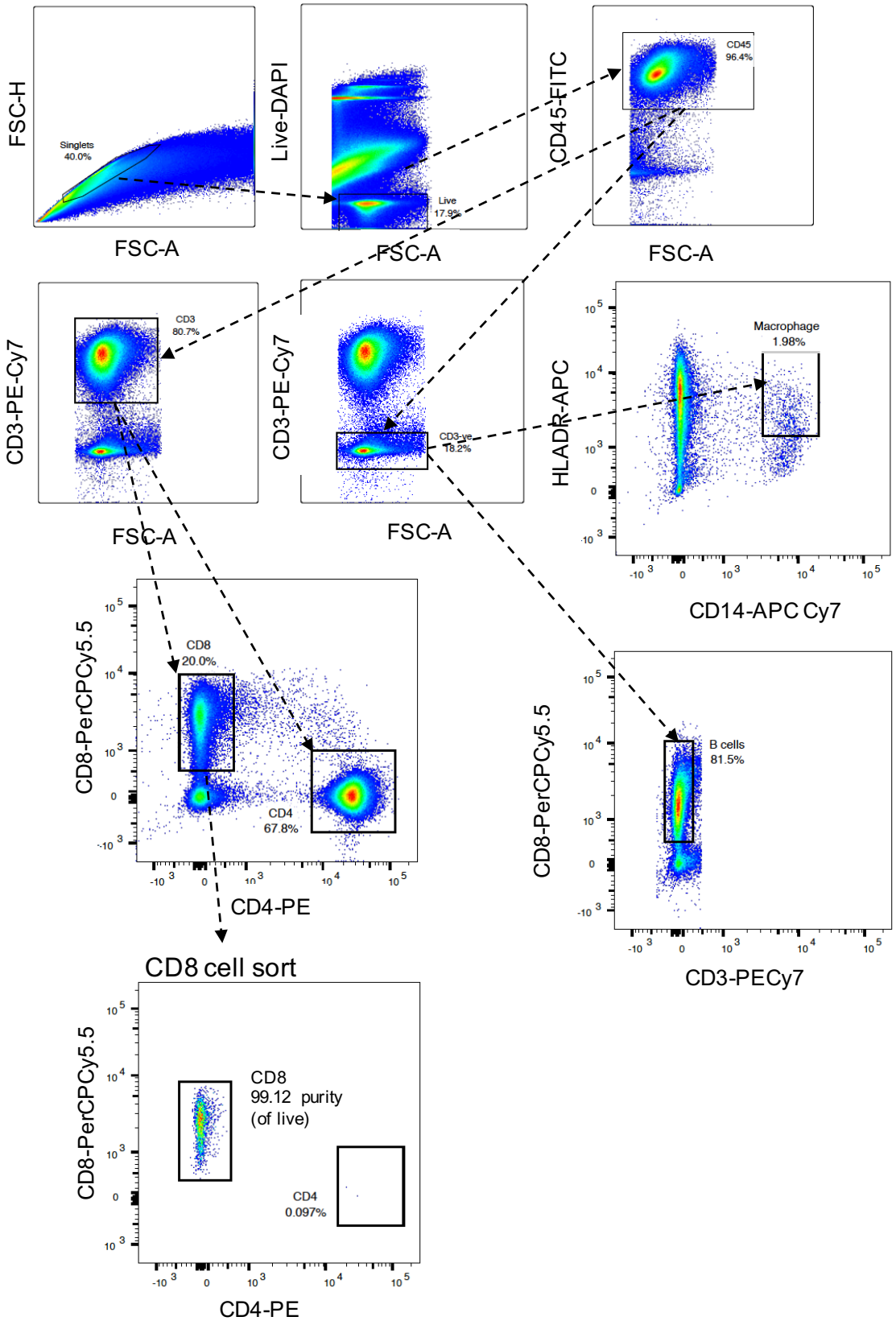
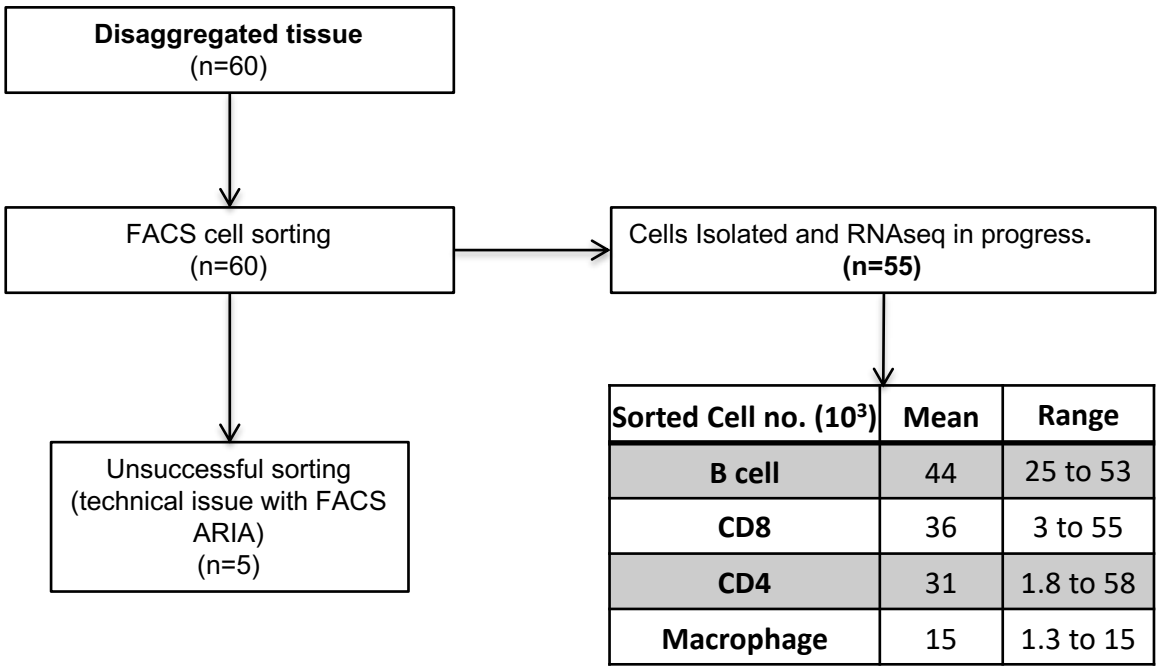

FSC-H
Singlets
40.0%
FSC-A
Live-DAPI
Live
17.9%
FSC-A
CD45-FITC
CD45
96.4%
FSC-A
CD3-PE-Cy7
CD3
80.7%
FSC-A
CD3-PE-Cy7
CD3-ve
18.2%
FSC-A
HLADR-APC
Macrophage
1.98%
CD14-APC Cy7
CD8-PerCPCy5.5
CD8
20.0%
CD4
67.8%
CD4-PE
CD8-PerCPCy5.5
B cells
81.5%
CD3-PECy7
CD8 cell sort
CD8-PerCPCy5.5
CD8
99.12 purity
(of live)
CD4
0.097%
CD4-PE

Figure 34. Immune cell sorting panel.
Immune cell gating strategy for isolation of CD8+, CD4+ T cells, B cells and Macrophages (activated monocytes) from tumour samples. (A) FACS panel tested on PBMC and (B) HNSCC tumour, both display high purity (>98%) for CD4+ and CD8+ T cells.

4.1.3.3 Immune cell distribution in HNSCC and NSCLC determined by FACS

4.1.3.3.1 Immune cell isolation in HNSCC

Using the optimised tissue-dissociation and flow-sorting panel, immune cells were isolated from HNSCC tumours. The average number of immune cells and range for the samples collected are outlined in Table 20, the target number of events for each sort was between 1000 and 50,000 cells for each population.



Sorted Cell no. (10^3)	Mean	Range
B cell	44	25 to 53
CD8	36	3 to 55
CD4	31	1.8 to 58
Macrophage	15	1.3 to 15

Table 20. Average and range for each sorted cell population in HNSCC.
The number of cells sorted for the CD8+, CD4+, B cells and macrophages (activated monocytes) using the immune cell sorting panels. Due to technical problems with the flow cytometer 5/ 60 cases were not sorted.

The distribution of the immune cells as determined by FACS, is shown in Figure 35 A and B; bar plots (geometric mean and 95% confidence interval) are shown as a percentage of the parent population. The immune cell proportion for each population from its parent gate is summarised in Table 21. The data show that CD4+ cells are present at a ratio of 2:1 to CD8 T cells ($p<0.001$) and B cells, with the B cells and CD8 cells having a similar abundance, macrophages (activated monocytes) were the least frequent cell in HNSCC.

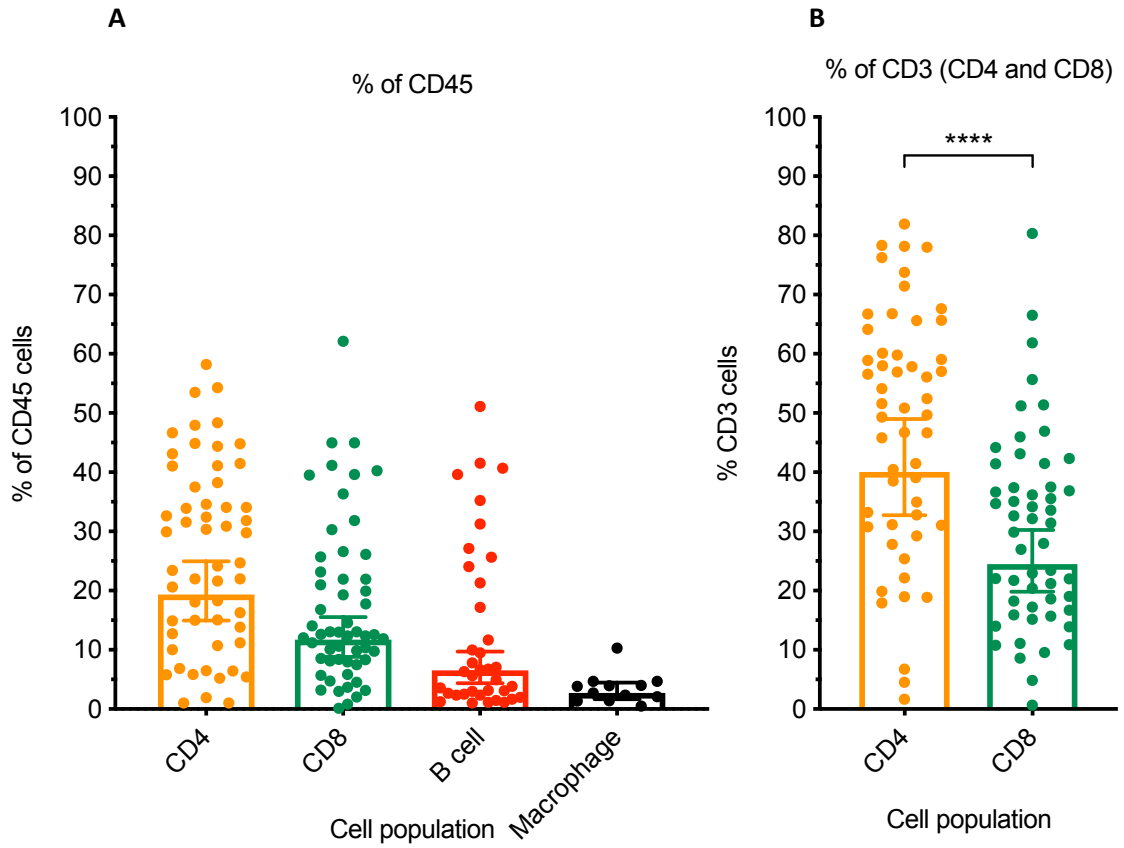


Figure 35. Immune cell percentages for CD8+, CD4+ T cells, B cells and Macrophages (activated monocytes) **in HNSCC.**

The geometric mean cell frequency (error bars represent the 95% confidence interval) for CD8+, CD4+, B cell and macrophages (activated monocytes) was plotted from CD45 and CD3 populations (n=55). (A) % CD4+ and (B) % CD3+. The CD45+ parent gate yields comparable cell frequencies to the reported literature.

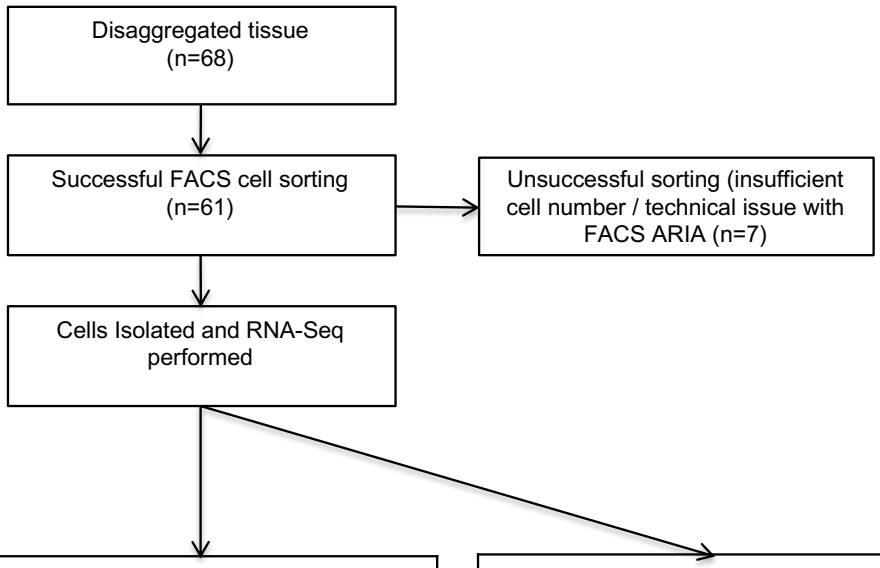
	Singlet	Live	CD45+	CD4+	CD8+	B cell	Macrophage (activated monocytes)
n	55	55	55	55	55	37	12
% Total	86.82	55.14	10.33	3.346	1.64	1.646	0.1306
% Singlet		62.64	12.06	3.838	1.897	1.76	0.1374
% Live			23.09	7.582	3.751	2.44	0.2191
% CD45+ve				26.04	17.1	12.63	3.496
% CD3+ve				47.33	29.92		

Table 21. The median percentage of each cell type derived from flow cytometry in HNSCC.

Each cell population was assessed from different parent gating strategies displayed as median % of each cell type. Cell populations as a defined by % of CD45 yield comparable results to the literature.

4.1.3.3.2 Immune cell sorting in Lung tumour and background lung.

A collaborative laboratory collection process was implemented to isolate immune cells from NSCLC as well as adjacent non-involved back ground lung tissue using the same process as for HNSCC. The Immune cell sorting used the same flow cytometry panel (Figure 34) allowing isolation of CD4+, CD8+, B cells and Macrophages (activated monocytes). The number of sorted cells for each population is shown in Table 22.



Lung tumour n=48		
Sorted Cell no. (10^3)	Mean	Range
B cell	16	1.4 to 53
CD8	22	1.6 to 51
CD4	27	4.8 to 54
Macrophage	24	2.7 to 53

Background lung tissue n=37		
Sorted Cell no. (10^3)	Mean	Range
B cell	17	1.7 to 51
CD8	33	18 to 52
CD4	39	11 to 54
Macrophage	33	1.4 to 52

Table 22. Average and range for each sorted cell population from NSCLC and non-involved background lung.
The number of cells sorted for the CD8, CD4, B cells and macrophages (activated monocytes) using the immune cell sorting panel. Due to technical problems with the flow cytometer 7 cases were not sorted.

The distribution of the immune populations (CD8+, CD4+, B cell and Macrophages (activated monocytes) between tumour and background lung gated on CD45 and CD3+ cells is shown in Figure 36 A and B as well as a summary in Table 23. The overall percentage of immune cells depended on which parent gate was used for the analysis, as was observed in HNSCC. The number of CD4+, CD8+ cells and B cells in the CD45+ gate (Figure 36 A) was significantly different between tumour and non-involved lung (NIL), with tumours having higher numbers, macrophages (activated monocytes) were similar in NIL (5.9%) and tumour (5%). When the cell numbers were calculated form the CD3+/- gate (Figure 36 B) only the B cells were significantly different between

the Tumour and NIL, however the CD4+ cells were higher in the tumour but did not reach significance (p=0.07).

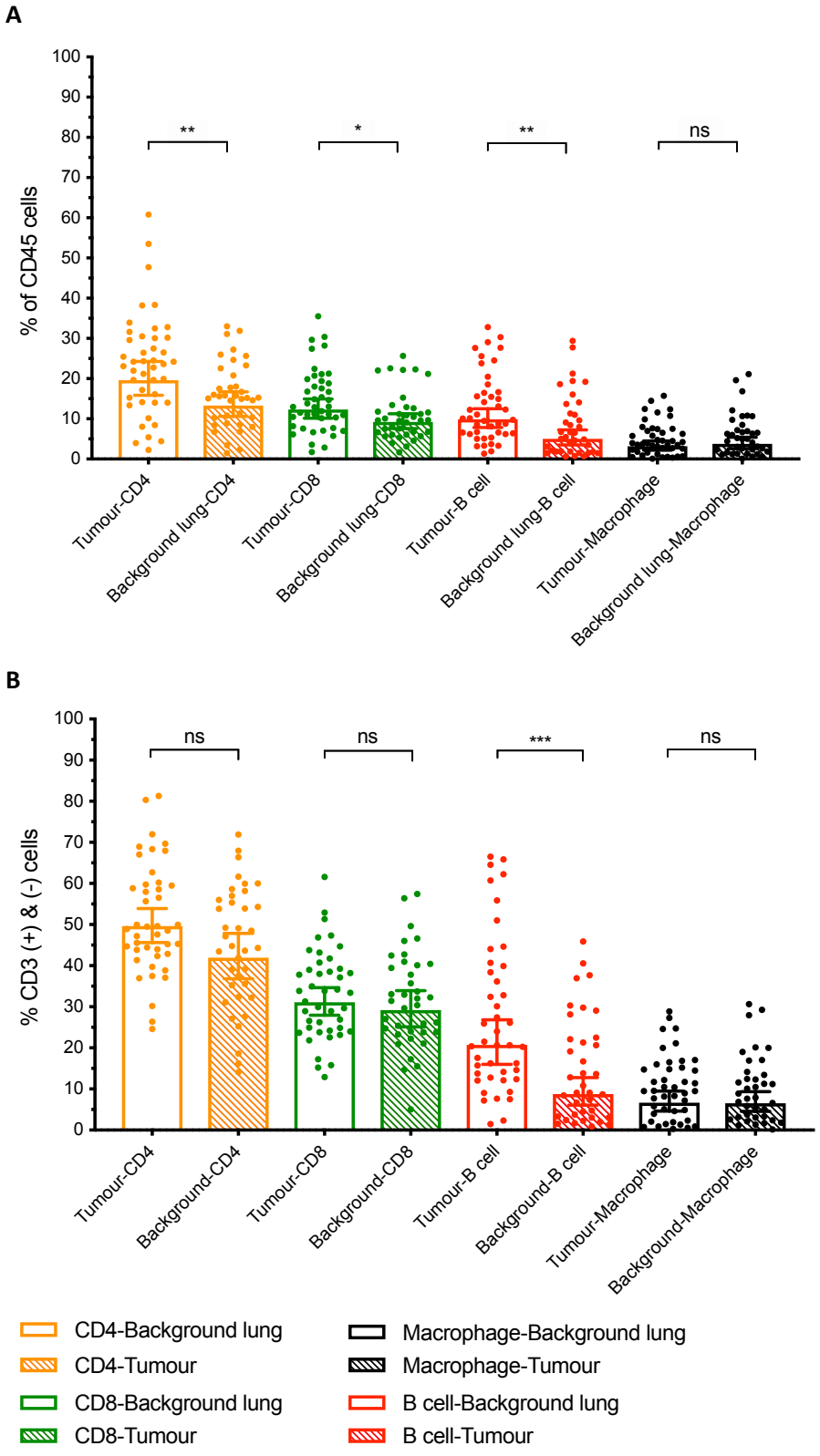


Figure 36. Immune cell percentages for CD8+, CD4+ T cells, B cells and Macrophages (activated monocytes) in Lung tumour and background lung tissue.

The geometric mean cell frequency (error bars represent the 95% confidence interval) or CD8+, CD4+, B cell and macrophages (activated monocytes) was plotted from the CD45+ and CD3+ cell

populations (n=55). (A) % CD45+ and (B) % CD3+. The CD45+ parent gate yields comparable cell frequencies to the reported literature.

Lung tumour

	Singlets	Live	CD45	CD3+	CD3-	CD4	CD8	B cell	CD14	Macrophage
n	43	43	43	43	43	43	43	43	43	43
% Total	73.1	40.9	8.8	4.1	4.2	2.1	1.4	1.1	0.8	0.4
% Singlet		54.3	12.2	5.6	6.0	2.9	1.9	1.6	2.6	5.0
% Live			33.5	16.0	16.2	9.0	5.0	4.6	9.6	14.1
% CD45+ve				44.5	50.6	23.5	14.5	12.7	9.6	5.0
% CD3(+) & (-)						51.3	32.9	27.3	19.8	10.4

Background lung

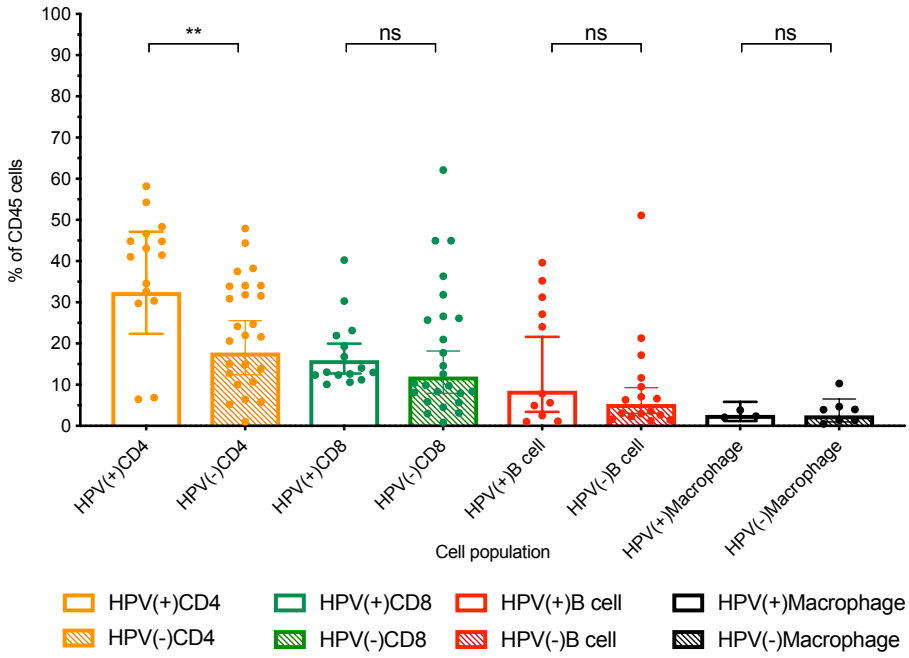
	Singlets	Live	CD45	CD3+	CD3-	CD4	CD8	B cell	CD14	Macrophage
n	37	37	37	37	37	37	37	37	37	37
% Total	67.9	40.8	17.8	5.8	10.7	2.5	1.7	1.5	3.5	1.0
% Singlet		57.4	28.0	8.9	17.2	3.9	2.7	2.1	10.0	5.9
% Live			58.5	20.2	35.0	9.2	6.5	4.4	16.8	21.9
% CD45+ve				34.7	59.0	15.9	10.8	8.3	16.8	5.9
% CD3(+) & (-)						44.8	31.7	14.5	27.8	10.0

Table 23. The median percentage of each cell type derived from flow cytometry in Lung tumour and non-involved background lung tissue.
Each cell populations assessed from different parent gating strategies was displayed as median % of each cell type. Cell populations as a defined by % of CD45 yield comparable results to the literature.

4.1.3.4 Analysis of immune cell frequencies in HNSCC and NSCLC in relation to tumour type.

The immune cell distribution in the subtypes HPV(+) and (-) HNSCC; Adenocarcinoma and Squamous cell lung carcinoma were assessed by gating on CD45+ cells (Figure 37 A and B). In HPV(+) HNSCC, only CD4+ cells were significantly higher compared to HPV(-) tumours; CD8+, B cells and Macrophages (activated monocytes) were not significantly different but showed a trend towards being higher in HPV(+) tumours (Figure 37 A). When comparing the histological category of Adenocarcinoma and Squamous cell lung carcinoma, there were no differences in immune cell percentages when gated on CD45+ cells, although adenocarcinoma appeared to have elevated CD4+ cell frequencies (Figure 37 B).

A



B

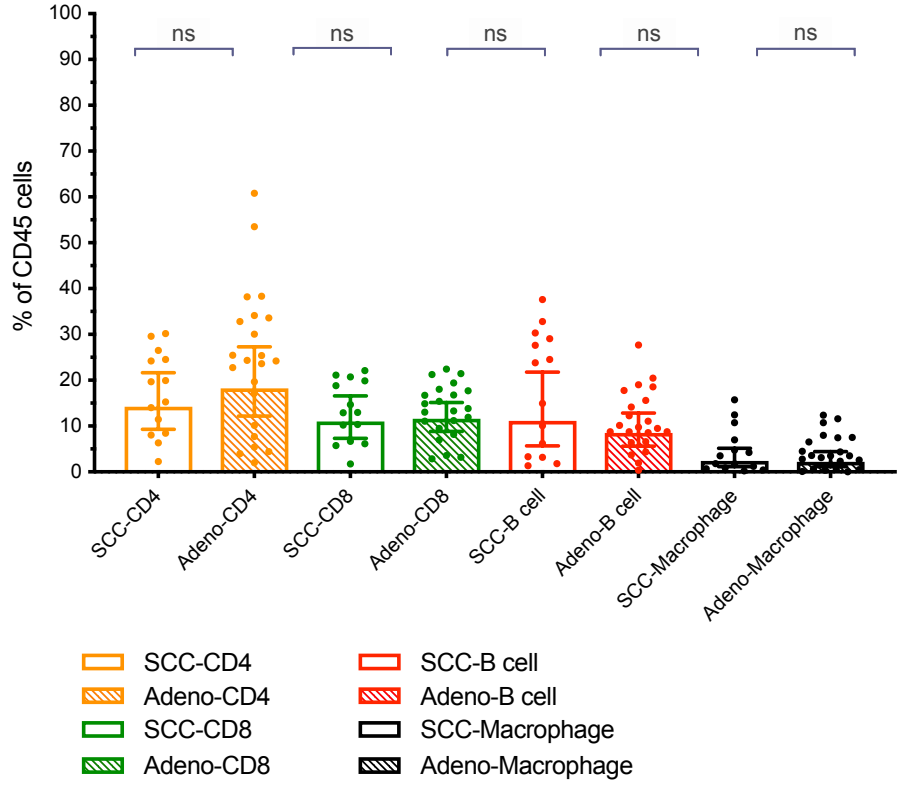


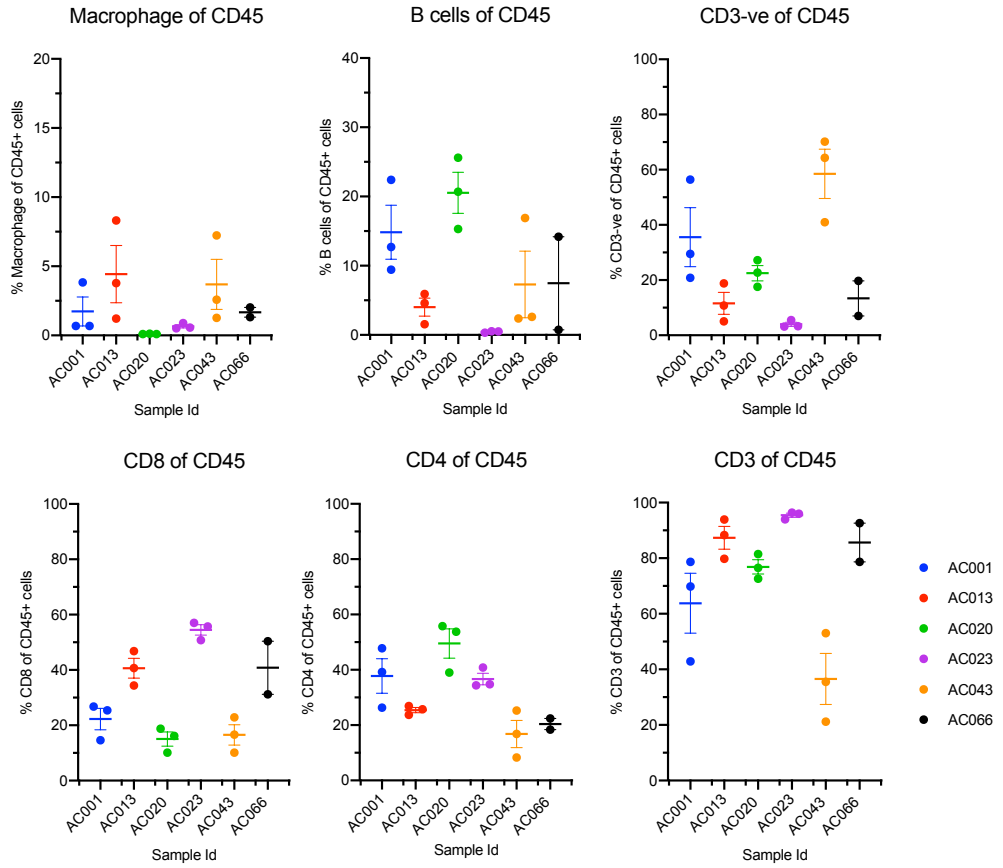
Figure 37. Comparison of immune cell infiltrate and tumour type in HNSCC and NSCLC.

The immune infiltrate (CD4+, CD8+, B cell and macrophage/ activated monocytes) gated from the CD45+ cells are shown as a percentage (geometric mean, where error bars represent 95% confidence interval). (A) The HNSCC cohort when divided into HPV(+) and (-) only shows a significant difference in CD4+ T cells. (B) The histologic subtypes of squamous cell carcinoma and adenocarcinoma display no significant differences across the immune cell types in NSCLC.

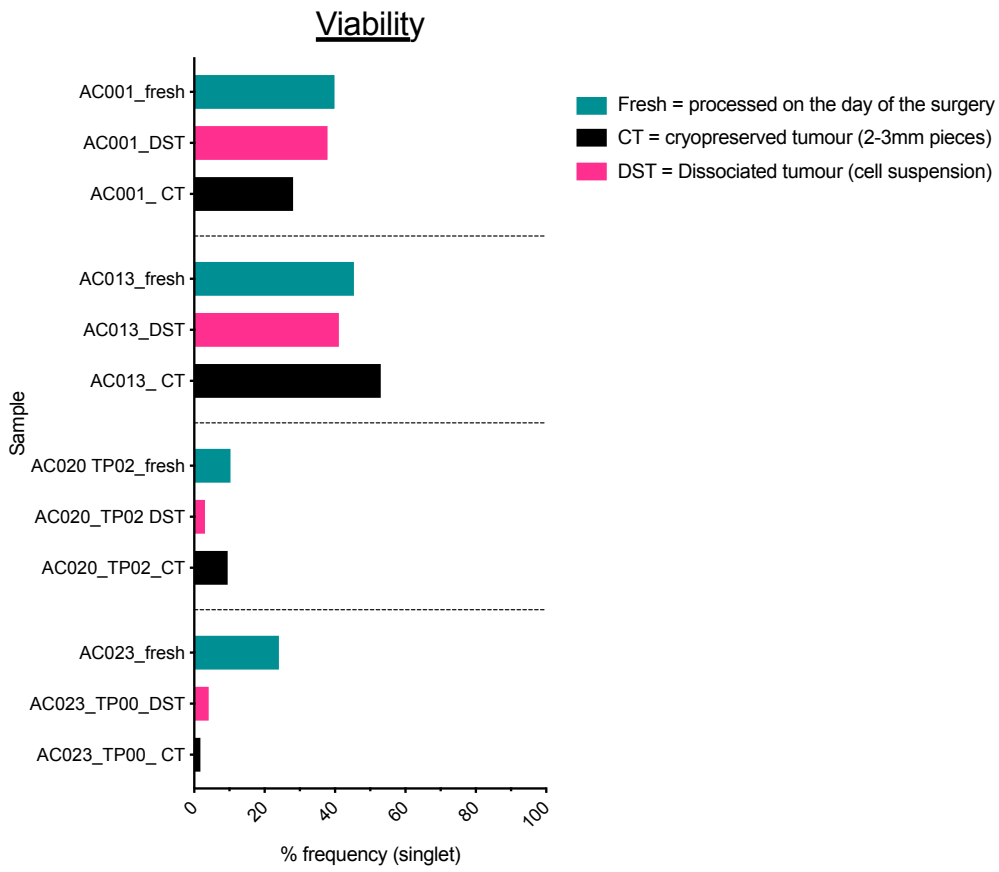
4.1.3.5 Evaluation of tissue storage methods on immune cell frequencies and gene expression profiles

Part of the on-going tissue processing optimisation involved identifying if tumour samples could be stored prior to sorting and analysis in RNA-Seq experiments. The different storage and separation methods were compared to see whether there were any significant differences in either the yield of cells or the detectable gene expression profiles. A small cohort of 6 cases were used (Appendix A.1.6), part of the freshly resected tumour was sorted as previously reported, part was preserved as a cell suspension (DST) and part was cryopreserved as multiple 2mm fragments (CT) In 90% FCS and 10% DMSO (see methods section 2.6.2). The cells/ tissue were then thawed and immune cells sorted as previously described, followed by RNA-Seq. The cell percentages gated on CD45+ % populations are displayed as a dot plots with the standard error of the mean (Figure 38 A). Overall the cell percentages between the different methods are consistent (fresh, CT and DST), a high or low B cell or CD8+ cell percentage is present in the fresh sort, CT and also DST. The individual cases, processing method and immune cell type (CD8, CD4, Macrophage/ activated monocytes and B cell) are also shown as bar plots in Figure 38 C. There were however, exceptions to this with AC066 displaying a larger error of mean for B cells and macrophages (activated monocytes). The tumour volume is likely to impact the consistency of the FACS data when comparing it to samples sorted in smaller aliquots. The viability of the cells was also consistent between the fresh, DST and CT processing methods (Figure 38 B) with the exception of case AC023, where the tissue ischaemic time and large tumour volume may have played a role in the variability.

A


Macrophage of CD45
% Macrophage of CD45+ cells
B cells of CD45
% B cells of CD45+ cells
CD3-ve of CD45
% CD3-ve of CD45+ cells
CD8 of CD45
% CD8 of CD45+ cells
CD4 of CD45
% CD4 of CD45+ cells
CD3 of CD45
% CD3 of CD45+ cells
AC001
AC013
AC020
AC023
AC043
AC066
Sample Id

B


Viability
AC001_fresh
AC001_DST
AC001_ CT
AC013_fresh
AC013_DST
AC013_ CT
AC020 TP02_fresh
AC020_TP02 DST
AC020_TP02_CT
AC023_fresh
AC023_TP00_DST
AC023_TP00_ CT
Sample
% frequency (singlet)
Fresh = processed on the day of the surgery
CT = cryopreserved tumour (2-3mm pieces)
DST = Dissociated tumour (cell suspension)

C

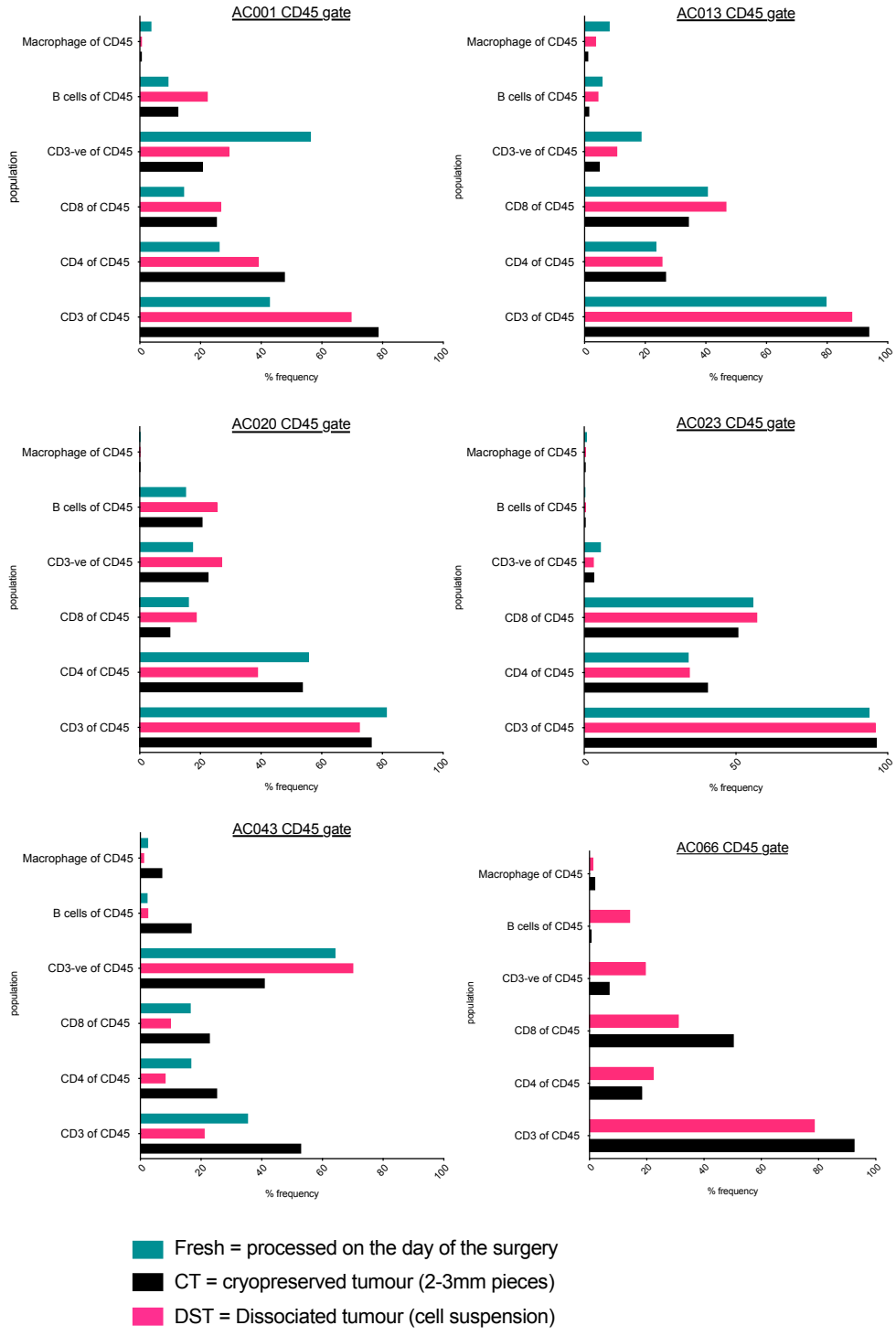


Figure 38. Assessment of processing method on the immune cell frequencies and viability as determined by FACS.

Tumour samples from patients were collected, processed and sorted on the day of surgery (fresh) as well as being cryopreserved as cell suspensions (DST) and 4 x 2-3mm pieces of tumour (CT). (A) Summary dot plot of each case for the cell populations, error bars represent standard error of mean (SEM). (B) The viability of the cells (% live of the singlet gate) using the nuclear stain DAPI also shows a consistency between the different storage conditions, except in the Fresh tissue AC023 sample, where a higher % of cells stained DAPI positive. (C) The cell frequencies for each case and each cell population (CD8, CD4, B cell and macrophages/ activated monocytes) are shown as a percentage of CD45.

Following sorting of the Immune cells, the CD8+ cells were subjected to RNA-Seq and gene expression evaluated, firstly by PCA of the top 4000 and 1000 most variable genes across the different samples (Figure 39 A) and then by visualisation of immune gene expression profiles using a heat map (Figure 39 B). The 4 cases with fresh, DST and CT CD8 T cells are colour coded and labelled with the tissue storage/ processing method, the PCAs illustrate that the matching samples cluster closer to each other than to different cases, using both the top 4000 and 1000 most variable genes. The fresh sample from AC023 displayed the largest variance and distance to its matched pairs. The immune gene expression focusing on markers of functionality and targets of immunotherapy showed consistent expression across all of the different samples, PDCD1 (PD1) is a good example where high and low expression levels were observed between cases, but was consistent within the matching samples of a case (Figure 39 B).

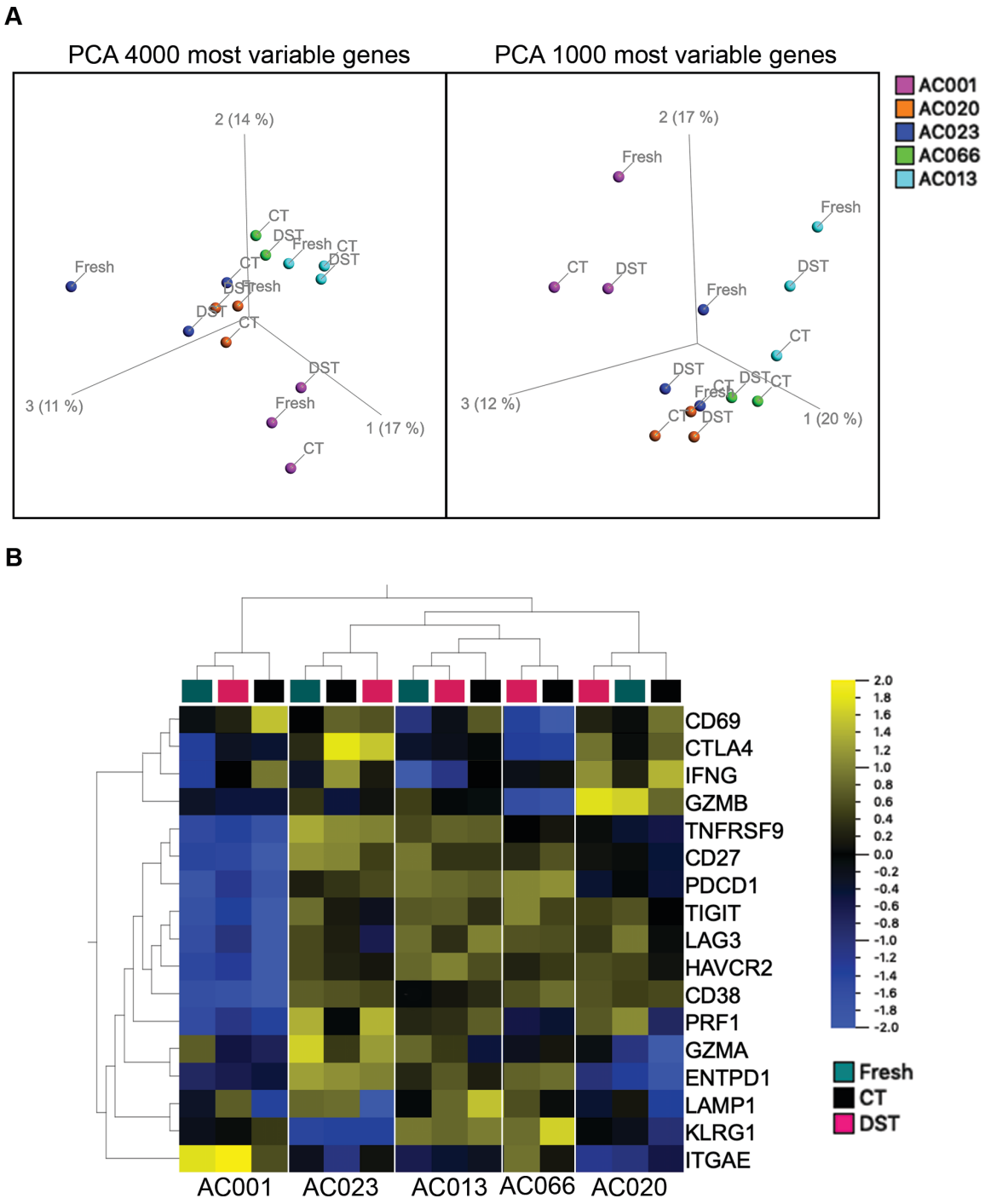


Figure 39. Analysis of gene expression profiles from cells isolated using different processing methods.

(A) To ensure global CD8 gene expression (bulk RNA-Seq) changes hadn't occurred between storage conditions the top 4000 and 1000 most variable genes across the samples and storage conditions were visualised as a PCA (principle component analysis), displaying the most variable genes emphasizes any global changes. Samples cluster on the PCA closely to their replicates. (B) To further assess differences in the CD8 RNA-Seq key immune genes were visualized as a heatmap, again showing good consistency among samples replicates.

4.1.4 Discussion

In order to gain further insight into tumour-infiltrating lymphocytes using transcriptomic approaches, they needed to be routinely isolated from solid tumours. Early testing with magnetic cell isolation techniques (Easy Sep, Stemcell and MACS, Miltenyi) yielded lower purified cell numbers and required sequential isolation steps to obtain multiple cells types (data not shown). The optimisation of cell sorting using Flow cytometry required both an effective tumour dissociation protocol and a functional FACS panel that could be implemented as a robust sorting procedure. A routine procedure for isolating immune cells was developed from existing work and published work (Maecker et al., 2012), this was firstly applied to head and neck tumours, followed by a collaborative expansion into lung cancer. Initial tumour cell suspensions were prone to the formation of cell aggregates; as a result EDTA (2mM) was added to the sorting buffer which alleviated this issue. Cells were sorted directly into Trizol LS (rather than lysis buffer) to lyse cells and preserve RNA. The flow cytometer setup was made structured and routine to ensure good stream setup and correct 4-way stream placement for effective sorting into 1.5ml tubes (SOP in Appendix A2). Tumour cell suspensions were sorted using a FACS panel for the isolation of CD4+ and CD8+ T cells, B cells (CD19+ and CD20+) and macrophages (activated monocytes, CD14+ and HLA-DR+), this was routinely applied to HNSCC and NSCLC samples.

The flow cytometry data shows that the overall immune density can vary greatly depending on the gating strategy with the CD45+ % having an average of 10% of all events; this increases to 23% if only live cells are used. Reassuringly, the overall immune cell percentages for HNSCC match those identified by other groups (Lechner et al., 2017), CD4+ T cells = 47% of T- lymphocytes (CD3+) and CD8 = 29%, with the CD4:CD8 ratio at a 1.5 – 2. The overall infiltration with B cells matched that of the CD8+ T cells. Again, the immune density in NSCLC showed a similar number of immune cells to those stated in the literature (Lechner et al., 2017), (Lizotte et al., 2016) (Stankovic et al., 2018). NSCLC had T cell (CD3+) percentages of similar levels to HNSCC with CD4+ T cells = 51% and CD8 +T cells = 33%, the non-involved lung (NIL) CD4+ T cells = 45% and CD8+ T cells = 32%, also of note were the number of B cells, 27% in tumours and 14% in non-involved lung. A significant difference was seen between tumour and non-involved lung for CD8+, CD4+ and B cells when gated on CD45+ cell, whereas, when gating on CD3, only B cells were significantly different. The next thing was to assess if differences existed between tumour subtypes in HPV(+) and (-) HNSCC and adenocarcinoma and squamous cell carcinoma (NSCLC). In HNSCC, a significant difference in CD4+ infiltrate was found between HPV(+) and (-) tumours. Our previous data looking at TIL rich HSNCC identified a difference in B cells between HPV(+) and (-) tumours, using RNA-Seq and IHC (Wood et al., 2016b). The data here is not significant although the HPV(+) tumours do split into a high B cell and low B cell group. In line with the literature, no

difference was observed in the immune distributions between adenocarcinoma and squamous lung carcinomas (Lizotte et al., 2016).

As part of the on-going tissue storage, samples that couldn't be sorted on the day of surgery were cryopreserved as pieces of tissue (CT) and/or cell suspensions (DST) to ensure that all samples are stored for research applications. This opened the question of how representative preserved tissues are compared to the fresh surgical specimen. In order to address this, a group of 5 tumours were divided into triplicate samples one of which was subjected to flow cytometry and sorting on the day of surgery (fresh), one was cryopreserved as pieces of tissue (CT) and one was disaggregated into cell suspensions (DST) which were analysed. The FACS data shows minimal differences between the three methods, with equivalent percentages of CD8+ and B cell (CD45 gate). The viability was also consistent except for AC023 which was a large tumour with likely issues due to ischaemia. To evaluate the impact on gene expression, CD8+ cells were selected for comparison. PCA analysis of the top most variable 1000 and 4000 genes across the samples revealed that samples grouped mostly by case, except for AC023 (previously identified as an outlier). This indicates that at a global gene expression level, the cases exhibited similar global expression patterns that were not perturbed by the cryopreservation process. This has also been shown at a single cell level (Guillaumet-Adkins et al., 2017). Immune gene expression was also assessed and shows similar expression levels between the different storage methods in CD8+ T cells, indicating that the cryopreservation methods is not affecting the gene expression readout.

4.1.5 Conclusions

Isolation of immune cells from tumours enables high resolution studies to be undertaken in different patient cohorts. A tumour dissociation protocol was developed in order to generate single cell suspensions for flow cytometry in HNSCC and NSCLC. In addition to this, a flow cytometry panel for the isolation of CD8+, CD4+, B cells and macrophages (activated monocytes) was implemented. A standardised protocol using flow cytometry (BD FACS ARIA) was then implemented for routine and consistent immune cell isolation in HNSCC and NSCLC. This process has subsequently been incorporated into two clinical trials in HNSCC (HARE-40 and AMG319; Appendices A2), where immune cell transcriptomics is being assessed under different treatment conditions. The analysis of the Flow cytometry data shows that the immune cell percentages fit with what is currently published. As part of the tumour sample collection strategy (CT and DST), it was also important to show that smaller biopsies and cell suspensions reflect the larger samples processed on the day of surgery. In addition to this, it was necessary to demonstrate that the cryopreservation process had not negatively impacted the gene expression profiles. This was done in a small cohort of samples and focused only on CD8+ cells, meaning additional cell types (CD4+,

B cell and macrophages/ activated monocytes) and tumour types would be required to answer this question fully. However, the data here showed a high level of consistency both at the cell frequency and gene expression level.

4.2 Impact of CD8+ Tissue resident memory cells in Head & neck and lung cancer

4.2.1 Introduction

Response to immunotherapy is highly heterogeneous and not all patients benefit from such approaches, with 1/3 of patients showing little or no response (Sharma and Allison, 2015). This opens the question of what defining features lead to a beneficial clinical response. The main cellular targets of immunotherapy are cytotoxic T lymphocytes (CTLs), primarily CD8+ T cells (Tumeh et al., 2014); the main aim of immunotherapies is to "release" them from a state of exhaustion. In many solid cancers, the density of CD8+ T cells is positively linked to survival, as previously outlined (Galon et al., 2006), whether the immune density is also linked to a qualitative difference is less well known. During the immune response to tumours, T cells become functionally unresponsive ("exhausted"), this is in part due to chronic activation leading to a specific phenotype, these states are also seen in chronic viral infections (Wherry, 2011a). Classical markers of T cell "exhaustion" are CTLA4, PD1, TIM-3 and LAG3; expression of these ultimately leads to a reduction in the cytolytic function of the CD8 T cells (Pardoll, 2012, Barber et al., 2006).

The success of anti-PD1 therapy in management of cancer has led to its widespread use, however significant adverse events can occur (Schachter et al., 2017, June et al., 2017), leading to the search for more specifically targetable features of tumour-reactive T cells. Transcriptomic approaches allow fine mapping of the immune cells and the resulting molecular features in different patient stratification groups. To characterise the CD8+ T cell infiltrate in solid tumours, the Immune cells were isolated (as previously outlined in section 4.1) from HNSCC and NSCLC, as well as from the non-involved background lung. Figure 40 A shows a schematic of the anatomic localisation of the tissue collection. In order to establish what unique features may exist in tumour CD8+ T cells, their transcriptional activity was assessed and differential gene analysis was performed. The differentially expressed genes between Lung tumour-derived CD8+ T cells compared to CD8+ T cells from non-involved lung are shown in a heatmap together with the tSNE plots in Figure 40 B-D. The transcriptional profiling revealed a distinct gene expression pattern in the tumour-derived CD8+ cells that was observed in both NSCLC and HNSCC. This was separate from that of the CD8+ T cells from non-involved lung, as seen in the heatmap and also the clustering dendrogram (Figure 40 B and C). A signature of markers indicating elevated T cell exhaustion and anergy was detected. Markers of T cell dysfunction (PDCD1, HAVCR2, LAG3 and CTLA4) were differentially expressed between non-involved lung and tumour-derived CD8+ T cells (Ganesan et al., 2017). However, evidence of greater clonal expansion in PD1 positive tumour

infiltrating CD8+ T-cells has also been identified (relative to PD1 negative CD8+ T cells)(Pentcheva-Hoang et al., 2007, Gros A, 2014). This highlights a dual role of CD8+PD1+ cells in the tumour. In depth analysis of the transcriptome in tumour-derived CD8+ T cells identified features of T cell activation and clonal expansion (41BB expression and higher numbers of unique TCRs). Included in the differential signature was a marker of T cell tissue residency, CD103 (ITGAE), an integrin that helps retain T cells in the peripheral tissues by binding to epithelial E-cadherin (Schenkel and Masopust, 2014).

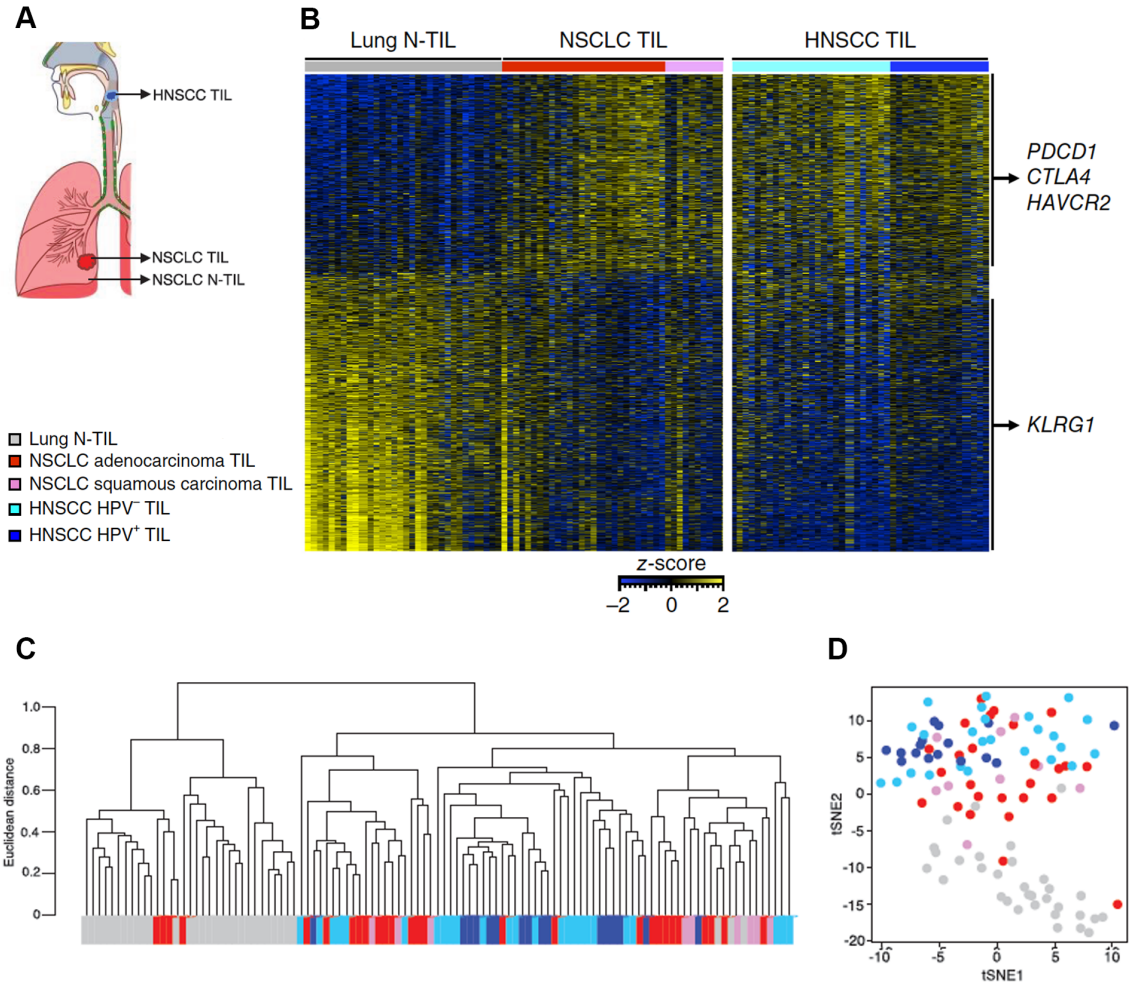


Figure 40. The core transcriptional profile of CD8 T cells from HNSCC, NSCLC and non-involved lung (NIL).

(A) Schematic representation of the tissue sites for the cohort. (B) RNA-Seq analysis of genes differentially expressed (DESeq2 analysis; Benjamini-Hochberg test, FDR<0.05, FC >1.5) between Lung tumour CD8 (n=36) and non-involved lung CD8+ T cells (n=32) (rows = genes , columns = patient tumours), presented as row-wise z-scores of normalized read counts in CD8+ TILs. The same genes were visualised in the HNSCC CD8+ T cells (n = 41), similarity between tumour CD8+ TILs was observed with a different transcriptional profile in the NIL. Hierarchical clustering analysis (Ward average linkage, Pearson's correlation distance metric) grouped the tumour CD8+ T cells with minimal outliers (C). This was further demonstrated with tSNE (perplexity = 6, iterations = 1000), where the tumour CD8+ T cells cluster together separate from the NIL (D). Adapted from (Ganesan et al., 2017).

4.2.2 Objectives

A gene expression pattern for "exhausted" CD8+ Tissue resident memory T cells (CD8+CD103+) was identified in lung tumours (Ganesan et al., 2017). However, this needed to be confirmed at a protein level using flow cytometry, in lung tumour, non-involved lung as a control tissue and with peripheral blood as an additional control sample. The impact of tissue- resident memory T cells (CD103+) on survival will also be evaluated by IHC on a retrospective cohort of NSCLC and HSNCC. CD8+ T cells from HPV(+) and (-) tumours will also be evaluated for evidence of a viral-specific gene expression profile in the virally-driven tumours. The core transcriptomic differences between the clinically relevant CD8+PD1+ and PD1- T cells will also be assessed in a small cohort of samples.

4.2.3 Results

4.2.3.1 Evaluating the tissue residency and exhaustion features of CD8+ T cells in cancer

4.2.3.1.1 Frequency of T_{RM} and exhaustion markers in tumours and control tissue using flow cytometry

A cohort (Appendix A.1.7) from whom matched tumour, non-involved lung (control) and peripheral blood (control) was obtained, was assessed on the day of the surgical resection, using a FACS panel that covers CD8 (CD8A), PD1 (PDCD1), CD103 (ITGAE) and 41BB (TNFRSF9). The CD8+ T cells (gating = singlet>live>CD45>CD3>CD8) were assessed for the frequency of CD103 (Figure 41), PD1 (Figure 42) and 41BB positive cells (Figure 43) in the different tissue compartments. Representative plots for the expression and frequency of the tissue residency marker CD103 are shown in Figure 41 A and B, with a bar chart summary (geometric mean and 95% confidence intervals) of the percentage of CD8+ CD103+ T cells in the three tissue compartments. The peripheral blood control has a very low number of CD8+ CD103+ T cells (1-2%) and is consistently low across the cohort, however they are not absent and represent an intriguing population that may have originated in the tumour. The non-involved lung (NIL) and tumour display a heterogeneous number of CD103+ cells with significantly higher numbers in the tumour ($p=0.0059$) relative to the NIL (Figure 41 C).

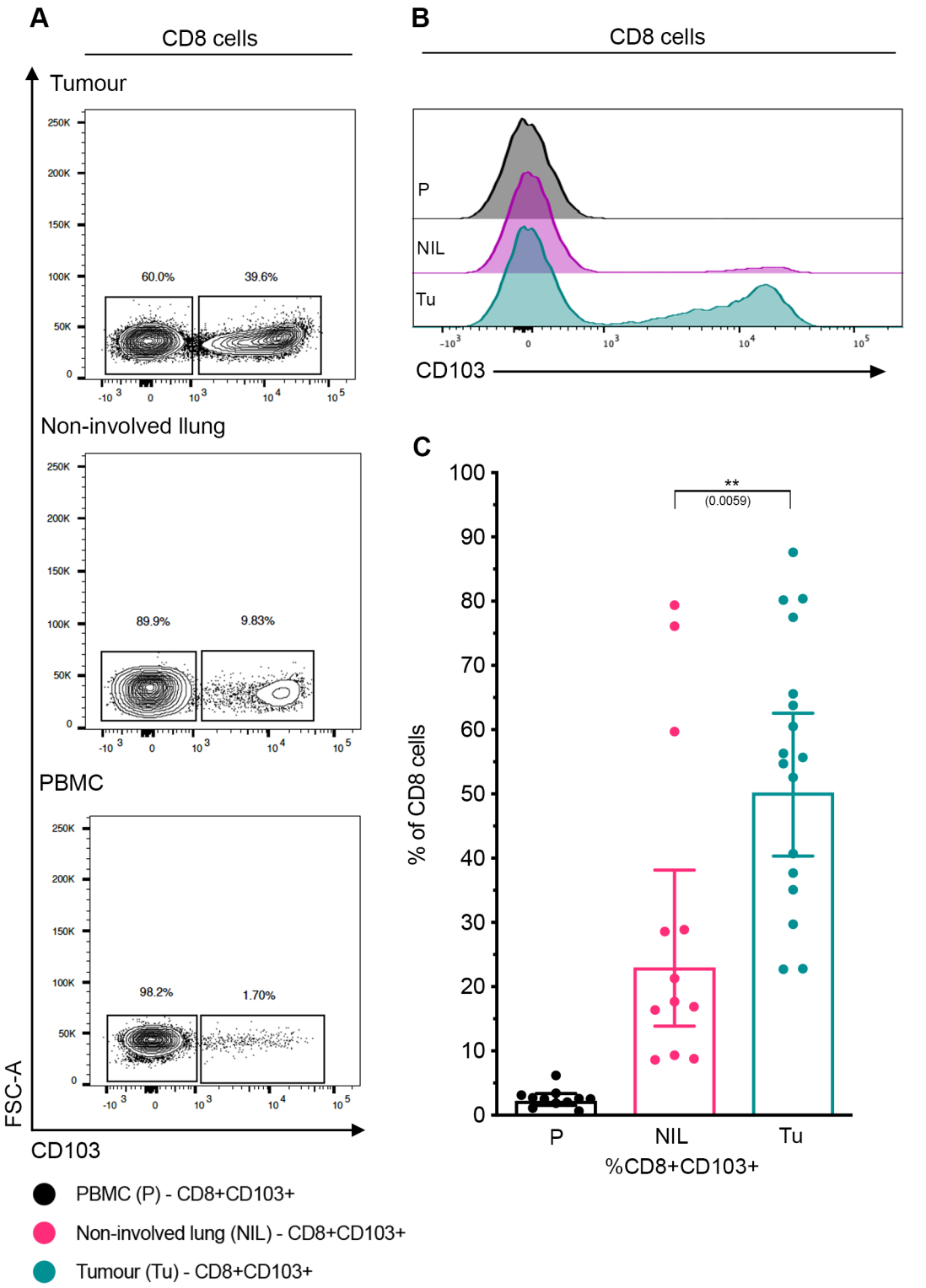


Figure 41. Evaluation of CD103 protein expression in CD8 T cells.

Flow cytometric analysis of CD103 (ITGAE) was carried out on PBMC, non-involved lung (NIL) and tumour (NSCLC). (A) Representative plots for the cell frequencies in PBMC, NIL and tumour, (B) Histogram for the same representative cases. (C) Bar and dot plots (geometric mean and 95% confidence interval) displaying the distribution of CD8+CD103+ T_{RM} cell frequencies in PBMC , Nil and tumour. Tumours exhibit a significantly higher number of CD8+CD103+ T_{RM} T cells (p=0.0059; Mann-Whitney test).

The finding of high CD8+CD103+ T_{RM} cells in the tumour led to the evaluation of whether the T cell "exhaustion" marker PD1 varied between the CD8+CD103+ and CD103- populations. The expression and distribution of PD1 between the CD103+ and CD103- CD8+ T cells is shown in Figure 42 A and B, where representative plots show a higher expression and frequency of PD1 in tumour T_{RM} (CD8+CD103+) cells compared to the NIL. A higher frequency of CD8+CD103+PD1+ cells was observed in the tumour compared to the NIL (p=0.0006). In addition, the frequency of PD1+ T cells was higher in the tumour-infiltrating CD103+ CD8+ T cells compared to the CD103- CD8+ cells (p=0.0011) (Figure 42 C). In PBMC, the CD8+CD103+ cells expressed very little PD1, however the CD8+CD103- T cells in peripheral blood expressed equivalent levels of PD1 to CD8+CD103- T cells in the normal lung and tumour.

This showed that the CD8+CD103+ T_{RM} cells were expressing high levels of PD1 in the tumour, as was seen in the transcriptomic data. Another feature identified from the transcriptomic analysis of tumour CD8+ T cells was the increased expression of 41BB, a marker of TCR engagement. This displayed a more heterogenous expression profile between the different tissue compartments, shown in Figure 43 A-C (PBMC, NIL and tumour), however a significant difference was identified in the number of CD8+CD103+41BB+ cells in the tumour relative to the NIL (p=0.0005). The tumour also displayed a higher number of 41BB+ cells between the CD103+ and CD103- populations (p=0.0015). The number of CD8+CD103-41BB+ cells was similar in PBMC and NIL, with the tumour displaying lower cell numbers.

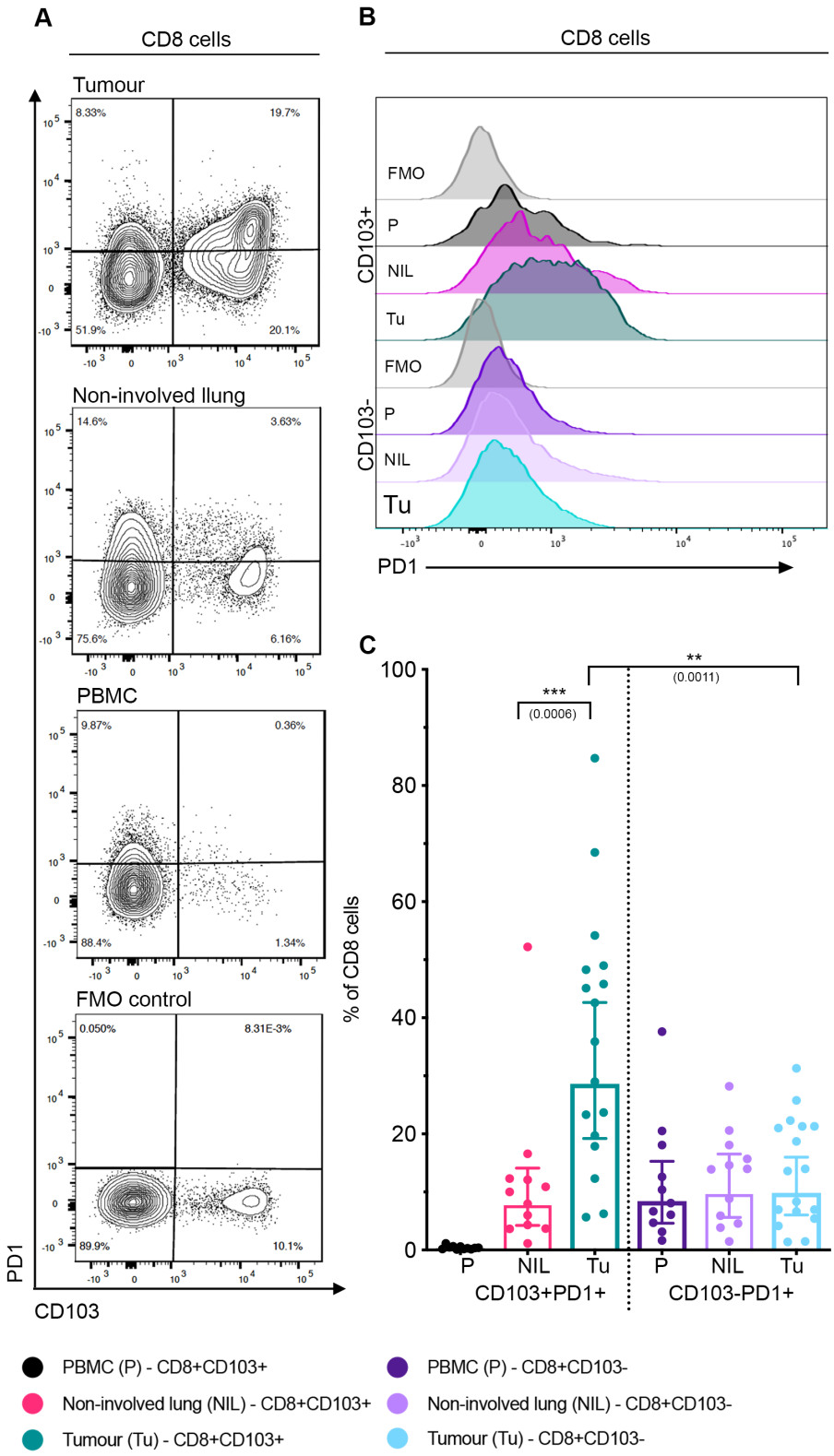


Figure 42. PD1 exhaustion marker cell frequencies in CD8 T_{RM} T cells.

Flow cytometric analysis of PD1 cell numbers in CD8 T_{RM} (CD103+) and non- T_{RM} (CD103-) in PBMC, non-involved lung (NIL) and tumour (NSCLC). (A) Representative plots for the cell frequencies in PBMC, NIL and tumour. A fluorescence minus one (FMO) control tube with the removal of the PD1 antibody was used to set the cut-off point for the gating, (B) Histogram for the same representative case. (C) Bar and dot plots (geometric mean and 95% confidence interval) displaying the distribution of CD8+CD103+ PD1+ T_{RM} and non- T_{RM} cell frequencies in PBMC, Nil and tumour. Significantly higher PD1 cell numbers were observed in the tumour CD8+ T cells (Mann-Whitney test).

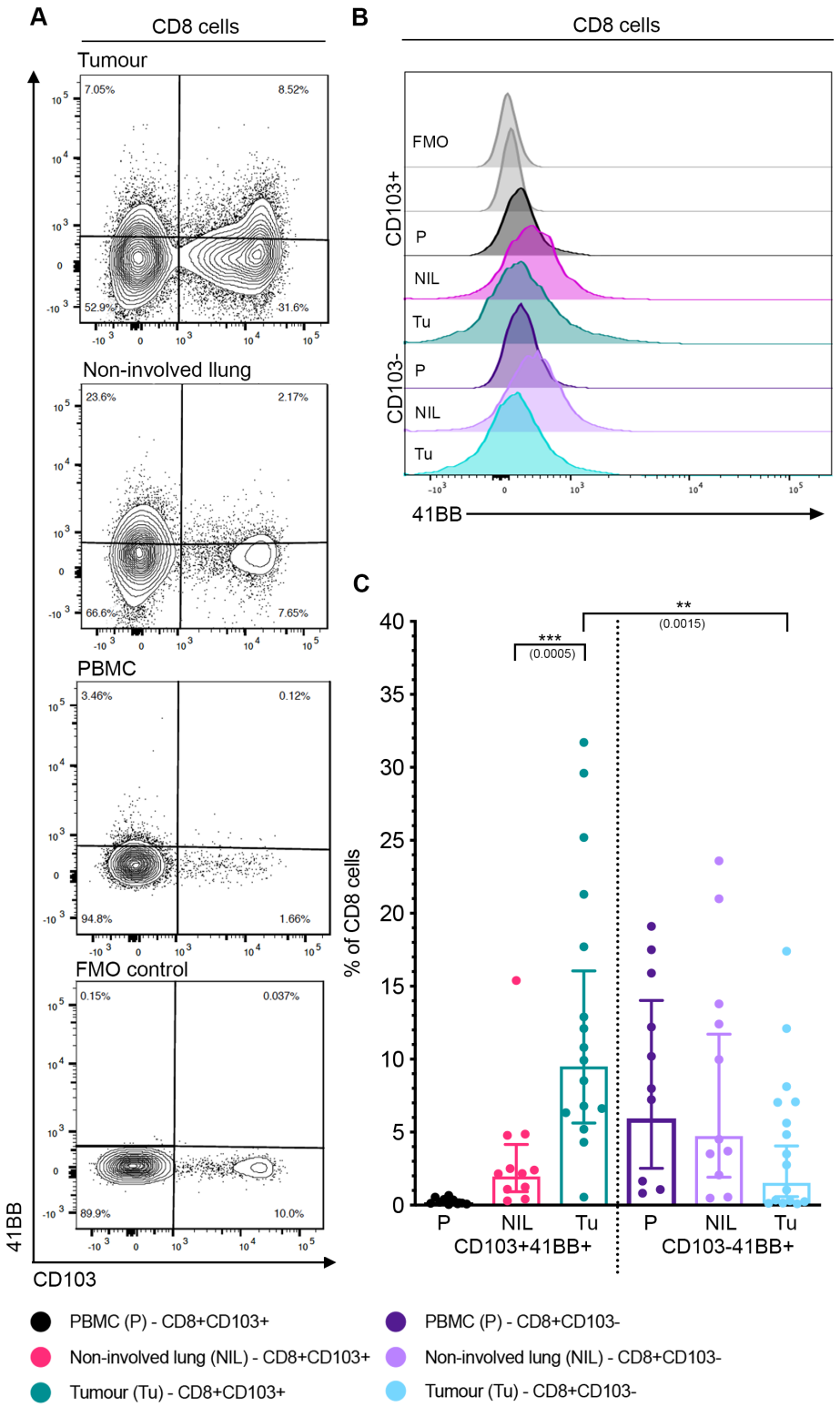


Figure 43. Analysis of 41BB cell activation marker in CD8 T_{RM} T cells.

Flow cytometric analysis of 41BB+ cell numbers in CD8+ T_{RM} (CD103+) and non- T_{RM} (CD103-) in PBMC, non-involved lung (NIL) and tumour (NSCLC). (A) Representative plots for the cell frequencies in PBMC, NIL and tumour. A fluorescence minus one (FMO) control tube with the removal of the 41BB antibody was used to set the cut-off point for the gating, (B) Histogram for the same representative case. (C) Bar and dot plots (geometric mean and 95% confidence interval) displaying the distribution of CD8+CD103+ 41BB+ T_{RM} and non- T_{RM} cell frequencies in PBMC, Nil and tumour. Significantly higher 41BB cell numbers were observed in the tumour CD8+ T_{RMs} (Mann-Whitney test).

4.2.3.1.2 Mapping the proliferative state of CD8+CD103+ T_{RM} T cells using Ki67

Evidence of a higher number of 41BB+ cells and an increase in the number of unique TCR (CDR3) repeats that are indicative of a T cell expansion (Ganesan et al., 2017), led to the interrogation of the proliferative state of the cells using Ki67 as a marker. This was performed using flow cytometry on a separate cohort of n=10 tumour samples (Appendix A.1.7; key markers = Ki67, PD1, CD103 and 41BB), where the frequency of Ki67 cells was analysed relative to the tissue type and T_{RM} status (CD103+ and CD103-). The representative FACS plot and histogram (Figure 44 A and B) shows a tumour with high numbers of CD8+CD103+Ki67+ T_{RM} cells. The number of CD8+CD103+Ki67+ cells in the NIL was extremely low compared to the tumour (p=0.0007). The comparison of the tumour CD8+ T_{RMs} (CD8+CD103+) with non- T_{RM} (CD8+CD103-) (Figure 44 C) cells again showed a significant difference, with the tumour T_{RM} cells displaying higher proliferating Ki67+ cell numbers (p=0.0433). The PBMC and NIL non-T_{RM} cells numbers displayed low numbers of Ki67+ cells.

In order to understand the co-expression of the markers PD1, 41BB, Ki67 and CD103, the flow cytometry data was visualised as a tSNE (t-distributed stochastic neighbour embedding) plot (Figure 45). The CD8+ T cells were randomly sampled to yield 10,000 cells per sample, then concatenated together and tSNE analysis performed using the limited number of markers (PD1, 41BB, Ki67 and CD103). The expression of CD103 (T_{RM}) clearly divides the CD8+ cells into two groups, most of the PBMC and NIL have less expression of CD103, as has been previously identified using conventional flow analysis (Figure 41). Cells from the tumour (teal) and NIL(pink) dominate the CD8+CD103+ signal in the tSNE. Two regions within the CD8+CD103+ cells were almost exclusively derived from the tumour (not shared with the NIL). The expression of PD1 was most prominent in this region which is also broadly divided by a small cluster of 41BB+ and a ki67+ cells (arrow). It shows that CD8+ T cells can co-express CD103+PD1+41BB+Ki67+ and that only a small subset were actively dividing (Ki67+) and display TCR engagement (41BB+).

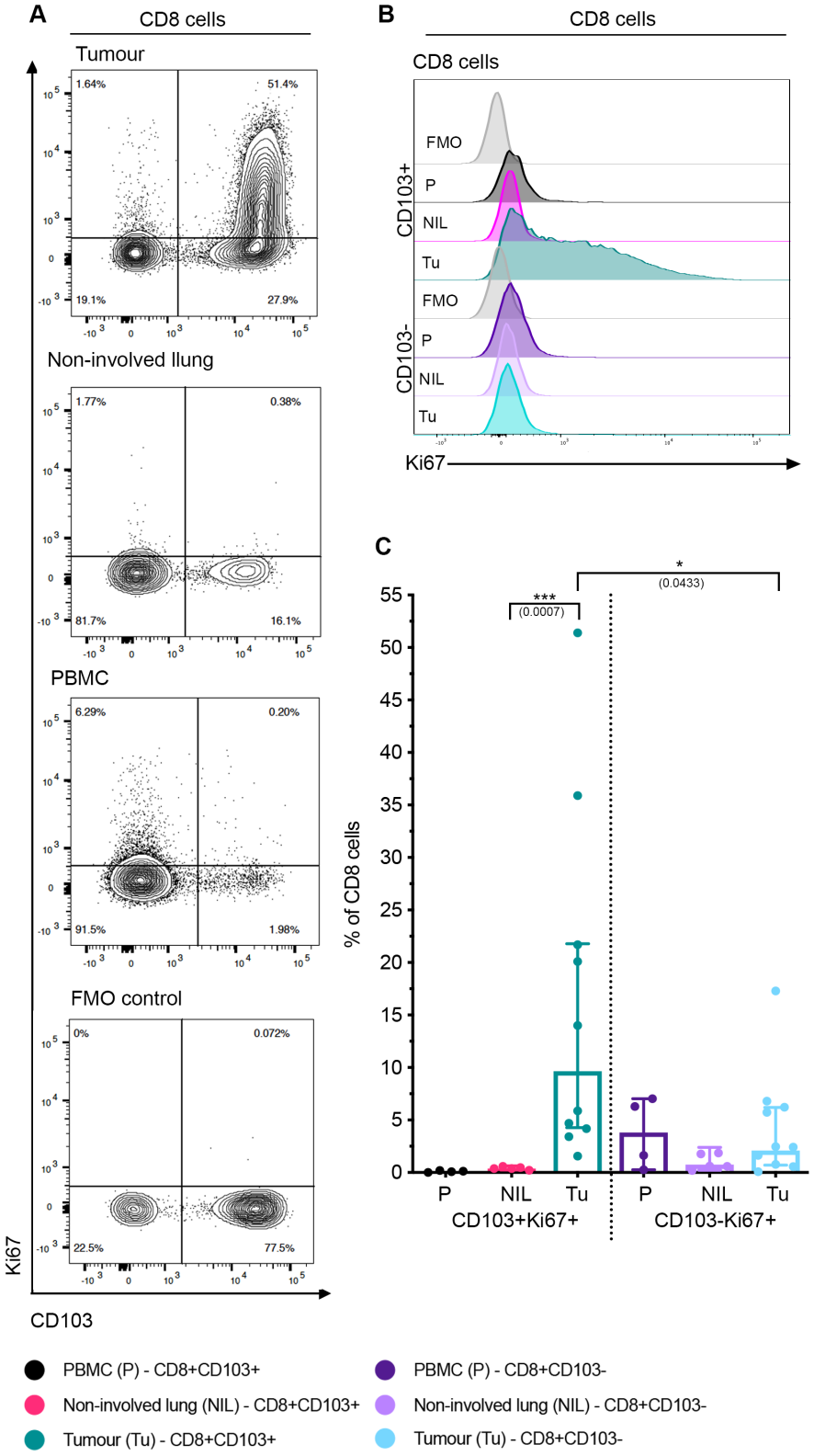


Figure 44. Proliferative capacity of the CD8 T_{RMs} in NSCLC.

Intracellular analysis of the nuclear proliferation marker Ki67 in CD8 T_{RM} (CD103+) and non- T_{RM} (CD103-) in PBMC, non-involved lung (NIL) and tumour (NSCLC). (A) Representative plots for the cell frequencies in PBMC, NIL and tumour. A fluorescence minus one (FMO) control tube with the removal of the Ki67 antibody was used to set the cut-off point for the gating, (B) Histogram for the same representative case. (C) Bar and dot plots (geometric mean and 95% confidence interval) displaying the distribution of CD8+CD103+ Ki67+ T_{RM} and non- T_{RM} cell frequencies in PBMC, Nil and tumour. Higher numbers of proliferating Ki67+ cells were identified in the tumour CD8+ T_{RM} (Mann-Whitney test).

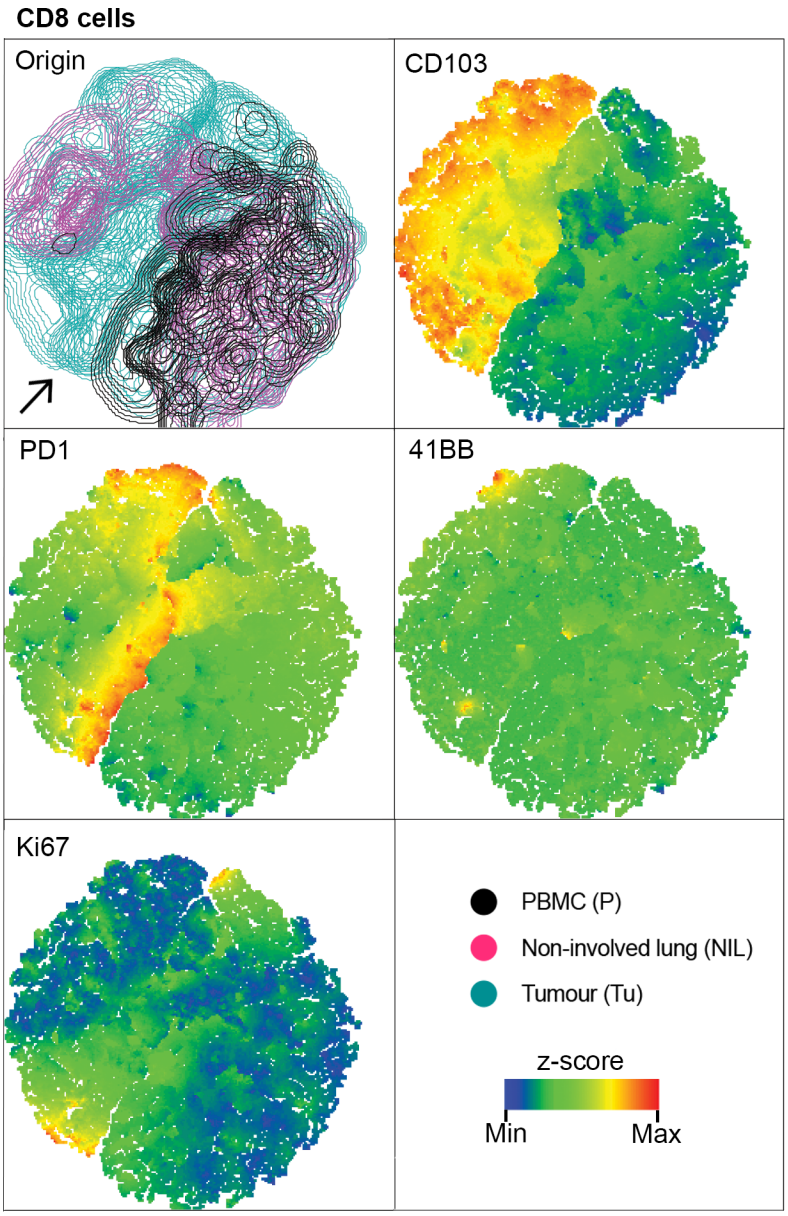


Figure 45. Co-expression of T cell activation and exhaustions markers in CD8 T_{RMs}.
Cytometric analysis of CD8+ cells from n= 3 PBMC, n= 6 NIL and n=10 tumour, these were reduced to 10,000 CD8+ cells and combined across all samples (Flowjo 10). The expression (z-score, blue low expression and red high expression) of the surface markers CD8, CD103, PD1 and 41BB and the nuclear proliferation marker ki67 were visualised as 2D tSNE plots (perplexity = 6, iterations = 800). The tissue of origin is shown as PBMC (black contours), Tumour (teal contours), and NIL (pink contours), the expression of the markers can then be seen across the different tissue origins. The origin of the CD8 cells (PBMC, Tumour and NIL) reveals unique regions on the tSNE plot in the tumour CD8+ T_{RMs} (teal). Co-expression of the activation (41BB) and proliferation (ki67) markers and PD1 exhaustion makers were present in the CD8+ T_{RMs} in the tumour (teal), the arrow marks an area where tumour CD8+ T_{RMs} exhibit co-expression. This demonstrates presence of antigen engaged and proliferating CD8+CD103+PD1+41BB+Ki67+ cells in the tumour, that are absent in the background lung and PBMC controls.

4.2.3.2 Comparison of CD8+ T cells derived from virally driven HPV(+) and viral independent (HPV-) head and neck cancer.

The analysis of immune dense HPV(+) and (-) HSNCC whole tumour RNA-Seq revealed that the immunological T cell signature was very consistent between the two tumour sub-types. To assess this further, the CD8+ T cells from HPV(+) and (-) tumours were compared with the aim of revealing any differences between virally-driven cancer relative to a viral-independent cancer. A sample size of n=12 HPV(+) was compared to n=20 HPV(-) CD8+ T cells using DESeq2.

4.2.3.2.1 Differential gene analysis between CD8+ T cells from HPV(+) and (-) tumours

A limited number of differentially expressed genes were identified (13 genes with an FDR <0.1 and FC of >1.5), these were visualised as a hierarchically clustered heatmap (Figure 46 A). The HPV(-) CD8+ T cells have a higher level of most of the genes except TSPAN33, which displayed very low transcripts per million gene counts (<10 TPM) and is likely a spurious result. Gene ontology analysis recovered no significant biological associations with the limited number of genes. However, in order to try and rule out these genes being linked to a viral response, they were cross-referenced with gene ontology terms linked to viral responses, and yielded no matches (GO-0009615 response to virus n=295 genes; GO-0051607 defence response to virus n=214 genes and GO-0098586 cellular response to virus n=47 genes). The genes from the viral-specific gene ontologies (n=288) were visualised as a heatmap, no clustering or grouping of the samples by HPV status was observed (Figure 46 B). The same was seen when tSNE was performed using the 288 genes where 3 clusters of CD8+ T cells formed based on gene expression similarity, but this was not linked to the HPV status of the samples (Figure 46 C).

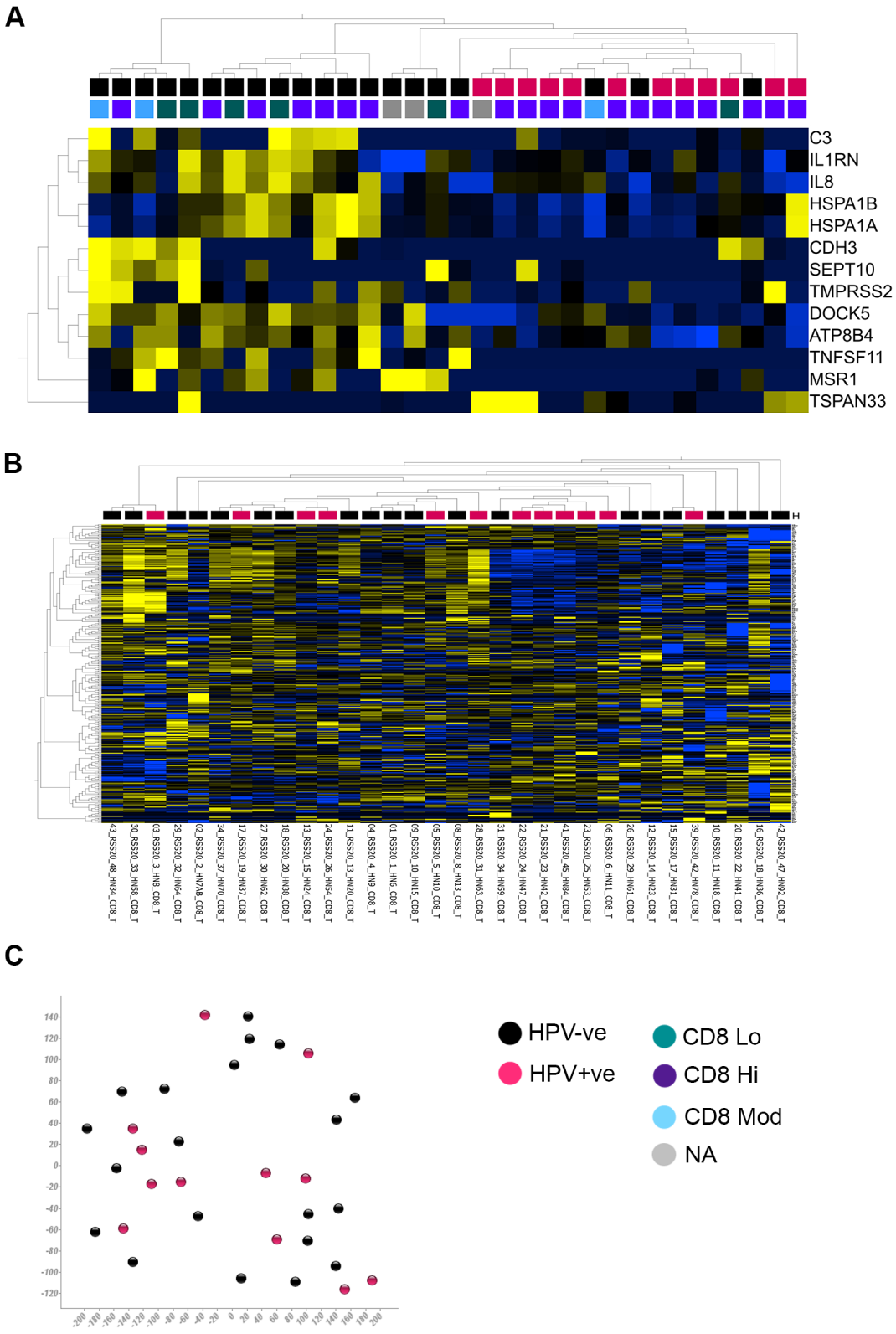


Figure 46. Gene expression differences between HPV(+) and (-) CD8 cells in HNSCC.

DESeq2 comparison (FDR<0.1, FC>1.5) of HPV(+) (n=12) and (-) (n=20) revealed few differentially expressed genes (n=13). (A) These are visualised as a hierarchically clustered (Person's correlation distance metric, Ward-average linkage) heatmap presented as row-wise z-scores of normalized read counts in CD8. (B) Gene ontologies GO-0009615 response to virus; GO-0051607 defence response to virus and GO-0098586 cellular response to virus (total 288 genes)

Visualised as a heatmap show little separation between HPV(+) and (-) CD8 cells. (C) tSNE (perplexity = 6, iterations =1000) again shows little grouping of the CD8 cells by HPV status.

4.2.3.3 Functional differences between PD1+ve and PD1-ve CD8+ T cells

In order to assess the finding that CD8+ T_{RMs}cells (CD8+CD103+) express markers of cytolytic function, proliferation as well as high levels of negative immune regulators like PD1, a small cohort of CD8+ T cells split into PD1+ and PD1-, were isolated by flow cytometric sorting (Appendix A.1.8). They were evaluated at the transcriptomic level to assess whether the features identified in the flow cytometry and bulk CD8 RNA-Seq were again present in cells subcategorised into PD1+ and PD1-. This also had the potential to reveal unique features and biomarkers specific to the PD1+ and PD1- CD8+ groups. Using DESeq2, differentially expressed genes (DEG) were identified between PD1+ and PD1- CD8+ T cells; genes with FDR<0.05 (false discovery rate or q-value) and FC (fold change) >1.5 were considered significant. A total of 133 genes were expressed to a greater extent in the CD8+PD1+ cells, with 147 genes expressed to a lesser extent in the CD8+PD1- cells relative to the CD8+PD1+ cells. The DEGs were visualised as a hierarchically clustered heatmap (Figure 47 A) in which, reassuringly, a clear difference between the two groups was observed, and in line with our previous findings, the genes PDCD1, HAVCR2, ITGAE, MKI67, GZMB and IFNG were higher in the CD8+PD1+ cells relative to the CD8+PD1- cells. The expression of negative regulators of the immune response were also visualised (Figure 47 B). This shows that the CD8+PD1+ cells co-express other negative regulators of the immune response. Key negative regulators, activation and functional markers of the immune response are represented as dot plots in Figure 47 C. All of the genes (PD1, TIM3, 41BB, Ki67, CD103, IFNG and GZMB) have a significant difference (FDR <0.05; range of q-values=1.1^{e-6} to 0.05) between PD1+ and PD1- CD8+ T cells (DESeq2), the CD8+PD1+ cells showing higher expression in all but one case (GZMB).

In order to gain further insight to the genes that were differentially expressed between PD1+ and PD1- CD8+ T cells, gene ontology analysis was performed on the greater and lesser expressed genes between the two groups (133 and 147 respectively). Genes expressed to a lesser extent were not significantly associated with any GO biological processes or pathways, however n =10 genes (SELL, NT5E, ITGA6, TNFRSF10D, CD320, S1PR1, MST1R, TLR1, TSPAN7, CCR7) linked to the tumour necrosis super family (TNFS) were identified (Table 24). The top 25 gene ontologies associated with the genes expressed to a greater extent in CD8+PD1+ cells were broadly linked to cell cycle progression and mitotic cell division. However, within this there were several immune-specific terms (GO:0002520 - immune system development; GO:0032652 - regulation of interleukin-1 production and GO:0050715 - positive regulation of cytokine secretion). The gene families related to tumour necrosis factor superfamily further identified key immune gene differences between PD1+ and PD1- CD8+ T cells (PDCD1, TNFRSF9, LAG3, ITGAE, TFRC, SIRPG, BMPR1A, CXCR6, CCR1, CCR5, ALCAM and VCAM1).

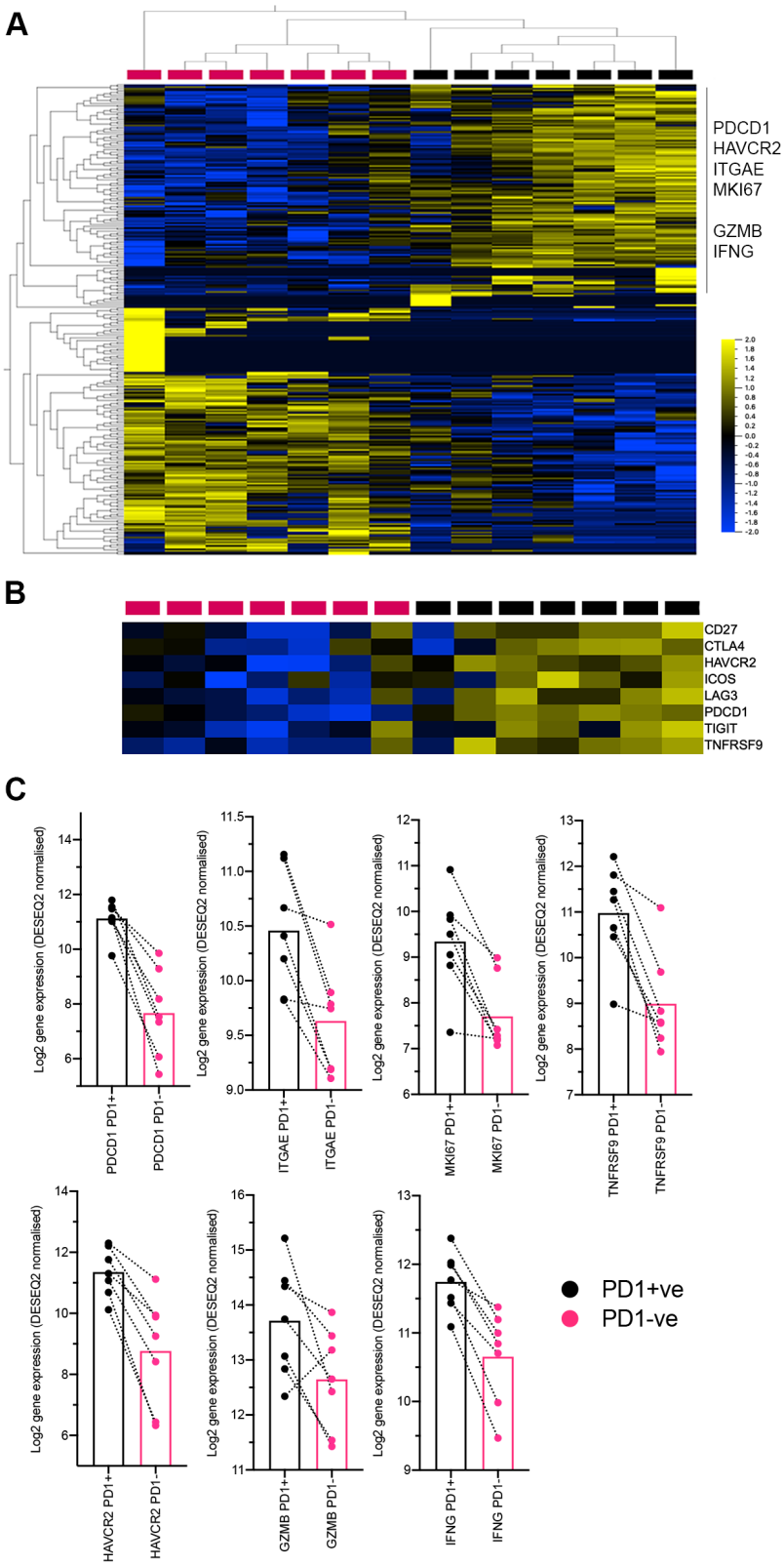


Figure 47. Transcriptional profile of PD1 Hi tumours.

(A) DESeq2 comparison (FDR<0.1, FC>1.5) of PD1 Hi (n=7) and PD1 Lo (n=7) CD8+ T cells revealed 143 greater expressed genes and 133 lesser expressed. These are visualised as a hierarchically clustered (Person's correlation distance metric, Ward-average linkage) heatmap presented as row-wise z-scores of normalized read counts. (B) Immunotherapy targets in PD1 Hi and Lo tumour CD8 cells. (C) Dot plots showing geometric mean and 95% confidence interval for key targets identified in the CD103 Hi CD8 cells (PDCD1, ITGAE (CD103), MKI67 (Ki67), TNFRSF9 (41BB) , HAVCR2 (TIM3), GZMB, and IFNG.

To further evaluate the link between PD1+ cells and CD103+ T_{RMs}cells, the genes expressed to a greater extent in the CD8+PD1+ cells were overlapped with the genes expressed to a greater extent in the CD8+CD103+ cells (n= 442 genes; (Ganesan et al., 2017)). An overlap of 60 genes were identified that contain many of the key functional markers of activation and proliferation (GZMA, GZMB, PRF1, IFNG and MKI67) as well as negative regulators of the immune response (HAVCR2, LAG3 and CTLA4) (Figure 48).

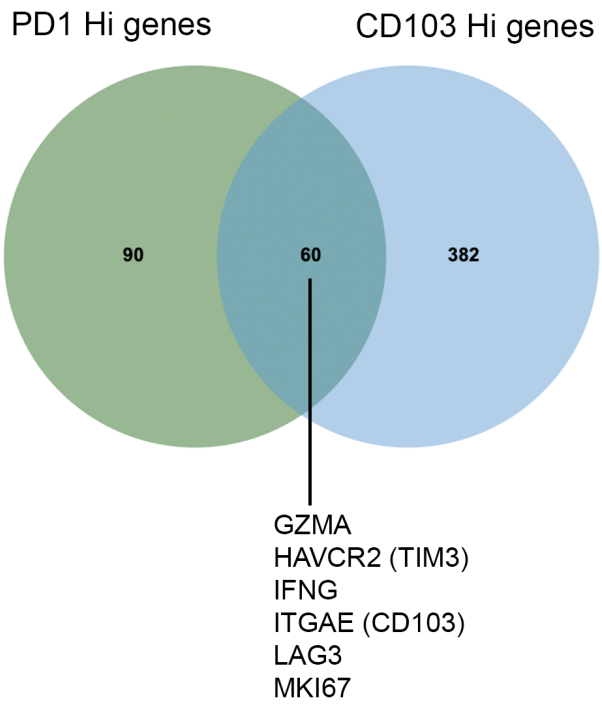


Figure 48. Overlap of genes expressed to a greater extent in CD8+PD1+ and CD8+CD103+ T cells. Overlap between the differentially expressed genes in CD8+CD103+ cells and CD8+PD1+ cells, 60 genes are shared. The RNA-Seq of CD8+PD1+ and CD8+PD1- T cells (n=7 patients) mirrors the findings from the CD8+CD103+ and CD8+CD103- comparison, where high levels of immune check point expression and cytolytic function are co-expressed.

Genes with upregulated expresison (PD1+ve compared to PD1-ve)

Category	ID	Name	q-value FDR	No. of hits	Total GO size	Pecentage	Hit in Query List
GO: Biological Process	GO:0000278	mitotic cell cycle	2.14E-10	32	1016	3.15	NUSAP1,CDKN2A,CDKN2B,CDKN3,KIF18A,PTTG1,CENPA,UBE2C,SMC4,MCM2,SKA2,STMN1,MCM10,PRC1,RRM2,MELK,GFI1,E2F8,CCNA2,CEP55,TYMS,AURKA,LMNB1,TPX2,CTDP1,BTG3,SKA3,CDCA5,PMF1,TOP2A,WDR62,KIF2C
	GO:1903047	mitotic cell cycle process	4.27E-10	30	931	3.22	NUSAP1,CDKN2A,CDKN2B,CDKN3,KIF18A,PTTG1,CENPA,UBE2C,SMC4,MCM2,SKA2,STMN1,MCM10,PRC1,RRM2,MELK,GFI1,CCNA2,CEP55,TYMS,AURKA,LMNB1,TPX2,CTDP1,SKA3,CDCA5,PMF1,TOP2A,WDR62,KIF2C
	GO:0022402	cell cycle process	4.31E-10	36	1385	2.60	RAD51,NUSAP1,CDKN2A,CDKN2B,CDKN3,KIF18A,PTTG1,TP53INP1,CENPA,UBE2C,SMC4,MCM2,SKA2,STMN1,MCM10,PRC1,RRM2,MELK,GFI1,E2F8,CCNA2,CEP55,TYMS,IFNG,AURKA,LMNB1,TPX2,CTDP1,MKI67,SLBP,SKA3,CDCA5,PMF1,TOP2A,WDR62,KIF2C
	GO:0007049	cell cycle	4.23E-09	39	1766	2.21	RAD51,NUSAP1,CDKN2A,CDKN2B,CDKN3,KIF18A,PTTG1,TP53INP1,CENPA,UBE2C,HJURP,SMC4,MCM2,SKA2,STMN1,MND1,MCM10,PRC1,RRM2,MELK,GFI1,E2F8,CCNA2,CEP55,TYMS,IFNG,AURKA,LMNB1,TPX2,CTDP1,MKI67,BTG3,SLBP,SKA3,CDCA5,PMF1,TOP2A,WDR62,KIF2C
	GO:0000819	sister chromatid segregation	2.00E-06	13	225	5.78	NUSAP1,KIF18A,PTTG1,CENPA,UBE2C,SMC4,SKA2,PRC1,CEP55,CDCA5,PMF1,TOP2A,KIF2C
	GO:0007059	chromosome segregation	2.90E-06	15	333	4.50	NUSAP1,KIF18A,PTTG1,CENPA,UBE2C,HJURP,SMC4,SKA2,PRC1,CEP55,SKA3,CDCA5,PMF1,TOP2A,KIF2C
	GO:0000280	nuclear division	5.88E-06	19	599	3.17	RAD51,NUSAP1,KIF18A,PTTG1,UBE2C,SMC4,SKA2,PRC1,CCNA2,CEP55,AURKA,TPX2,CTDP1,MKI67,SKA3,CDCA5,PMF1,TOP2A,KIF2C
	GO:0051301	cell division	5.88E-06	20	668	2.99	NUSAP1,CDKN2A,PTTG1,UBE2C,SMC4,SKA2,STMN1,PRC1,E2F8,CCNA2,CEP55,AURKA,TUBA1B,TPX2,CTDP1,SKA3,CDCA5,PMF1,TOP2A,KIF2C
	GO:0048285	organelle fission	1.29E-05	19	636	2.99	RAD51,NUSAP1,KIF18A,PTTG1,UBE2C,SMC4,SKA2,PRC1,CCNA2,CEP55,AURKA,TPX2,CTDP1,MKI67,SKA3,CDCA5,PMF1,TOP2A,KIF2C
	GO:0098813	nuclear chromosome segregation	1.54E-05	13	283	4.59	NUSAP1,KIF18A,PTTG1,CENPA,UBE2C,SMC4,SKA2,PRC1,CEP55,CDCA5,PMF1,TOP2A,KIF2C
	GO:0000070	mitotic sister chromatid segregation	8.51E-05	9	136	6.62	NUSAP1,KIF18A,PTTG1,UBE2C,SMC4,PRC1,CEP55,CDCA5,KIF2C
	GO:0044772	mitotic cell cycle phase transition	1.09E-04	16	529	3.02	CDKN2A,CDKN2B,CDKN3,UBE2C,MCM2,MCM10,RRM2,MELK,GFI1,CCNA2,TYMS,AURKA,LMNB1,TPX2,CTDP1,CDCA5
	GO:0044770	cell cycle phase transition	1.90E-04	16	555	2.88	CDKN2A,CDKN2B,CDKN3,UBE2C,MCM2,MCM10,RRM2,MELK,GFI1,CCNA2,TYMS,AURKA,LMNB1,TPX2,CTDP1,CDCA5
	GO:0000226	microtubule cytoskeleton organization	5.37E-04	14	464	3.02	NUSAP1,KIF18A,CENPA,GAPDH,SKA2,STMN1,PRC1,AURKA,TUBA1B,TPX2,PRKCZ,SKA3,WDR62,KIF2C
	GO:0002520	immune system development	8.30E-04	20	949	2.11	IL10,CDKN2A,CDKN2B,PDCD1,ACP5,HMGB2,MELK,GFI1,TFRC,IFNG,HAVCR2,CSF1,SNX10,CCL3,CCR1,PRKCZ,SEMA4A,ITM2A,TOP2A,VCAM1
Gene Family	471	CD molecules|Tumor necrosis factor superfamily	1.68E-05	12	394	3.05	PDCD1,TNFRSF9,LAG3,ITGAE,TFRC,SIRPG,BMPR1A,CXCR6,CCR1,CCR5,ALCAM,VCAM1
	735	Secretory carrier membrane proteins	7.85E-03	2	5	40.00	SCAMP2,SCAMP5
	1087	LEM domain containing	1.09E-02	2	7	28.57	LEMD1,TMPO
	1091	C-C motif chemokine	1.74E-02	2	10	20.00	CCR1,CCR5

Genes with downregulated expresison (PD1+ve compared to PD1-ve)

Category	ID	Name	q-value FDR	No. of hits	Total GO size	Percentage	Hit in Query List
Gene Family	471	CD molecules|Tumor necrosis factor superfamily	1.34E-03	10	394	2.54	SELL,NT5E,ITGA6,TNFRSF10D,CD320,S1PR1,MST1R,TLR1,TSPAN7,CCR7
	1179	Apolipoproteins|Sushi domain containing	5.45E-03	4	57	7.02	SELL,SUSD4,SNED1,GABBR1
	1217	Brain expressed X-linked family	5.45E-03	2	5	40.00	BEX5,BEX2
	782	Tumor necrosis factor receptor superfamily	6.44E-03	3	29	10.34	TNFRSF25,TNFRSF10D,LTBR
	738	Proteases, serine	4.98E-02	3	63	4.76	PRSS35,TMPRSS4,ST14

Table 24. Gene ontology and gene family analysis of Differentially expressed genes in CD8+PD1+ and CD8+PD1- cells.

Biological processes (Top 15 gene ontologies shown) and gene families that are significantly (FDR <0.05) associated with the differentially expressed genes. The Genes expressed to a greater/ lesser extent in CD8+PD1+ cells compared to CD8+PD1- cells (only the top 20 significant biological processes are shown).

4.2.3.4 Stratification of tumours by Immune density and evaluation of T_{RM} targets using IHC

The differences between the tumour and background lung CD8+ T cells highlighted an emerging T_{RM} signature within the tumour. However, assessment of the distribution of the CD8+CD103+ T_{RM} from the flow cytometry identified a range of samples where lower CD8+CD103+ T_{RM} numbers occur in some cases. This led to the analysis of the tumour CD8 cells, taking in to account immune density; it highlighted that features of Immune cell exhaustion were heterogeneous across the cohort. They were broadly linked to T cell density with $CD8^{Hi}$ tumours showing increased levels of the negative regulators of the immune response in both NSCLC and HNSCC (Figure 49 A and B). The $CD8^{Hi}$ and ^{Lo} groups were further characterised by comparative DESeq2 analysis, this was extended to included tumours with high T_{RM} cell numbers ($CD103^{Hi}$ and $CD103^{Lo}$). The transcriptomic analysis revealed that differences are present in CD8 cells that reside in $CD8^{Hi}$ and ^{Lo} tumours, with a differential expression of activation, functional and negative regulation markers between CD8+ T cells in the $CD8^{Hi}$ and ^{Lo} tumours (Ganesan et al., 2017).

4.2.3.4.1 Tissue distribution of T_{RM} targets in $CD8^{Hi}$ and ^{Lo} tumours

The Immune density which is linked to a favourable outcome has previously focused on CD8 as one of the most effective prognostic markers of an improved survival. Elevated frequencies of CD8+ T cell CD103+ (T_{RM}), PD1+ and 41BB+ cells were observed in the tumour relative to the non-involved lung. Representative images for CD8, CD103, PD1 and Ki67 confirm the presence of the markers identified by transcriptomics but also highlight the differences that exist between $CD8^{Hi}$ and ^{Lo} tumours. $CD8^{Lo}$ tumours exhibit much lower CD8 infiltration but also lower CD103 levels. Ki67+ cells were diffuse in both the $CD8^{Hi}$ and ^{Lo} with the staining most likely attributed to tumour cells (Figure 50).

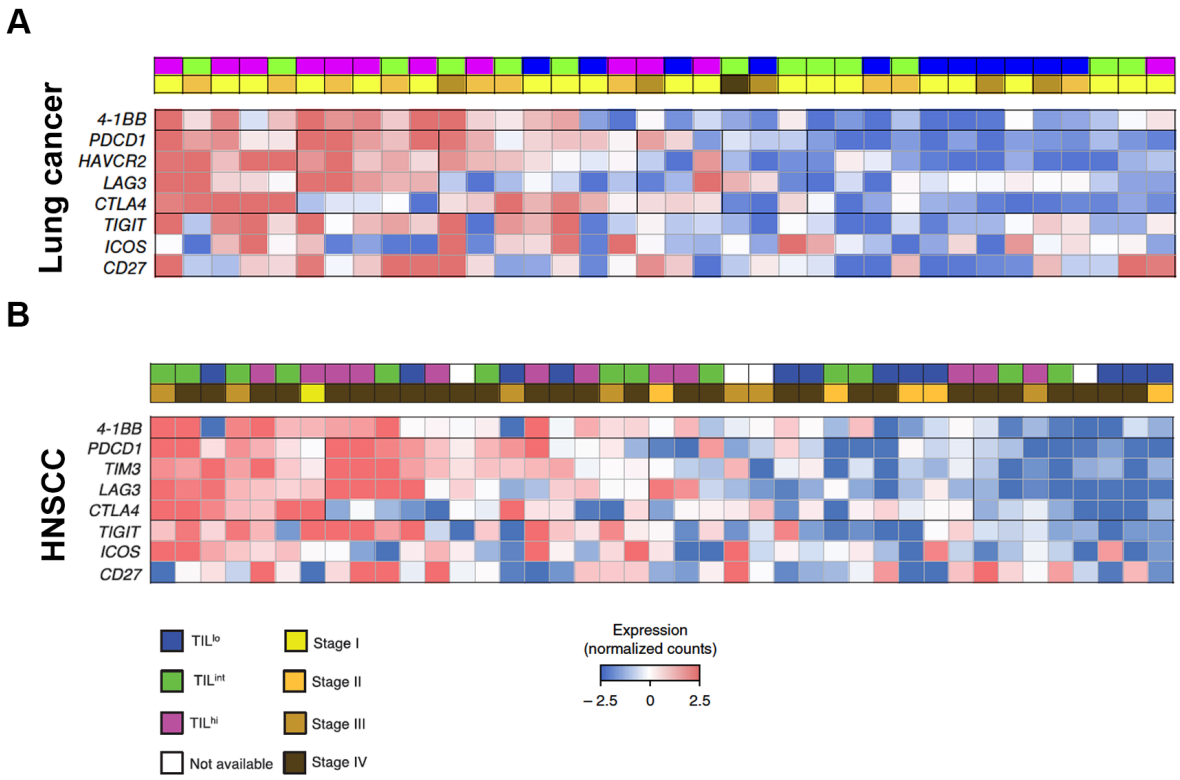


Figure 49. Heterogeneous expression of immunotherapy targets in CD8+ T cells from HNSCC and NCSLSC.

RNA-Seq analysis (normalised counts) of specific immune targets in CD8+ TILs from patients with (A) NSCLC and (B) HNSCC. CD8+ TIL density and tumour stage are annotated. Expression of the immunotherapy targets in baseline samples is heterogenous, however it is broadly elevated in immune dense tumours (CD8) and could be used to guide therapy. Adapted from (Ganesan et al., 2017).

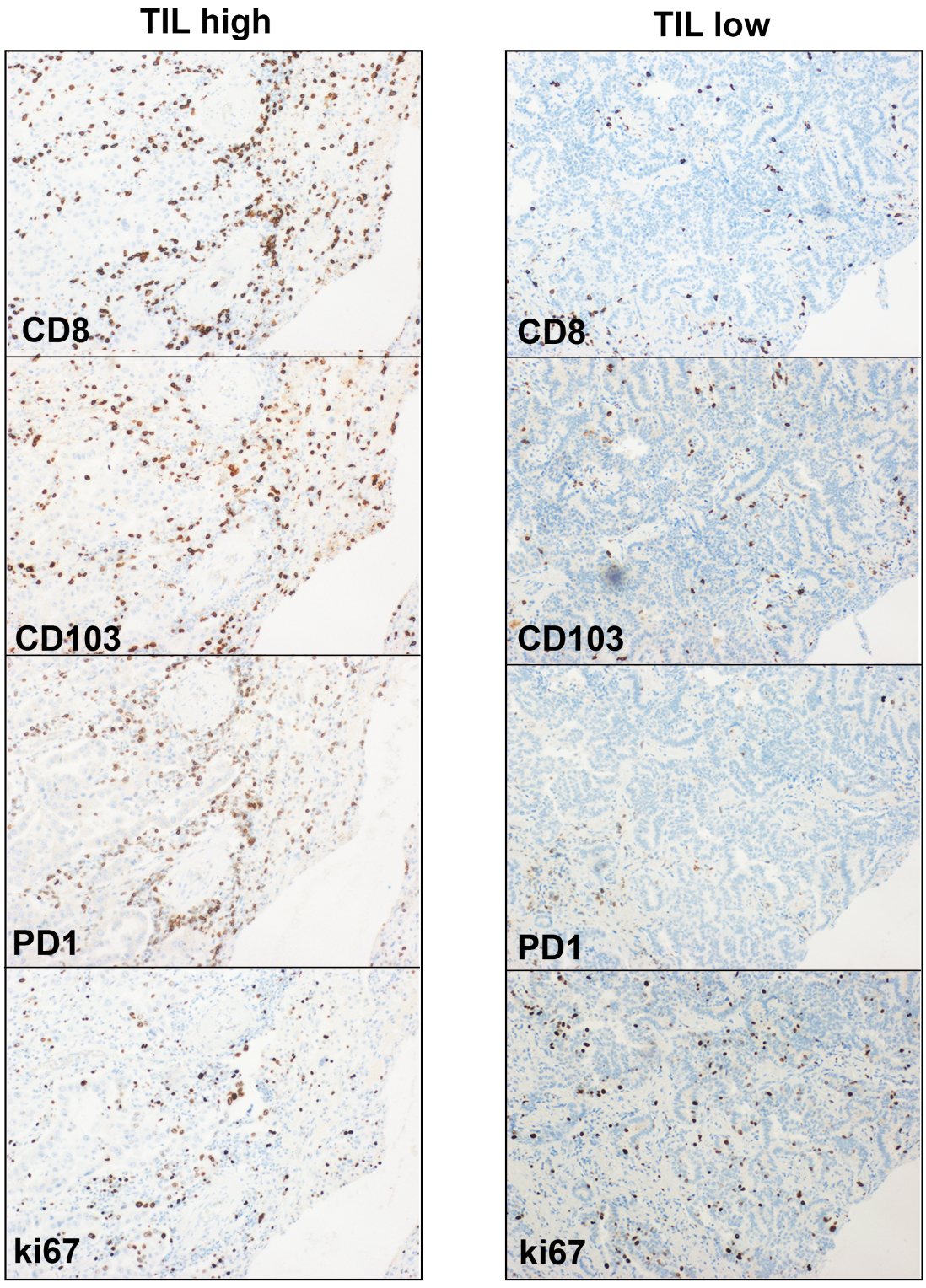


Figure 50. Immunohistochemistry for the surface expression of CD8, CD103, PD1 and Ki67 in TIL Hi and Lo NSCLC.

Representative images of CD8, CD103, PD1 and Ki67 immunohistochemistry in $CD8^{Hi}$ and $CD8^{Lo}$ tumours, Images were captured at 100x magnification. Clear differences in the CD103 and PD1 staining were visible between $CD8^{Hi}$ and $CD8^{Lo}$ tumours. (IHC slides prepared by research histology – Monette Lopez and Maria Machado).

4.2.3.4.2 CD103+ T cell density is predictive of survival

Tumour tissue from a retrospective cohort of n=689 NSCLC patients that had previously not been reported and a cohort of n=124 previously published HNSCC patients (Ward et al., 2014b), were available as a TMA (tissue microarray). They were stained for CD8 and CD103 and the immune density evaluated using the TIL Hi, Mod and Lo classification as used in (Ward et al., 2014b).

4.2.3.4.2.1 NSCLC survival data

In order to assess the NSCLC survival data and ensure the data fit with what has previously been reported, the tumour stage and gender were assessed (Figure 51 A and B respectively). The patient survival is tightly linked to the staging of the tumours, with stage I conferring the best prognosis (~80% at 5 years) and stage IV displaying the worst survival (~15% at 5 years). Another internal check of the survival data is gender, where males have a poorer outcome relative to females (p=0.031). The identification of a unique transcriptomic profile between CD8+ T cells from $CD8^{Hi}$ and ^{Lo} tumours, where CD103 was tightly linked to the cytolytic activity of the CD8+ T Cells meant it may also confer a survival benefit superior to CD8 alone. Analysis of the CD8 and CD103 density in the TMA was carried out using the TIL Hi, Mod and Lo categorisation. A higher density of CD8+ T cells was associated with a trend towards better survival (~55% at 5 years) but did not reach significance (p=0.086) (Figure 52 A). Assessment of $CD103^{Hi}$ tumours revealed a significant survival benefit (~62% at 5 years) over $CD103^{Lo}$ tumours (p=0.043) (Figure 52 B). The CD8 and CD103 survival data were combined in Figure 52 C, this highlighted clearly that $CD103^{Hi}$ tumours have an improved survival relative to the $CD8^{Hi}$ tumours. In order to assess the possibility of sub categories within the $CD8^{Hi}$ group, only the $CD8^{Hi}$ tumours were analysed for CD103 density. The sub-group of $CD8^{Hi}$ /$CD103^{Lo}$ tumours displayed a significantly (p=0.036) poorer survival (~20% at 5 years) relative to the $CD8^{Hi}$ /$CD103^{Hi}$ patients (~62% at 5 years) (Figure 52 D). This suggests that patients with elevated T_{RM} ($CD103^{Hi}$) numbers in their tumour have a better long-term survival even in $CD8^{Hi}$ tumours. Conversely, it shows that in $CD8^{Hi}$ /$CD103^{Lo}$ tumours patients have a worse prognosis reiterating the importance of CD103+ T_{RM} cells in the tumour.

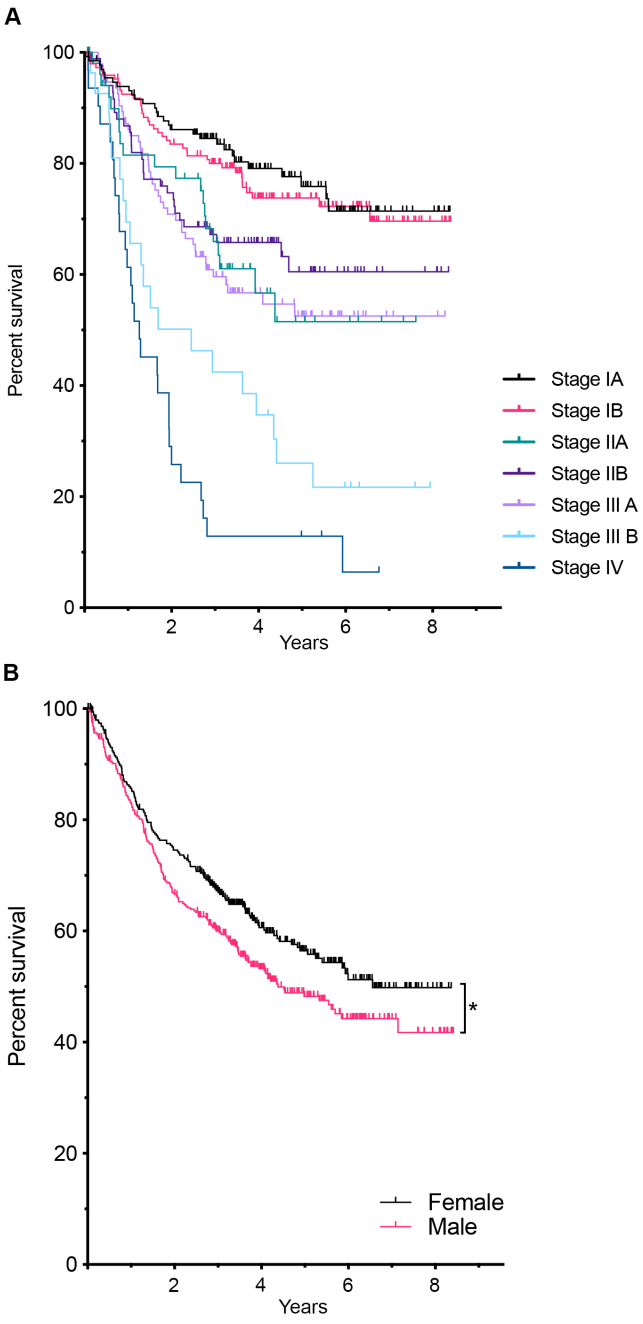


Figure 51. Stage and gender survival differences in lung cancer.

Survival data from n=689 patients with lung cancer. (A) Tumour stage and survival presented as Kaplan-Meier curves, where increasing stage of disease is linked to a worse outcome. (B) Differences between male and female survival (p=0.031; log-rank test presented as a Kaplan-Meier curve). Survival data collected by Dr S. Chee University of Southampton, published in (Ganesan et al., 2017).

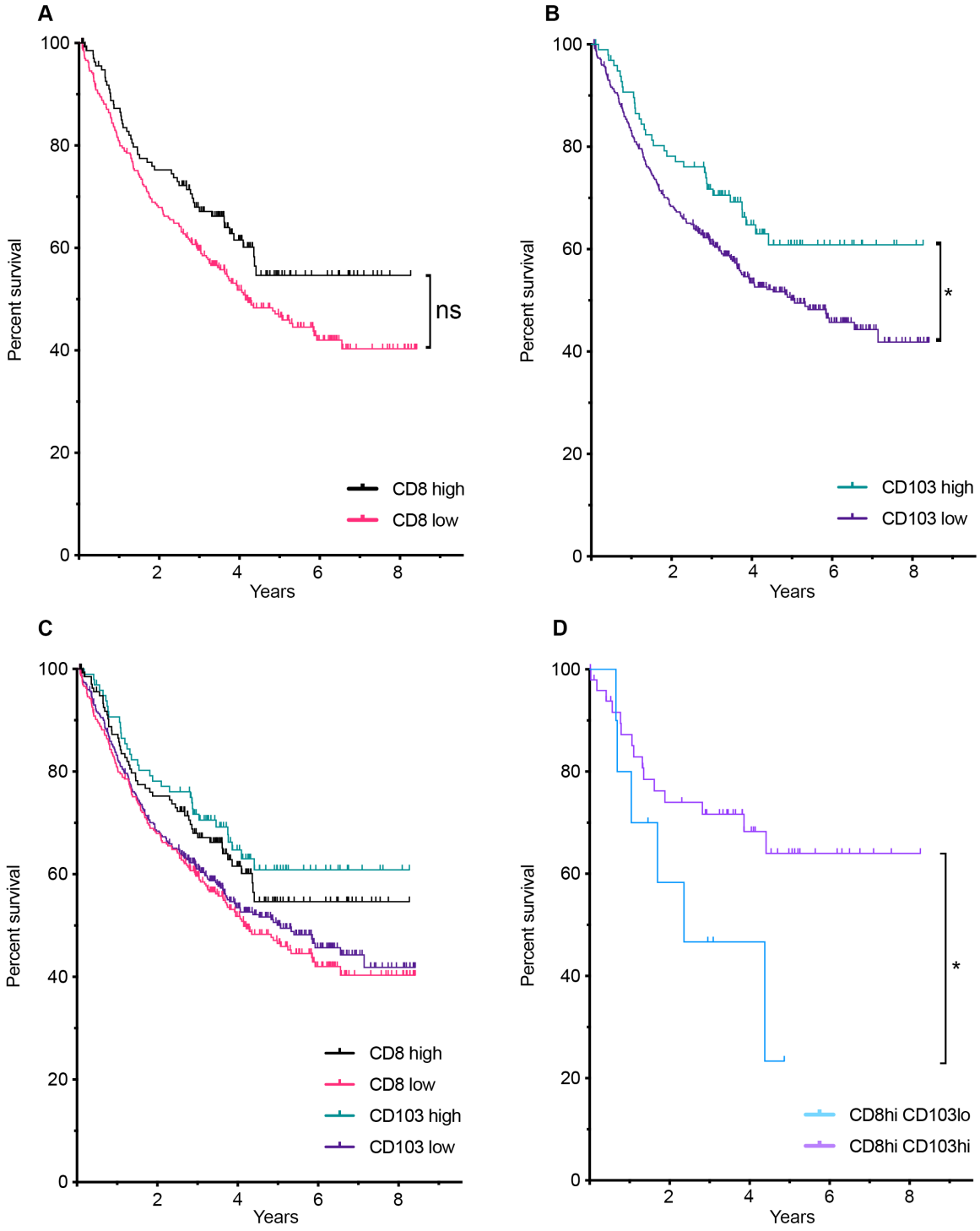


Figure 52. CD103 density predicts for survival in lung cancer.

Survival data from patients with lung cancer (n = 689; presented as Kaplan Meier-curves with log-rank test significance) with a low density ($CD8^{Lo}$) or high density ($CD8^{Hi}$) of CD8 cells in tumours, p=0.086 (A), high or low density of CD103 cells ($CD103^{Lo}$ and $CD103^{Hi}$) p=0.043 (B). Survival data for CD8 and CD103 was also combined in (C). Survival of patients with $CD8^{Hi}$ tumours subcategorised by CD103 density ($CD103^{Lo}$ or $CD103^{Hi}$) in (D) p=0.036. The survival data highlights that $CD103^{Hi}$ tumours confer an improved survival over CD8. Survival data collected by Dr S. Chee University of Southampton, published in (Ganesan et al., 2017).

4.2.3.4.2.2 HNSCC survival data

The same analyses were then repeated in a different epithelial tumour arising in the head and neck (HNSCC) to assess the impact of CD103 positivity on patient survival (Figure 53 A-D). The density of the CD8+ cells were again linked to a favourable outcome (p=0.0182), as has been previously reported. The CD103 density (Hi and Lo) conferred a slight improvement in survival over CD8 alone, although the difference between $CD103^{Hi}$ and ^{Lo} did not reach significance (p=ns). The subcategory of $CD8^{Hi}$ /$CD103^{Hi}$ tumours showed a trend towards a more favourable outcome relative to $CD8^{Hi}$/$CD103^{Lo}$ (p=ns). Numerically this cohort of patients was under powered (n=123) leading to less significance in the result, the same trend was observed in HNSCC that was seen in NSCLC.

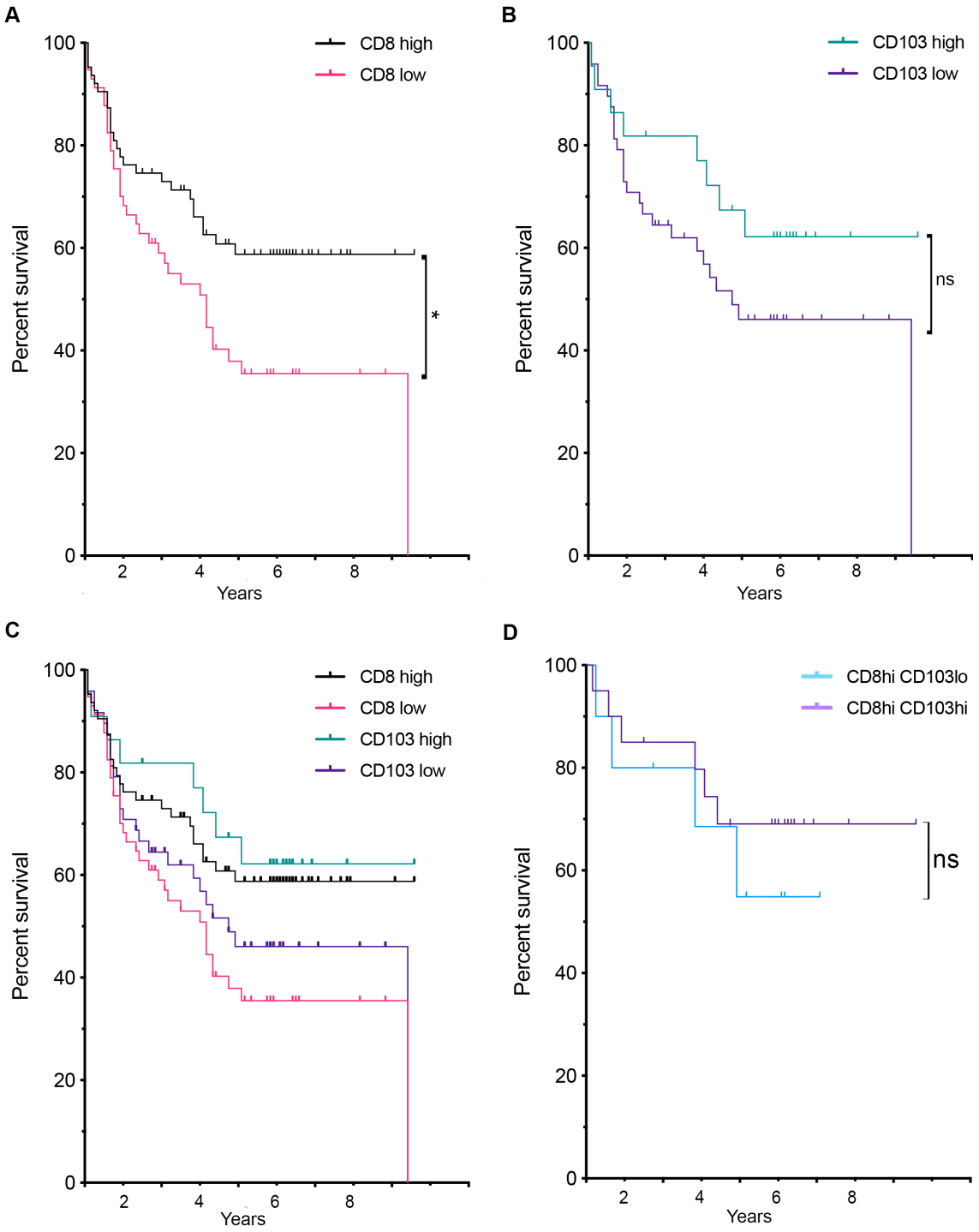


Figure 53. CD103 and patient survival in head and neck cancer.

Patient survival in HNSCC (*n* = 123; presented as Kaplan Meier-curves with log-rank test significance). (A) Patient survival of $CD8^{Hi}$ and $CD8^{Lo}$ tumours from the available cases; * = p=0.0182. (B) CD103 density ($CD103^{Hi}$ and $CD103^{Lo}$) and patient survival, displaying a small improvement over CD8 p=ns (low cohort size impacted significance). (C) Combined Kaplan-Meier curves for CD8 and CD103 and (D) survival of patients with $CD8^{Hi}$ tumours subcategorised into $CD103^{Lo}$ or $CD103^{Hi}$ p=ns. The CD103 shows a trend towards improved Survival, data collected by Dr M. Ward University of Southampton, published in Ward et al., 2014b.

4.2.4 Discussion

Isolation and transcriptomic evaluation of CD8+ T cells from HNSCC and NSCLC has allowed characterization of novel features of tumour specific CD8+ T cells. The initial transcriptomic analysis identified a clear CD8+ T_{RM} T cell signature in the tumours relative to the control tissue (non-involved lung) (Ganesan et al., 2017). Use of flow cytometry for protein validation of CD103 on CD8+ T cells in tumour and non-involved lung, confirmed the findings of the transcriptomic analysis with CD8+CD103+ T_{RM} T cells being present at higher numbers in the tumour compared to the control tissue. T_{RM} T cells play an integral role in the "re-call" response to antigen in peripheral tissues (Schenkel et al., 2014).

T_{RM} cells develop a unique transcriptional profile that is distinct from circulating memory cells, and are retained in the tissue via CD103 (ITGAE), which binds to E-Cadherin and anchors the CD8+ T cells in the epithelial tissues awaiting secondary encounter with antigen (Mackay et al., 2013). The features of T -cell "exhaustion" in the CD8+ T cells were also of interest as they represent effective targets of immunotherapy. The immunotherapy target PD1 was identified at high levels in the CD8+ T cells of the tumour and was present at higher frequencies relative to the NIL; in particular it was elevated in the tumour CD8+CD103+ T_{RM} T cells relative to the non-T_{RM} T cells (CD8+CD103-). PD1 expression can be a marker of clonally expanded antigen-specific T cells in the tumour, indicating that the CD8+ T_{RM} cell population is likely to be a tumour-reactive population (Gros et al., 2014). To test this and the proliferative capacity of the T_{RM} cells and non-T_{RM} cells, 41BB and Ki67 were evaluated at the protein level with flow cytometry. Elevated 41BB+ cell frequencies were identified in the CD8+ T cells between the tumour and NIL, but 41BB+ cells were also increased in the tumour CD8+ T_{RM} cells relative to the non-T_{RM}. Ki67 was also increased in the tumour CD8+ T cells relative to the control tissue and again, was higher in the CD8+ T_{RM} relative to the non-$T_{RM.}$ Ki67 has also been linked with the expression of PD1 in viral infections (Hong et al., 2013). This is indicative of activation through the TCR that leads to increased 41BB expression and cellular proliferation, likely linked to a tumour antigen-driven response and the formation of T_{RM} cells in the tumour (Willoughby et al., 2014) (Nam et al., 2005). The co-expression of PD1, 41BB and Ki67 in the tumour CD8+ T_{RM} cells (FACS, tSNE) indicates an interesting population of recently activated proliferating cells that have increased expression of PD1, classically considered as a marker of dysfunction (Pentcheva-Hoang et al., 2007). These activated tissue resident CD8+ T cells may be crucial in the anti-tumour response, they express high levels of markers that can be targeted by immunotherapy (e.g. PD1). If these cells are present in high enough numbers in the tumour they have the potential to elicit potent anti-tumour immune responses.

Assessment of the cellular density of CD8, CD103, PD1 and Ki67 (IHC) in CD8 TIL^{Hi} and ^{Lo} tumours highlighted differences between the CD8 TIL^{Hi} and ^{Lo} tumours in terms of CD103 density. Immune cell-rich tumours represent an on-going protective immunological reaction to TAAs, conversely TIL^{Lo} tumours likely represent immunologically silent tumours or tumours with stromal barrier features (Djenidi et al., 2015),(Hanley et al., 2018). When taking into account the CD8+ TIL density ($CD8^{Hi}$ and ^{Lo}) and CD103 expression ($ITGAE^{Hi}$ and ^{Lo}) in the CD8 transcriptomic data, further differences were revealed between CD8+ T cells from immune-rich tumours relative to immune-low tumours. Increased levels of molecules indicating cytolytic function and active killing of tumours (GZMB, PRF1 and IFNG) were present, suggesting that CD8+ T_{RM} T cells were functionally superior to the CD8+ non-T_{RMs} in the tumour (Ganesan et al., 2017). Due to CD8+CD103+ cells (T_{RMs}) displaying improved ability to control tumours, the T_{RM} marker CD103 was used in two retrospective survival cohorts (NSCLC and HNSCC), where $CD103^{Hi}$ infiltrates in tumours more clearly appeared to confer an improved patient survival compared to CD8+ T cells. This was significant in the NSCLC cohort, with a trend to significance in HNSCC, the small cohort size in HNSCC likely impacted the significance level. When only the $CD8^{Hi}$ tumours are considered, the $CD103^{Hi}$ and ^{Lo} patient groups separate, with $CD8^{Hi}$ /$CD103^{Lo}$ patients displaying a worse outcome compared to the $CD8^{Hi}$ /$CD103^{Hi}$ group. This is likely due to differences in the subtypes of CD8+ T cells within the tumour, where an enrichment of the CD8+CD103+ cell (T_{RM}) subtype confers a survival benefit. Data from a Breast cancer cohort also showed that CD103 was prognostic when combined with CD8 and that the intraepithelial density was more beneficial than intrastromal infiltrate (Wang et al., 2016). The evaluation of CD103 (ITGAE) expression in the HNSCC and NSCLC TCGA data, also shows the same survival benefit. Where higher expression is associated with prolonged survival when compared to lower expression of ITGAE (Duhen et al., 2018). A variety of novel potential immunotherapy targets were also identified from the transcriptomic analysis in the CD8 TIL^{Hi} T_{RMs} (CD38, CD39, BATF, NAB1, KIR2DL4 and SIPRG). The cell surface ectonucleotidase CD39 (ENTPD1) is an interesting target that converts ATP to AMP, it may protect T_{RMs} (among other immune cells) from ATP-induced cell death (Ganesan et al., 2017). The survival benefit linked to the density/ expression of CD103 indicates that these are important cells that can be targeted by immune check point blockade, or could function as a key marker of patients that may respond to immune check point blockade.

In HNSCC that is divided into HPV(+) and (-) disease, it was unclear if differences in the CD8+ T cells may exist due to a viral tumour driver. Differential gene expression testing between HPV(+) and (-) CD8+ T cells was unable to reveal a substantial change between the CD8+ T cells derived from the two tumour types. It was hypothesised that the HPV(+) tumours may generate a different virus-specific response compared to HPV(-) tumours (Frazer, 2009). No overlap of the DEGs was observed with gene ontologies reflecting virus-specific immune responses as well as no

obvious grouping of cases by HPV status. This suggests very surprisingly that, on the whole, CD8+ T cells are transcriptionally similar between the two tumour subtypes. It argues that a virus-specific response (HPV viral infection i.e. HPV capsid proteins L1 and L2) is not the determining factor in the difference in immune density, and that the key factor is the immunogenic E6 and E7 proteins that are abundant in all HPV(+) tumours. This then leads to the hypothesis that the immune density in HPV(-) tumours is determined in part, by how immunogenic and visible it is to the immune system (less visible than HPV(+) tumours). This fits with the concept of higher immune infiltrates in tumours that display higher mutational loads, with an increase in potential neoantigens (McGranahan et al., 2016, Sahin et al., 2017). This is however a relatively small cohort and doesn't account for the fraction of episomal and integrated viral DNA in the HPV(+) tumours, which could impact the immune response (Olthof et al., 2014). It is possible that the level of episomal and integrated viral DNA leads to the differential immune responses that are observed in HPV(+) tumours coupled together with the variability in HPV gene expression (E1, E2, E6 etc.) within the tumours (Wood et al., 2016b). The transcriptional similarities between the CD8 cells in different tumours types (NSCLC and HNSCC), and also the sub-categories has important implications for immunotherapies. It argues that the cells are functionally similar, meaning therapies that are effective in one tumour are also likely to work in a another. The caveat is that the density of the immune cells, in particular the CD8+ T_{RM} infiltrate needs to be considered.

Examination of the transcriptomic features of CD8+PD1+ T cells also mirrored the findings that were identified in the CD8 bulk transcriptome, where cells expressing high levels of PD1 exhibited the same features as CD8+ T_{RM} Hi tumours (CD8+CD103+), a T cell exhaustion profile (TIM3, LAG3, CTLA4 and PD1) and at the same time a high level of antigen engagement (41BB) and cytolytic activity (GZMA/B). A large number of genes were shared between $CD103^{Hi}$ (Ganesan et al., 2017) and $PD1^{Hi}$ tumour CD8+ T cells. This demonstrated that the cells expressing high levels of PD1 or CD103 displayed similar characteristics of a protective adaptive immune response, and that PD1 and CD103 were co-expressed in tumour CD8+ T cells. Features of tissue residency (T_{RM}) and clonal expansion have also been identified in CD8+PD1+ cells in chronic inflammatory disorders, this includes the paradoxical elevation of cytolytic and antigen experience markers (GZMA, GZMB and 41BB) (Petrelli et al., 2018). This was initially counter-intuitive as the expression of PD1 is classically considered a marker of T cell dysfunction not function (Wherry EJ, 2007). It argues that upon activation, the negative regulators of the immune response are switched on during the active killing phase of the CTL, before they ultimately lead to cellular inactivity. One study highlighted that CD8+PD1+ cells exist in two states, one that is terminally differentiated and incapable of reactivation with the other displaying cellular plasticity that can be reprogrammed to an active state; these features were determined by the epigenetic state of the cells (Philip et al.,

2017). Another key marker that has emerged is TIM3 (HAVCR2) that was highly elevated in both $CD103^{Hi}$ and $PD1^{Hi}$ groups of CD8+ T cells; this has become a marker of hyper-activated CD8+ T cell responses in tumours (Fourcade J, 2010). The $CD8+PD1^{Lo}$ T cells exhibited a central memory/ Naïve phenotype expressing CCR7 and CD62L. Finding high numbers of CD8 T_{RM} features in the PD1+ cells adds further evidence to this group of cells being key targets for immunotherapy.

Data leading on from the interrogation of tumour CD8+ T cells presented here focused on the dissection of CD8+ T_{RM} and non-T_{RMs} at the bulk and single cell levels. This identified three unique Tumour-specific T_{RM} populations that expressed high levels of cytolytic markers that also co-expressed TIM3, PD1 and 41BB. The TIM3+ T_{RM} cells were found to be clonally expanded and also elevated in patients that respond to anti-PD1 therapy (Clarke et al., 2019). These subsets have the potential to be used as immunotherapy response biomarkers for patient stratification, adoptive cell therapies and tumour-specific TCR recovery. Interestingly the transcription factor TCF7 was expressed at higher levels in the $PD1^{Lo}$ group relative to $PD1^{hi}$; TCF7 has been identified as a marker of Immunotherapy responder versus non-responder. In this single cell study it was expressed on early activated cells and effector memory cells but not cells displaying high cytotoxic activity as observed from the previous CD8 T_{RM} PD1+TIM3+ study (Sade-Feldman et al., 2018). The exact cellular mechanisms and cell state transitions that take place during an effective immune attack on the tumour are still being elucidated, whether early TCF7 expression and activation leads to T_{RM} formation in the tumours, expression of immune checkpoints and ultimately a T cell contraction phase are still being elucidated.

4.2.5 Conclusion

Evaluation at the transcriptomic level of purified CD8+ T cells from NSCLC and HNSCC has revealed unique features that were indicative of tissue residency (CD8+CD103+). The CD8+ T_{RMs} (CD103+) expressed markers of T cell exhaustion (PD1) and activation (41BB, Ki67), this was observed transcriptomically and also at the protein level. Features of superior functionality were seen in CD8+ T_{RMs} from $CD8^{Hi}$ tumours. Retrospective analysis of the CD8+ and CD103+ density in NSCLC and HNSCC identified an improved prognosis linked to higher numbers of CD103+ cells when compared to CD8+ T cells alone. This has implications for patients, where higher CD103 levels would predict for a better prognosis and potentially improved immunotherapy responses. The work on CD8+ T_{RMs} and their role in cancer immunity was investigated further in the publication by Clarke et al., 2019.

4.3 Tumour infiltrating B lymphocytes in NSCLC

4.3.1 Introduction

The role that B cells play in the immune response to tumours has been investigated far less than that of their T-Cell counterparts. Tumour-infiltrating B lymphocytes (TIBs) have the potential to regulate immune function through antibody, presentation of antigens, co-stimulation and cytokines (Martin and Chan, 2006). B cells represent a heterogenous population (cell subtypes outlined in section 1.4.2) with distinctive functions that can have both pro-tumour and anti-tumour properties.

Current thinking about the positive impact of B cells in tumours has resulted from observations that higher B cell infiltrates are linked to better patient outcomes. In NSCLC, CD20+ B cells were prognostic in both stromal and epithelial locations (Al-Shibli et al., 2008). Another key element of the B cell response to tumours was the formation of tertiary lymphoid structures (TLS). These arise within the tumour and contain the same cellular components and morphology as germinal centres in secondary lymphoid tissues (dark and light zones), they express the markers BCL6, CD23, Ki67 and AID (Activation-induced cytidine deaminase). Reactivity to tumour antigens by IgA and IgG indicates that isotype switching can occur in the tumour and/or lymph node (Germain et al., 2014). In Breast cancer, the formation of TLS and high numbers of class-switched plasma cells (also co-localised with CD8+, CD4+ and CD20+ cells) were linked to a more favourable prognosis, the antibody isotypes IgG1 and IgG3 were found to dominate responses (Kroeger et al., 2016). In pancreatic ductal adenocarcinoma, B cells were identified as being prognostic but only in TLS-high tumours (not CD20+ B cells alone), which were again, co-localised with CD8+ cells (Castino et al., 2016). The clonal expansion of T cells has also been correlated to the density of TLS and B cells, showing a link between the B cells and the diversity of the adaptive immune response (Zhu et al., 2015).

Mining of the CGA (Cancer genome atlas) data for Breast and ovarian cancer identified a B cell gene expression signature that was linked to improved survival; a lower BCR diversity was also consistent with an antigen driven clonal response (Iglesia et al., 2014). It is not currently known whether the characteristics of B cells in immune dense tumours is different to that of immune sparse tumours. However, the improved survival that is linked with higher numbers of co-localised B cells and CD8+ T cells (Kroeger et al., 2016) indicates that differences may exist.

The other side of the humoral response to tumour involves a negative impact on the protective function of adaptive immune T cells. Immunosuppressive B cells can arise in different stages of B cell development where immature B cells differentiate into regulatory B (Breg) cells. These cells

represent a product of their environmental cues (IL21, IL6, IL35, CD40) rather than lineage specific transcription factors like FOXP3 for Tregs (Matsumoto et al., 2014, Rosser and Mauri, 2015). Breg cells function to impede the immune response and have a negative effect on a wide range of immune cells (CD4+ TH1, TH17 and Treg, NKT, CD8+ CTLs, DCs and monocytes) via the secretion of IL10, TGFβ and IL35 (Matsumoto et al., 2014, Rosser and Mauri, 2015). Use of anti-CD20 to deplete B cells in squamous cell carcinoma (mouse model) led to increase in the CD8+ T cell infiltrates by modulating the tumour macrophages towards a protective M1 phenotype (Affara et al., 2014). The elimination of Immunosuppressive plasma cells expressing IgA, IL10 and PDL1, also enabled the eradication of tumours in a mouse model of prostate cancer (Shalapour et al., 2015). Likewise CD8+ CTL function was restored by depletion or blockade of IgA+IL10+PDL1+ Breg cells that accumulate at the site of hepatocellular carcinoma (Shalapour et al., 2017). The phenotype of CD4+ T cells is also modulated by B cells in NSCLC, where activated TIBs (CD19+CD20+CD69+CD27+CD21+) were associated with an effector T cell response (IFNγ+ CD4+ TILs) and the presence of TLS. Exhausted TIBs (CD19+CD20+CD69+CD27–CD21–) were associated with a regulatory T-cell phenotype (FoxP3+ CD4+ TILs) and absence of TLS (Bruno et al., 2017). In Breast cancer, conversion of CD4+ T cells to CD4+ Tregs can also be induced by TGFβ released from Bregs, further hampering the immune response. This study identified the TIBs as activated but non-proliferative mature B cells expressing STAT3, CD19, $CD25^{Hi}$, $CD69^{Hi}$, $B7\text{-}H1^{Hi}$, $CD81^{Hi}$, $CD86^{Hi}$, $CD62L^{Lo}$ and IgM^{Hi} (Olkhanud et al., 2011). Tregs that secrete IL21 (also secreted by Tfh cells) can induce GZMB in B cells, GZMB+ Bregs express the regulatory markers IL10, IDO and CD25 and are considered immunosuppressive (Lindner et al., 2013). Another source of IL21 are the Tfh cells in the germinal centre, where IL21 is secreted upon interaction with B cells (Belanger and Crotty, 2016, Crotty, 2011). The formation of GZMB+ Bregs can occur in the absence of CD40L co-stimulation and in the presence of IL21, the resulting GZMB+ Bregs suppress T cell functionality, although GZMB+ Bregs have also been shown to display cytotoxic effects on tumour cells (Hagn et al., 2012). These regulatory B cell features again may vary in different tumours settings, in particular tumours with different immune cell densities (e.g. $CD8^{Hi}$ and $CD8^{Lo}$ tumours), where more Bregs could lead to less immune cell infiltration or impeded functionality.

B resident memory cells (B_{RMs}) are a subset of B cells in peripheral tissues, they are not currently well-defined. Much like T_{RM} cells (CD8+CD103+), they reside in the tissue and are established in response to local antigen encounter. In the lung, B_{RMs} have been identified as CXCR3+, CD62L- and are evenly divided between CD73+/ - (NT5E gene) expression. They play a role in accelerated antibody responses following antigenic challenge (Allie et al., 2019), although the role/ presence of B_{RMs} in tumours has yet to be established.

B cells are abundant in NSCLC and represent 15.9% of the CD45 cells (our FACS data shows B cells as 12.9% of CD45+ cells). Naïve B cells make up a small proportion (1.6%), plasma cells were the

least abundant (0.8%), memory cells (4.2%) and GC B cells (2.2%) were the most abundant (memory B cells = CD27+CD38+/-; GC B cells = IgM+IgD- B cells) (Stankovic et al., 2018, Kaminski et al., 2012).

The application of single cell (SC) transcriptomics has changed the shape of cell biology and broadened our understanding of it. The field of cancer immunology has embraced this methodology which enhances insight into both the tumour and immune cells. Publications reporting transcriptomic analysis of single cells (SC) that contain B cells, are summarised in Table 25 (Chung et al., 2017, Peng et al., 2019, Sade-Feldman et al., 2018, Lambrechts et al., 2018, Tirosh et al., 2016, Puram et al., 2017). SC analysis of B cells in breast cancer identified germinal centre B cells (centroblasts/ centrocytes) and cells with memory/ naïve profiles (Chung et al., 2017).

Title	Tumour type	No. B cells	No. of patients	RNA-Seq method	Summary of B cell data in the paper	Reference
Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq.	Melanoma	515	10	Smart_Seq2	B cell clusters defined in tSNE, cluster specific genes listed in supplemental data, B cell data not used in discussion.	Tirosh et al., 2016
Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer.	Breast cancer	175	11	Smart_Seq	B cell cluster identified and cluster specific genes subjected to GO analysis (B cell activation, Ig like, BCR signalling). Gene set variation analysis (GSVA) using B cell gene signatures. Two sub classes identified = centroblast/ cenrocytes (GC B cells) and memory /naïve B cells. B cell data not used in discussion.	Chung et al., 2017
Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer.	HPV(-) HNSCC	138	18	Smart_Seq2	B cells identified as a single cluster (B/ Plasma cells), B cells from HPV(-) tumours in this paper were only found in lymph node samples. No B cell data used in discussion.	Puram et al., 2017
Phenotype molding of stromal cells in the lung tumor microenvironment.	NSCLC	4806	5	10x genomics	B cells identified as a cluster from tumour isolate, also stated that B cells were elevated in tumour compared to background lung. B cells (CD79A/B) re-clustered separately into 9 clusters, 6 clusters were enriched in tumours.Follicular B cells expressing high levels of CD20 (MS4A1), CXCR4 and HLA-DRs. Plasma cells (IGHG1) and. MALT B cells (IGHA1, JCHAIN). No B cell data used in discussion.	Lambrechts et al., 2018
Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma.	Melanoma	1428	48	Smart_Seq2	Two B cell clusters identified. B cells (CD79A, MS4A1) and Plasma cells (IGH genes). State that responders to ICB have higher B cell numbers (not Plasma cells). No B cell data used in discussion.	Sade-Feldman et al., 2018
Single-cell RNA-seq highlights intra-tumoural heterogeneity and malignant progression in pancreatic ductal adenocarcinoma.	Pancreatic	2416	24	10x genomics	Global tSNE identifies B cells, these are then sub divided into 6 clusters. Key B cell markers highlighted. State that using Cell state trajectory, B cells transition from Naïve to Plasma blast. No B cell data used in discussion.	Peng et al., 2019

Table 25. Overview of current B lymphocyte single cell data.
Multiple single cell data sets are summarised including the overall findings relating to B cells in each publication.

Analysis of patients either responding or non-responding to immune check point blockade, identified two B cell clusters, CD20+ B cells were enriched in responders compared to non-responders while plasma cells were equivalent between the two groups (Sade-Feldman et al., 2018). B cells isolated from pancreatic ductal carcinoma formed 6 distinct clusters that displayed a transition from Naïve to activated plasmablasts (Peng et al., 2019). Single cell analysis of B cells from NSCLC defined them as Follicular (CD20+, CXCR4+ and HLA-DR), mucosal associated lymphoid (MALT) tissue-derived (higher relative IGHM, IGHA1 and IGJ expression) and plasma cells (higher relative IGHG1-4 expression). This type of classification is not entirely appropriate as IGHM, IGHA1-2 and IGHG1-4 are also expressed by other B cell subsets, in this case it refers to higher relative expression (not unique expression) in those B cell clusters. The B cells formed 6 clusters and were enriched in the tumour relative to the non-involved lung tissue (Lambrechts et al., 2018). These studies so far have not assessed B cells in patients with different levels of CD8+ T cell infiltrates ($CD8^{Hi}$ and $CD8^{Lo}$) or elaborated beyond the core B cell markers what the transcriptional profile of the B cells in each cluster represent.

4.3.2 Objectives

Evaluation of transcriptomic features of B lymphocytes from $CD8^{Hi}$ and $CD8^{Lo}$ tumours in NSCLC. B cells from adjacent non-involved lung tissue and lymph nodes will also be evaluated. Guided analysis will use markers known to differentiate B cells, following this, cluster-specific differential gene analysis will be used to reveal B cell features that are specific to $CD8^{Hi}$ and $CD8^{Lo}$ tumours.

4.3.3 Results

4.3.3.1 Evaluation of B cell density and formation of aggregates in the cohort

The cohort of $CD8^{Hi}$ (n=6) and $CD8^{Lo}$ (n=6) NSCLC tumours collected for single cell transcriptomic analysis were evaluated using IHC for CD20 (MS4A1) (Figure 54 A and B). The broad B cell marker CD20 was able to capture the overall immune density of B cells, as well as the number of B cell aggregates that form tertiary lymphoid structures (TLS). Analysis of the density of TIBs revealed that the $CD8^{Hi}$ cases were heavily enriched for B cells and showed a significant difference for those that resided within dense aggregates and TLS (p=0.0024) (Figure 54 A and B). However, although $CD8^{Lo}$ tumours had substantially fewer B cells and fewer B cell aggregates, they were not absent and resided predominantly at the tumour margin.

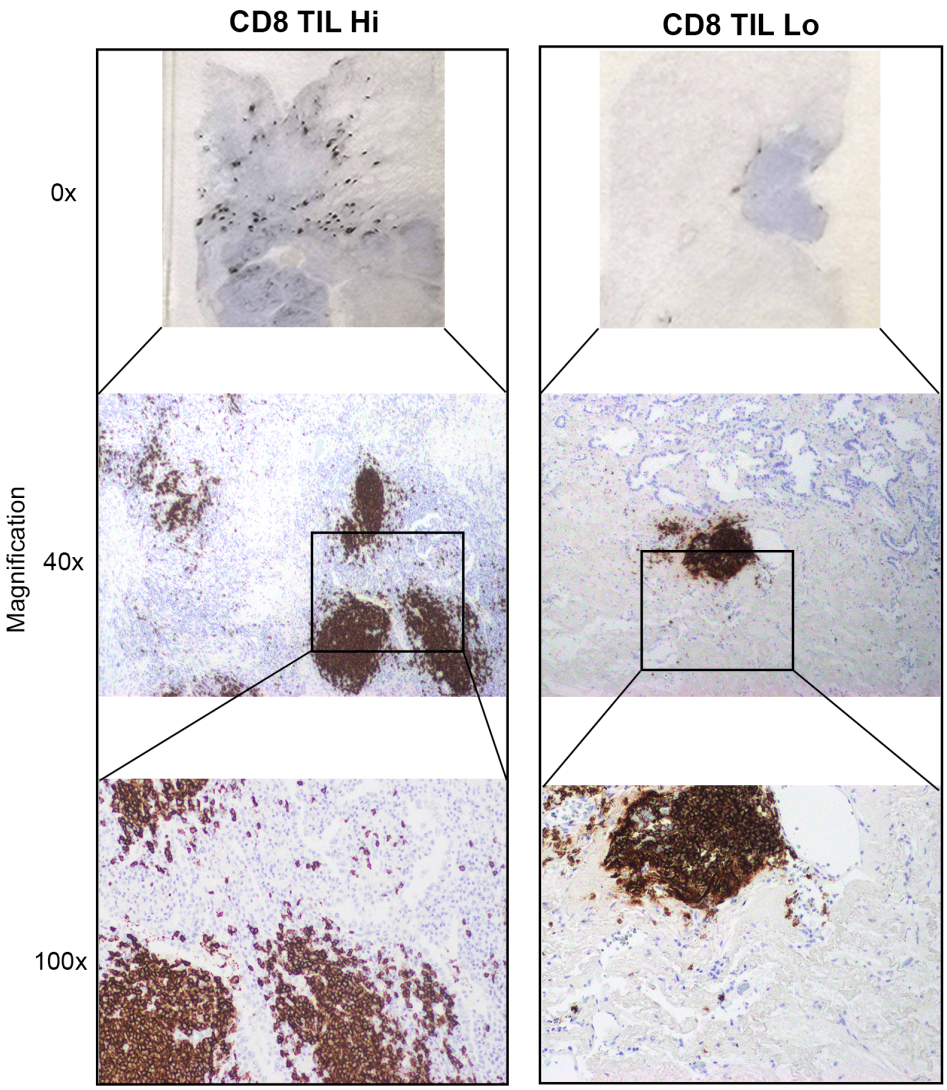
4.3.3.2 Single cell data analysis framework

After filtering and normalisation of the single cell data (Seurat v3.0), 4989 cells were taken forward for analysis of cell lineage and cluster-specific genes - see Table 26 for the number of B cells derived from each tissue origin. Linear dimensionality reduction was performed using PCA, followed by assessment of the data across the PCAs using the JackStraw plot to identity significant principle components (PCs) (Macosko et al., 2015). Additionally, the 'Elbow plot' displaying the standard deviation of the PC's was used to determine the number of PCs to take into the cell clustering (Appendix D1.1 and D1.2 respectively). Significant PCs were identified as having a p-value <0.05, the significance of the PCs drops after PC12 in the Jackstraw plot. This is also evident in the elbow plot where the standard deviation of the data reduces forming an 'elbow' at PC12. Principle components 1 to 12 were taken into the cell clustering with a resolution of 0.6.

Tissue origin	**Number of B cells**
$CD8^{Hi}$ tumours	1547
$CD8^{Lo}$ tumours	1165
Non-involved lung	756
Lymph node	1521
Total	**4989**

Table 26. Number of B cells from each tissue compartment.
The number of B cells recovered and taken forward for cluster evaluation for each tissue compartment ($CD8^{Hi}$and $CD8^{Lo}$ tumours, Non-involved lung and Lymph node).

A



B

Tumour Id	CD8 density	CD20 density	No. of aggregate	Location of TLS
TL853	High	Moderate	23	Intratumoural + Tumour edge
TL846	High	High	64	Tumour edge
TL873	High	High	59	Intratumoural + Tumour edge
TL925	High	High	34	Intratumoural + Tumour edge
TL938	High	High	61	Intratumoural + Tumour edge
TL920	High	High	46	Intratumoural + Tumour edge
TL29111	Low	Low	4	Tumour edge
TL939	Low	Low	15	Tumour edge
TL956	Low	Low	17	Tumour edge
TL959	Low	Moderate	28	Tumour edge
TL933	Low	Low	3	Tumour edge
TL924	Low	Moderate	26	Tumour edge

Figure 54. Evaluation of B cell aggerates in $CD8^{Hi}$ and $CD8^{Lo}$ tumours.

(A) Representative CD20 (MS4A1) IHC images from $CD8^{Hi}$ and $CD8^{Lo}$ tumours at 0x, 40x and 100x magnification. (B) Qualitative assessment of B cell density and number of B cell aggregates from the single cell data cohort. $CD8^{Hi}$ tumours have significantly higher numbers of B cell aggregates compared to $CD8^{Lo}$ tumours (unpaired t-test p=0.0024) (IHC slides prepared by research histology – Monette Lopez and Maria Machado).

4.3.3.3 Multiple B cell clusters emerge in NSCLC and background lung

Following dimensionality reduction and tSNE (t-distributed stochastic neighbour embedding) the B cells formed 9 transcriptionally distinct clusters (C0-8): these are colour coded by cluster and visualised in Figure 55.

4.3.3.1 Identification of B cell subtypes across the clusters

Following on from the identification of different B cell clusters across the different tissue compartments, B cell subtype allocation was undertaken using lineage specific markers (Table 27) (Stankovic et al., 2018, Kaminski et al., 2012). The key lineage markers covering the main subtypes of B cells were visualised by tSNE (Figure 56), where naïve B cells (CD27-IGHD+), activated/ memory cells (CD27+IGHD-Isotype switched+/-), Plasmablasts ($CD20^{lo}$CD27+CD38+Isotype switched+/-) and Plasma cells ($CD20^{Lo}$SDC1+Isotype switched+/-) can be distinguished. However, the expression profiles for the designated lineage markers were not distinct and formed a continuum of expression across several of the different clusters. In addition, not all surface markers were successful at differentiating B cells at the transcriptional level; for example CD24 and CD21 display absent/ very low expression and cannot be used for cell type allocation.
The naïve B cells (IGHD+CD27-) were elevated in C0 and C2 and a small fraction of C1. Activated/ memory cells expressing CD27, class switched Ig isotypes (IGHA1, IGHG1) and CD20 (MS4A1) were distinguishable from the naïve cells (CD27-IGHD+), and were identified in clusters C1, C2, C4, C5, C6. Plasmablasts are evident in C7 where low CD20 and abundant CD27, CD38, SDC1 and high expression of isotype switched Ig was observed (IGHA1 and IGHG1). Cluster C8 represented Plasma cells with high levels of SDC1 and a lack of CD20 and CD27 expression. The mean normalised Seurat expression and mean expressing fraction for the lineage markers for each cluster are shown as bar charts in Figure 57. In each cluster the genes display the same overall expression pattern and cell allocation complementing the tSNE. Very high normalised and percentage expression of Ig genes (IGHA1, IGHG1 and IGHM) was observed in the plasmablast cluster C7.

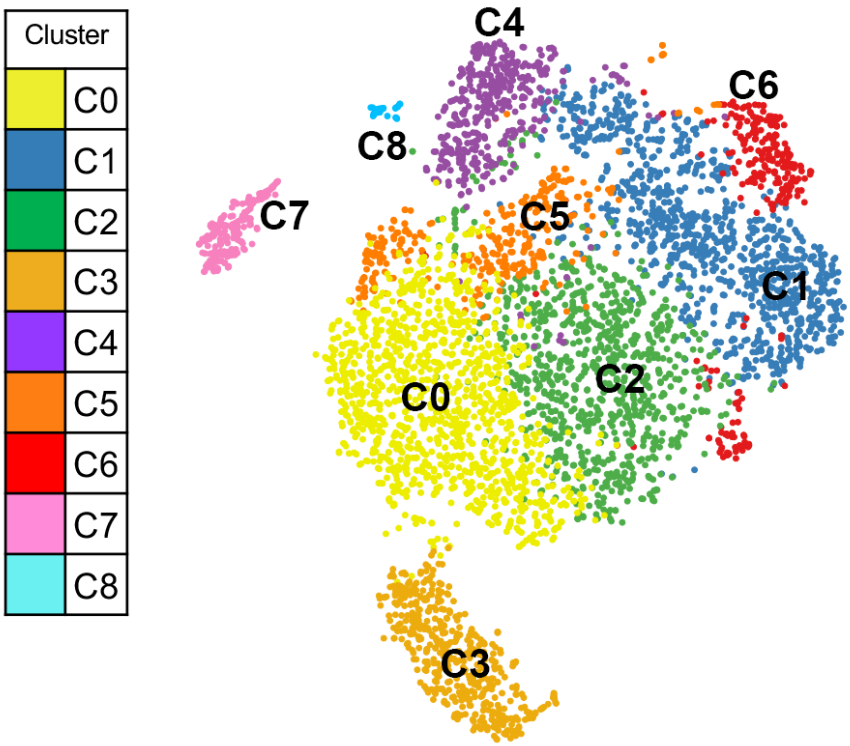


Figure 55. B cell clustering and visualisation of Single cell data by tSNE

Visualisation by tSNE (t-Distributed Stochastic Neighbour Embedding) of the 4989 B cells from tumour, non-involved lung and lymph node, each dot represents a cell, the colours indicate the 9 (C0 to C8) B cell clusters (Seurat v3.0 clustering).

Name	Type	Phenotype	Markers to further subset	Ascribed functions
Transitional	T1	**$IgD^{+}CD27^{neg}$** $CD10^{+}CD24^{high}$ $CD38^{high}$		Precursor to T2; IL10 production
	T2	**$IgD^{+}CD27^{neg}$** $CD10^{+}CD24^{high/+}$ $CD38^{high/+}$		Precursor to T3; IL10 production
	T3	**$IgD^{+}CD27^{neg}$** $CD10^{neg}CD24^{+/low}$ $CD38^{+/low}$		Precursor to mature-naïve; IL10 production
Mature-naïve		**$IgD^{+}CD27^{neg}$** $CD10^{neg}CD24^{+/low}$ $CD38^{+/low}$	CD23, CD69, CD80, CD86	Precursor to GC, memory, and antibody-secreting cells
Memory	Double-negative	**$IgD^{neg}CD27^{neg}$**	CD21, CD24, CD95, CXCR3	Recall responses (and effector functions)
	Non-switched	**$IgD^{+}CD27^{+}$**	CD1c, CD21, CD24	Immunoprotective self antibody ; circulating MZ-like ; regulatory
	IgM-only	IgM^{+}**$IgD^{neg}CD27^{+}$**	CD1c, CD21, CD24	Immunoprotective self antibody ; circulating MZ-like ; regulatory
	Switched	IgM^{neg} **$IgD^{neg}CD27^{+}$**	CD21, CD24, CD95, CXCR3	Pathogen protection; autoimmune pathology
Antibody-secreting cell	Plasmablast	**$IgD^{neg}CD27^{high}$** $CD38^{high}CD138^{neg}$	CD20, HLA-DR	Antibody secretion
	Plasma cell	**$IgD^{neg}CD27^{high}$** $CD38^{high}CD138^{+}$	CD20, HLA-DR	Antibody secretion

*Bold indicates core markers

Table 27. B cell lineage markers for subtype allocation

Table of phenotypic surface markers successfully used in flow cytometry to determine B cell subtypes. Adapted from (Kaminski et al., 2012).

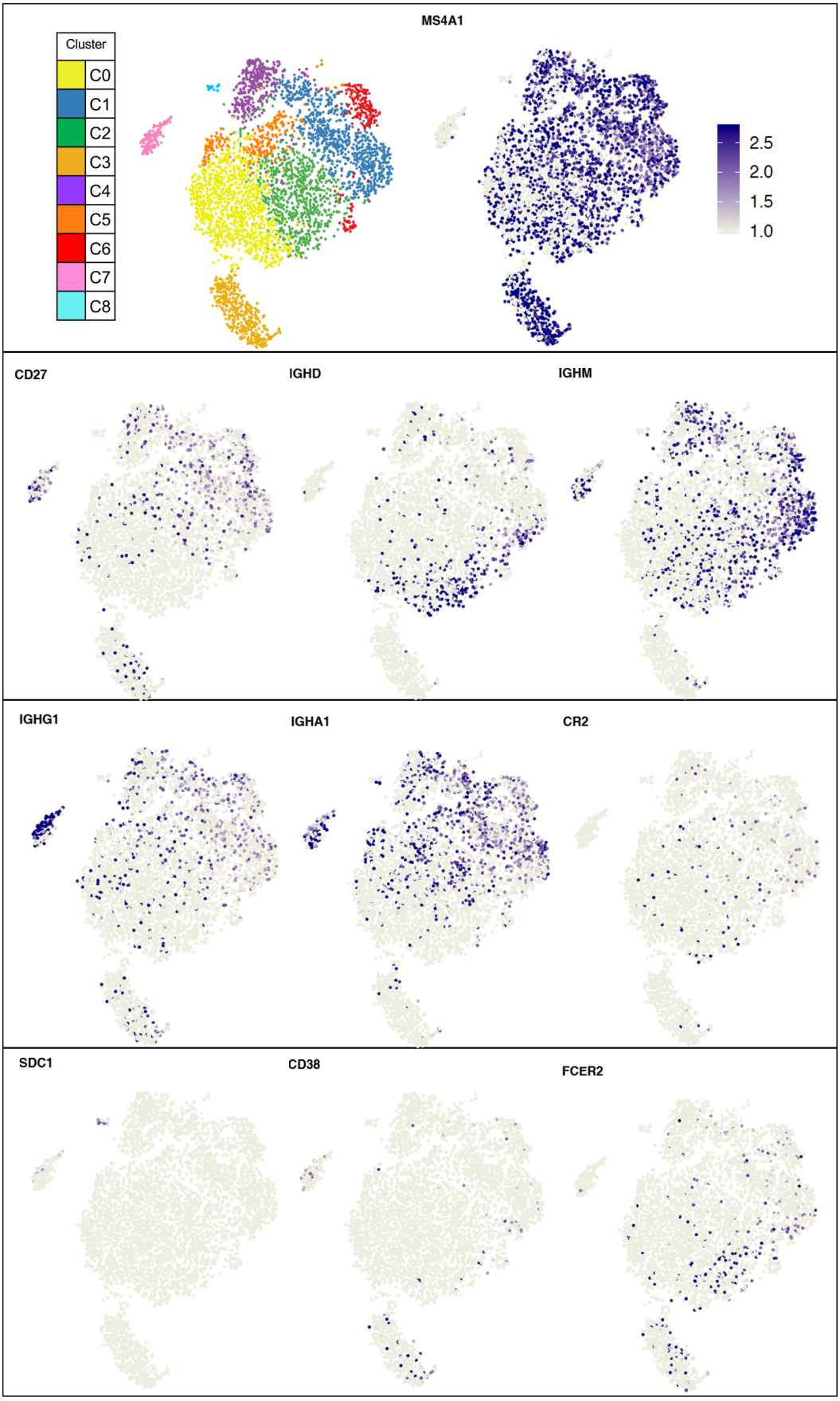


Figure 56. B cell subtype allocation using cell surface phenotypic markers.

Seurat normalised expression is overlaid across the tSNE plots with the expression level represented by the colour scale (blue= highest expression), the C0-C8 cluster allocation tSNE is also shown as a guide. The key B cell markers are able to distinguish possible B cell subtypes.

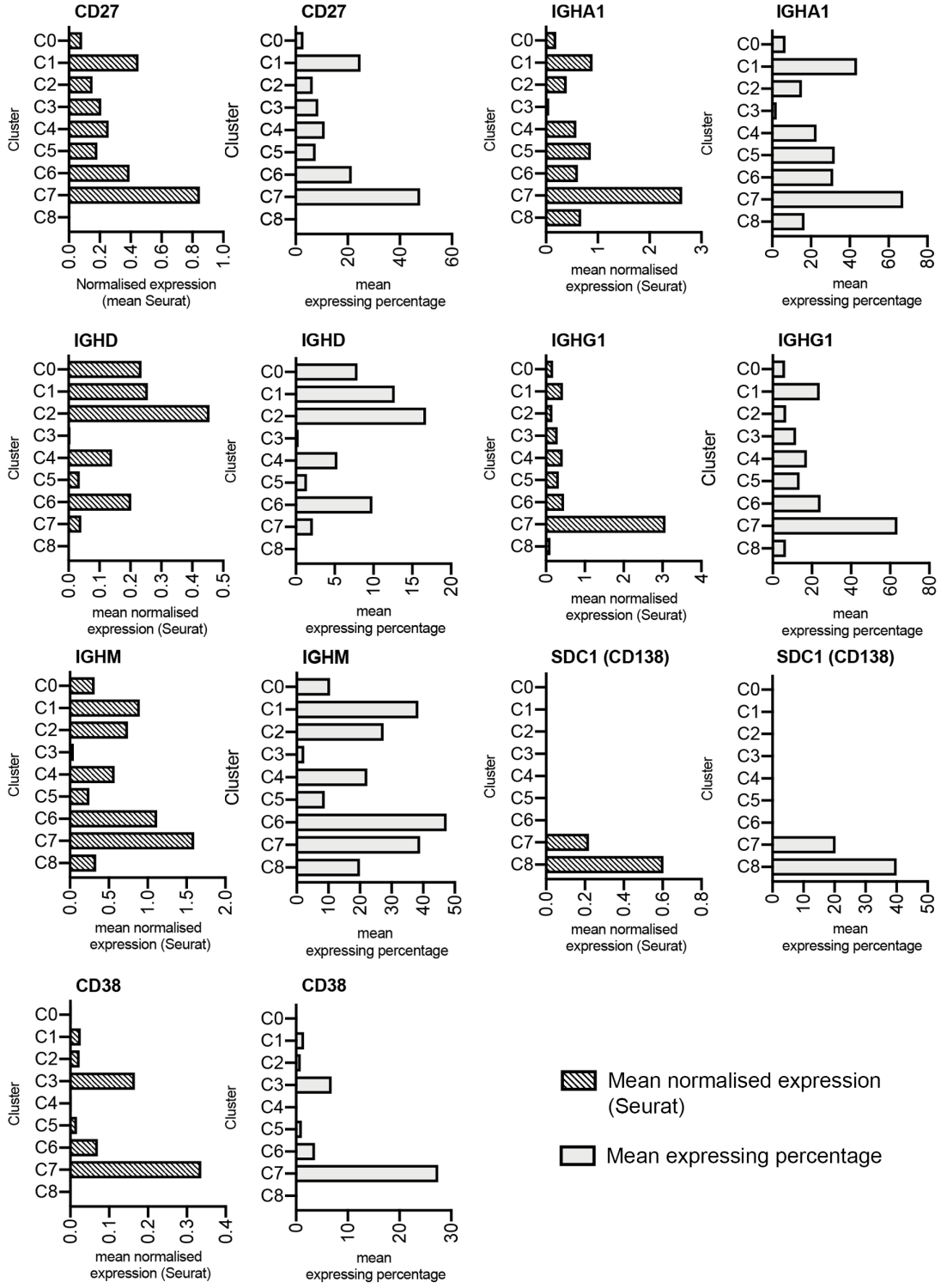


Figure 57. B cell phenotypic marker expression and expressing fraction across the cell clusters.
The B cell subtype markers (CD27-activated, IGHD-naïve, SDC1-plasmablast/ plasma cell, IGHA1 and IGHG1 class switched cells) are displayed as bar plots of mean Seurat expression (striped bar) and mean expressing fraction as a percentage (grey bar). The bar plots complement the B cell subtypes observed in the tSNE.

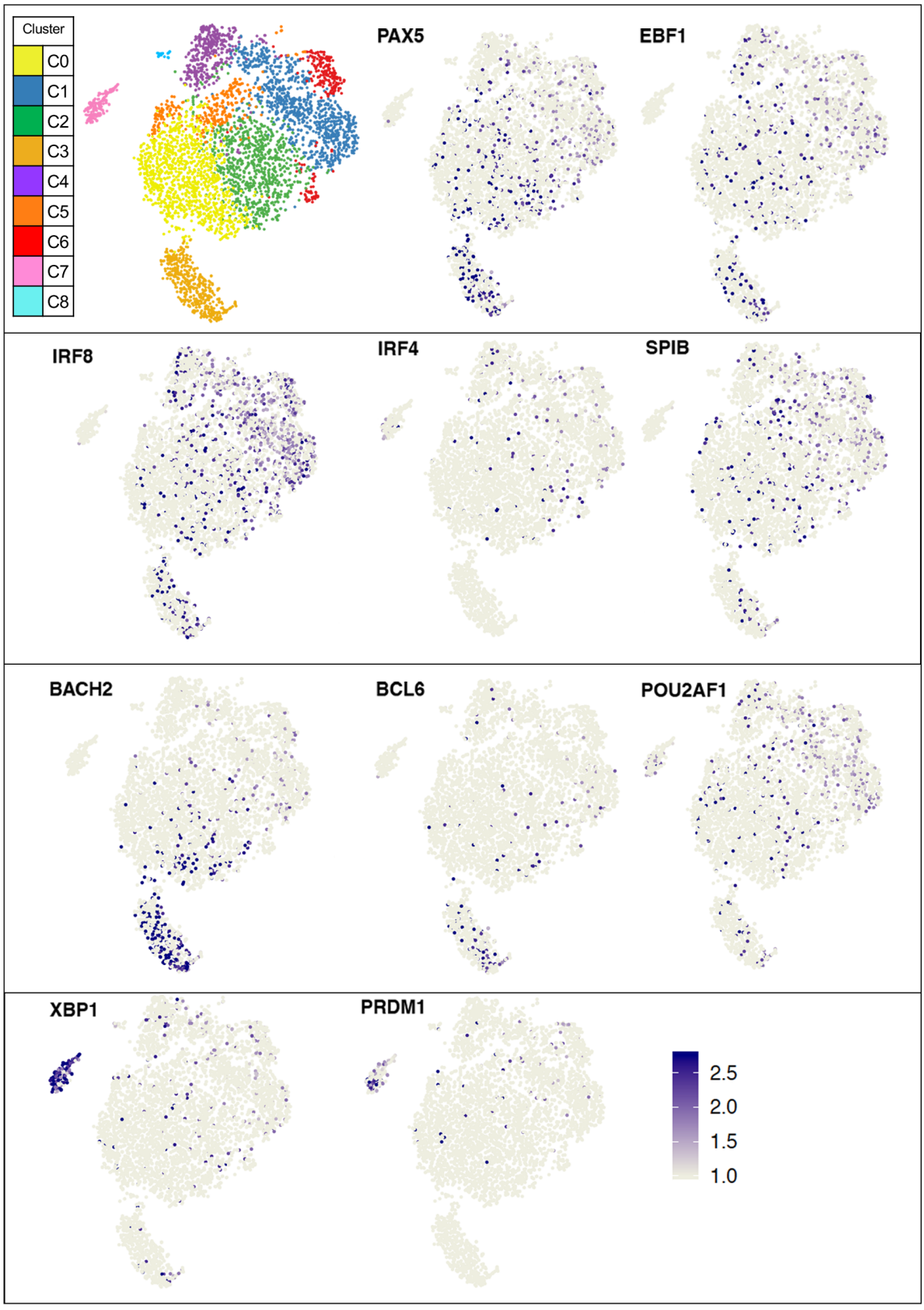


Figure 58. B cell subtype allocation using canonical transcription factors.

Seurat normalised expression is overlaid across the tSNE plots with the expression level represented by the colour scale, the C0-C8 cluster allocation tSNE is also shown as a guide. The B cell transcription factors assist in allocating B cell subtypes (Naïve=PAX5, EBF1; Activated=BACH2, IRF4^{lo}, IRF8; Germinal centre=BCL6, BACH2, PAX5; Memory cells=PAX5, POUF2AF1 (OBF1), SPIB; Plasmablast and plasma cell=IRF4, BLIMP1, XBP1 (Basso and Dalla-Favera, 2015, Nutt et al., 2015).

An additional layer of B cell subtype allocation was achieved with canonical transcription factors (Naïve=PAX5, EBF1; Activated=BACH2, IRF4^{lo}, IRF8; Germinal center=BCL6, BACH2, PAX5; Memory cells=PAX5, POUF2AF1 (OBF1), SPIB; Plasmablast and plasma cell=IRF4, BLIMP1, XBP1) (Basso and Dalla-Favera, 2015, Nutt et al., 2015). Again, these display a continuum of expression across the B cell clusters, except for XBP1 and PRDM1 (Figure 58), violin plots for each marker are shown in Appendix D.2.1. Naïve B cells expressing PAX5 and EBF1 were identified in C0 and C2. Memory B cell features (PAX5, OBF1 (POU2AF2) and SPIB) as well as features of activation (IRF8), were expressed in C1, C2 and C4-C6. A key feature of germinal centre B cells is the expression of BCL6 along with BACH2 and PAX5, which were apparent in C3. The plasmablast and plasma cell genes XBP1 and PRDM1 (BLIMP1) were observed in C7 and C8.

4.3.3.2 B cells from CD8^{Hi} and CD8^{Lo} tumours are enriched in different clusters

It was very interesting to observe that an enrichment of B cells in different clusters depended on whether they were from either CD8^{Hi} (C1 and C6) or CD8^{Lo} (C2, C4 and C5) tumours (Figure 59 A and B). The combination of lineage cell surface markers and transcription factors for B cell subsets (Figure 56 to Figure 58) were used to assign cell subtypes where possible. The cell subtypes for mature naïve/ memory B cells (MB), lymph node germinal centre B cells (LN GC), Plasmablasts (PB) and Plasma cells (PC) are shown in Figure 60. The tSNE plots showing each tissue origin and the corresponding B cell clusters shows enrichment of particular clusters depending on the tissue of origin, the clusters that were enriched in either B cells from CD8^{Hi} or CD8^{Lo} tumours were annotated CD8^{Hi} B1-B2 and CD8^{Lo} B3-B5 (Figure 60). Analysis of the clusters by proportions from each tissue compartment using stacked bar charts (Figure 61) reiterates that B cells in B1 CD8^{Hi} (Blue) and B2 CD8^{Hi} (Red) were enriched in the CD8^{Hi} tumours, whereas the B cells in CD8^{Lo} tumours were enriched in B3 CD8^{Lo} (Green), B4 CD8^{Lo} (Purple) and B5 CD8^{Lo} (Orange). B cells originating from the lymph node dominated LN GC (lymph node germinal centre B cells), with the remaining B cells distributed across MB (mature naïve/ memory B cells), clusters B3, B5 and PB (plasmablast); the non-involved lung B cells were enriched in cluster MB and B5.

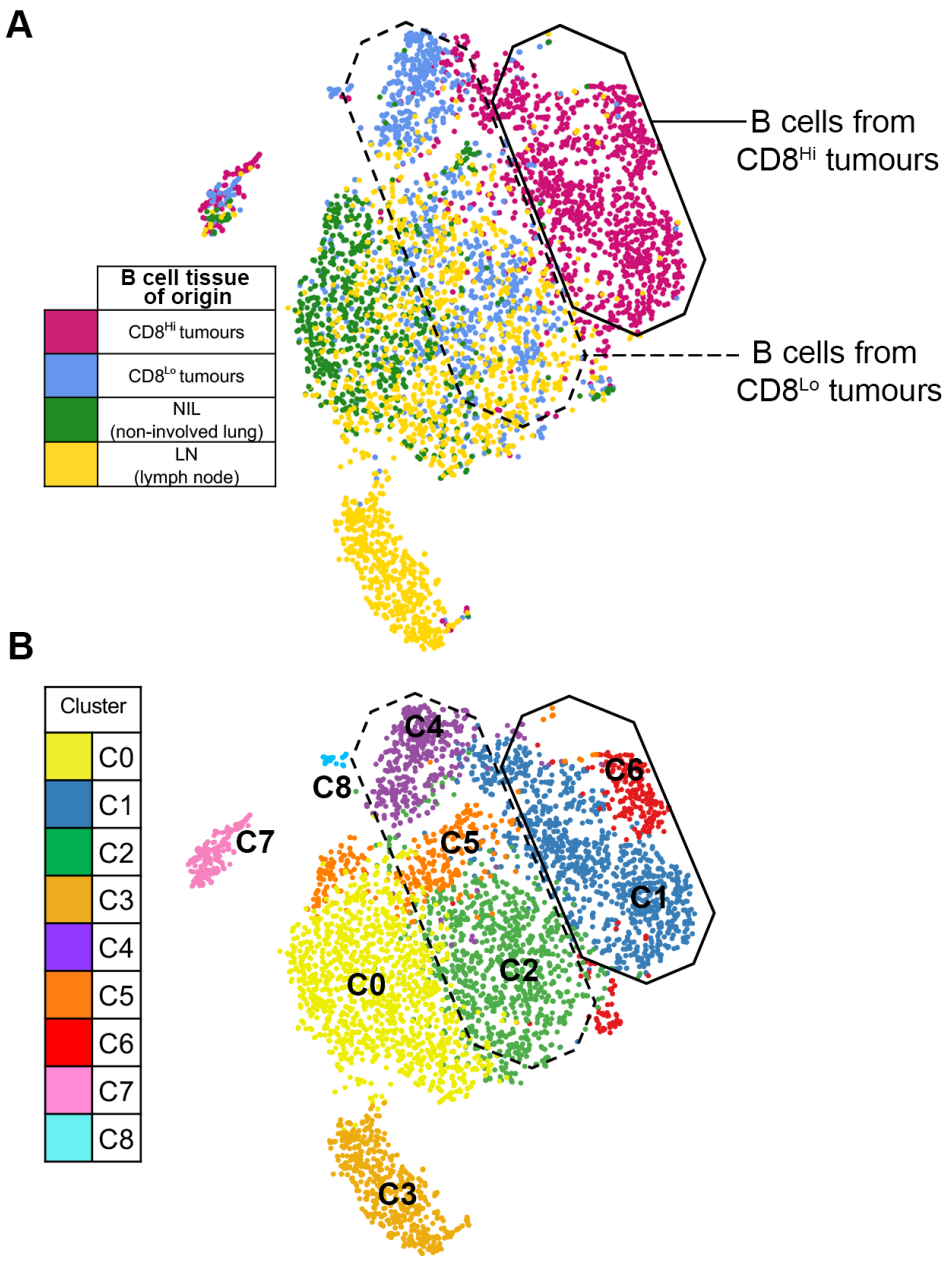


Figure 59. Distribution of clusters from different tissue origins.

(A) tSNE identifying the tissue of origin of B cells as $CD8^{Hi}$ and $CD8^{Lo}$ tumours, non-involved lung (NIL) and lymph node (LN). (B) As a guide the tSNE of the 9 clusters is included. The B cells from $CD8^{Hi}$ (C1 and C6 - solid line) and $CD8^{Lo}$ (C2, C4 and C5 - dashed line) tumours are enriched in different clusters.

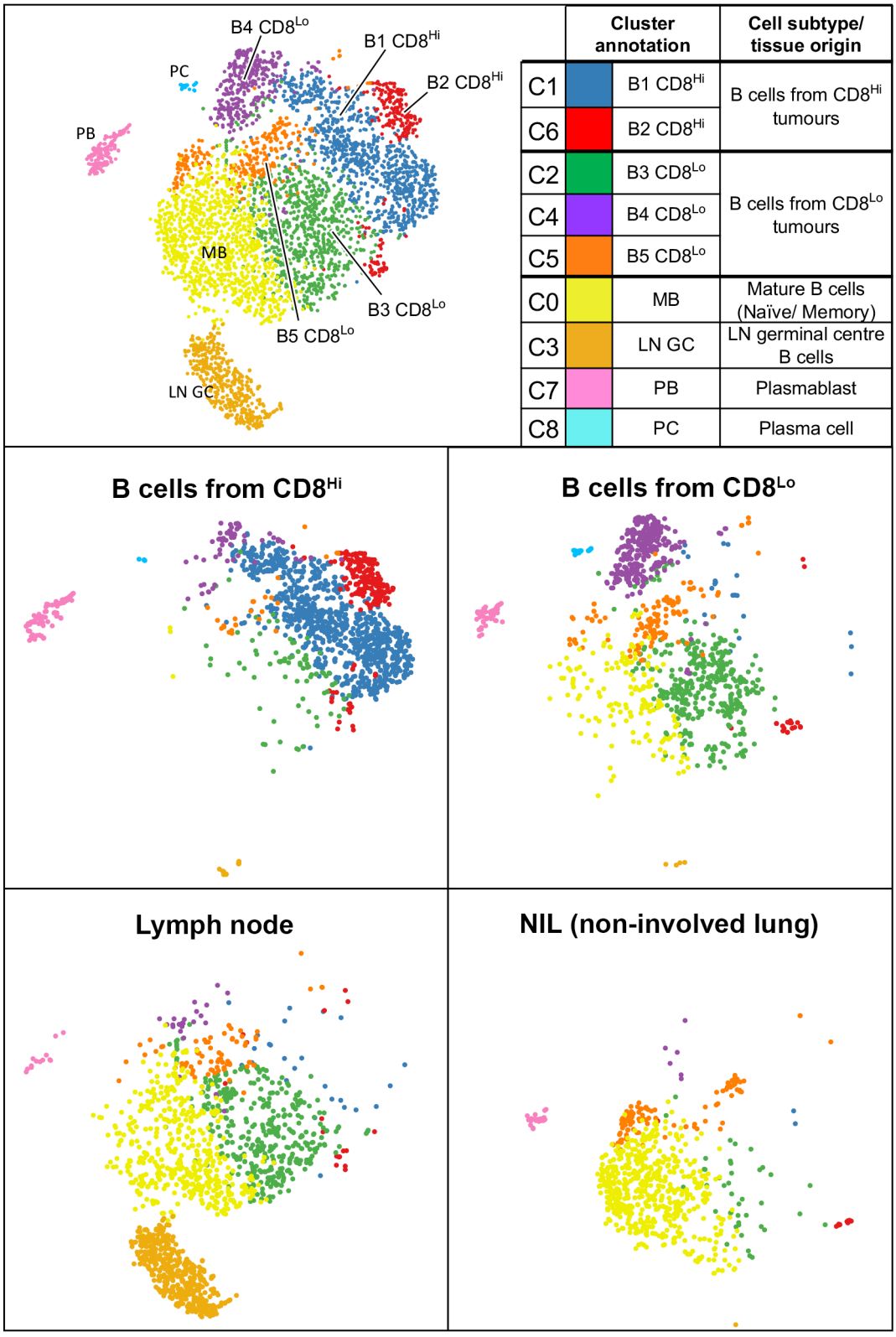


Figure 60. Cell subtype allocation and distribution of cell clusters from different tissue origins.
The Cell subtypes and clusters that were enriched in CD8^{Hi} and CD8^{Lo} tumours are annotated. The contribution of each tissue to the different clusters is shown by tSNE (CD8^{Hi} and CD8^{Lo} tumours, non-involved lung and lymph node). B cells from the CD8^{Hi} tumours were enriched in B1 (Blue) and B2 (Red). The B cells from CD8^{Lo} tumours were enriched in B3 (Green), B4 (Purple) and B5 (Orange). Cell subtypes have also been assigned for mature naïve/memory B cells (MB), lymph node germinal centre B cells (LN GC), Plasmablasts (PB) and plasma cells (PC).

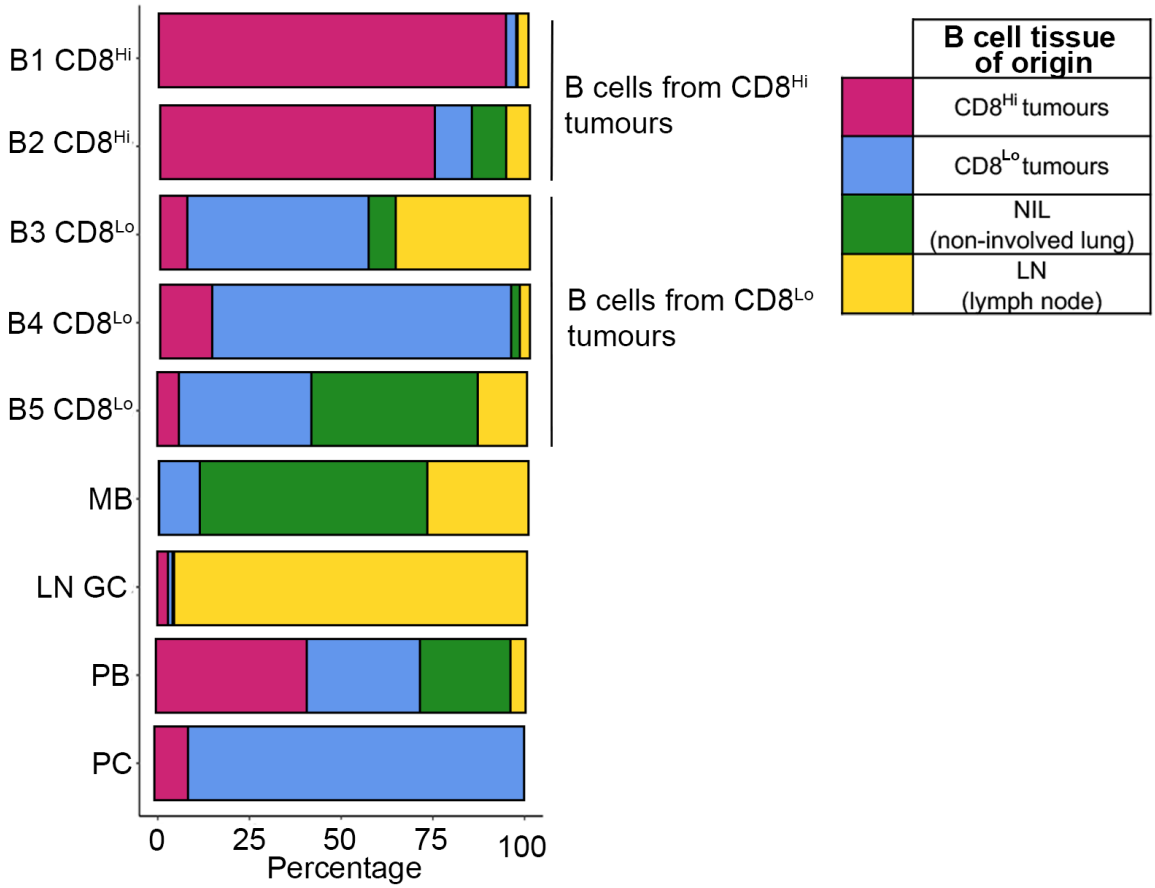


Figure 61. Distribution of cell clusters from different tissue origins.
A representative number of cells (n=756) were used for assessing the proportions of clusters and tissues origin ($CD8^{Hi}$ and $CD8^{Lo}$ tumours, non-involved lung and lymph node). Stacked bar plots showing the percentage contribution of each tissue origin to each cluster. B cells from different tissues display enrichment of clusters (e.g. $CD8^{Hi}$ tumours and cluster B1).

4.3.3.3 Identifying differentially expressed genes in each B cell cluster

In order to identify novel features associated with each cluster, differential gene testing of each cluster was performed (cluster of interest compared to all remaining cells) using MAST (model-based analysis of single cell transcriptomics), a cut-off to identify genes with significantly elevated expression was set as a fold change of >0.25 (log2) and an FDR of <0.05. The number of differentially expressed genes for each cluster is annotated next to the cluster allocations (Figure 62 A). B cells from $CD8^{Hi}$ tumours B1 $CD8^{Hi}$ (n=151 genes) and B2 $CD8^{Hi}$ (n=256 genes), B cells derived from $CD8^{Lo}$ tumours were enriched in clusters B3 $CD8^{Lo}$ (n=41 genes), B4 $CD8^{Lo}$ (n=50 genes) and B5 $CD8^{Lo}$ (n=84 genes); B3 $CD8^{Lo}$ also contained cells derived from the lymph node. Lymph node B cells accounted for >90% in the LN GC cluster (n=334 genes). The top 25 differentially expressed genes for each cluster (Individual clusters compared to remaining cells using MAST) are displayed as a heat map in Figure 62 B, where columns are cells and rows are genes, with the cluster colour annotated above. The gene expression pattern for the clusters was most distinctive in clusters LN GC, PB and PC, this was also mirrored by the spatial positioning of

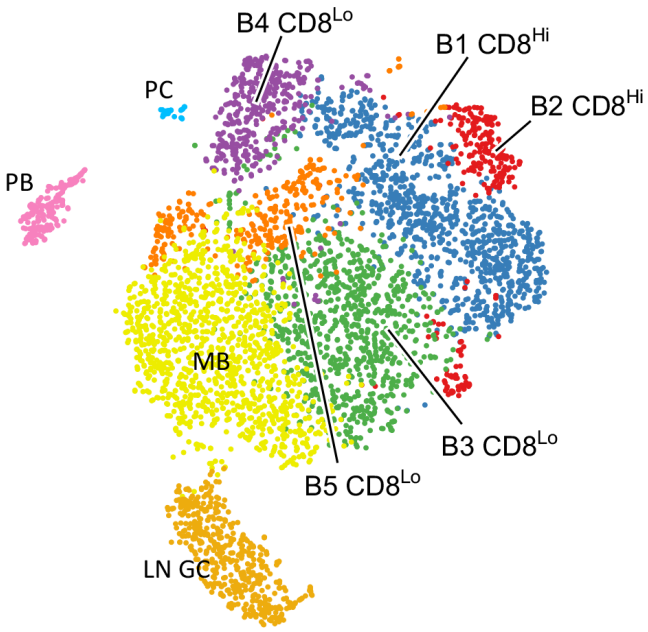
the clusters in the tSNE. The remaining clusters (B1-B5 and MB) share more gene features and occupy a closer relationship by tSNE.

4.3.3.4 B cells display variable features of activation depending on the tissue origin

The differentially expressed genes for each cluster were subjected to gene ontology analysis to identify significant gene families and biological processes using ToppGene Suite (Chen et al., 2009). From the differential gene testing, some common activation features (CXCR4, CD69 and GPR183) were shown to be differentially expressed in clusters B1, B2 $CD8^{Hi}$ and B3 $CD8^{Lo}$; these are displayed as violin plots of normalised expression, tSNE plots with overlaid expression, grey = low; navy blue = high expression and the mean expressing fraction (%) of the cells in each cluster (Figure 63 A-C). The average expression of CXCR4, CD69 and GPR183 was highest in B1 $CD8^{Hi}$ and B2 $CD8^{Hi}$ ($CD8^{Hi}$ tumours), although the average expression was also high in B3 $CD8^{Lo}$, B4 $CD8^{Lo}$ and B5 $CD8^{Lo}$. The genes IGHG1, IGHA1 and IGHM are also displayed, this allows the distribution of isotype switched cells across the clusters to be determined. Expression of the IGH genes was highest in PB (plasmablasts), with IGHG1, IGHA1 and IGHM expression also high in B cell clusters B1 $CD8^{Hi}$, B2 $CD8^{Hi}$ ($CD8^{Hi}$ tumours), B4 $CD8^{Lo}$ and B5 $CD8^{Lo}$ ($CD8^{Lo}$ tumours). The average expression of IGHA1 was 1.5-2x higher than that of IGHG1 in clusters B1-B5, with the highest average expression in B1 $CD8^{Hi}$ (excluding the PB cluster).

A

MAST differentially expressed genes	
Cluster Id	No. of genes
B1 CD8^{Hi}	151
B2 CD8^{Hi}	256
B3 CD8^{Lo}	41
B4 CD8^{Lo}	50
B5 CD8^{Lo}	84
MB	30
LN GC	334
PB	335
PC	1037



B

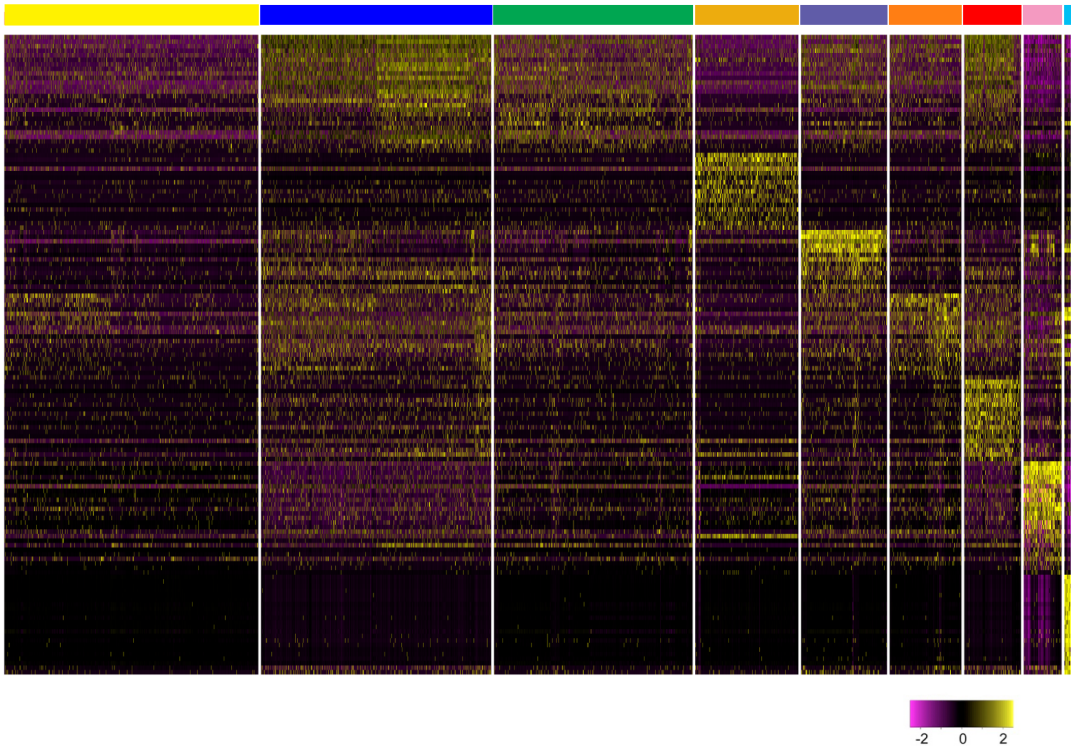


Figure 62. Summary of the differentially expressed genes in each cluster.
Differentially expressed genes were identified by MAST comparing each cluster to the remaining cells, significant genes were identified with a log2 fold change > 0.25 and FDR <0.05. (A) lists the number of differentially expressed genes identified in each cluster alongside a tSNE of the clusters. (B) shows the top 25 expressed genes for each cluster identified from the MAST analysis.

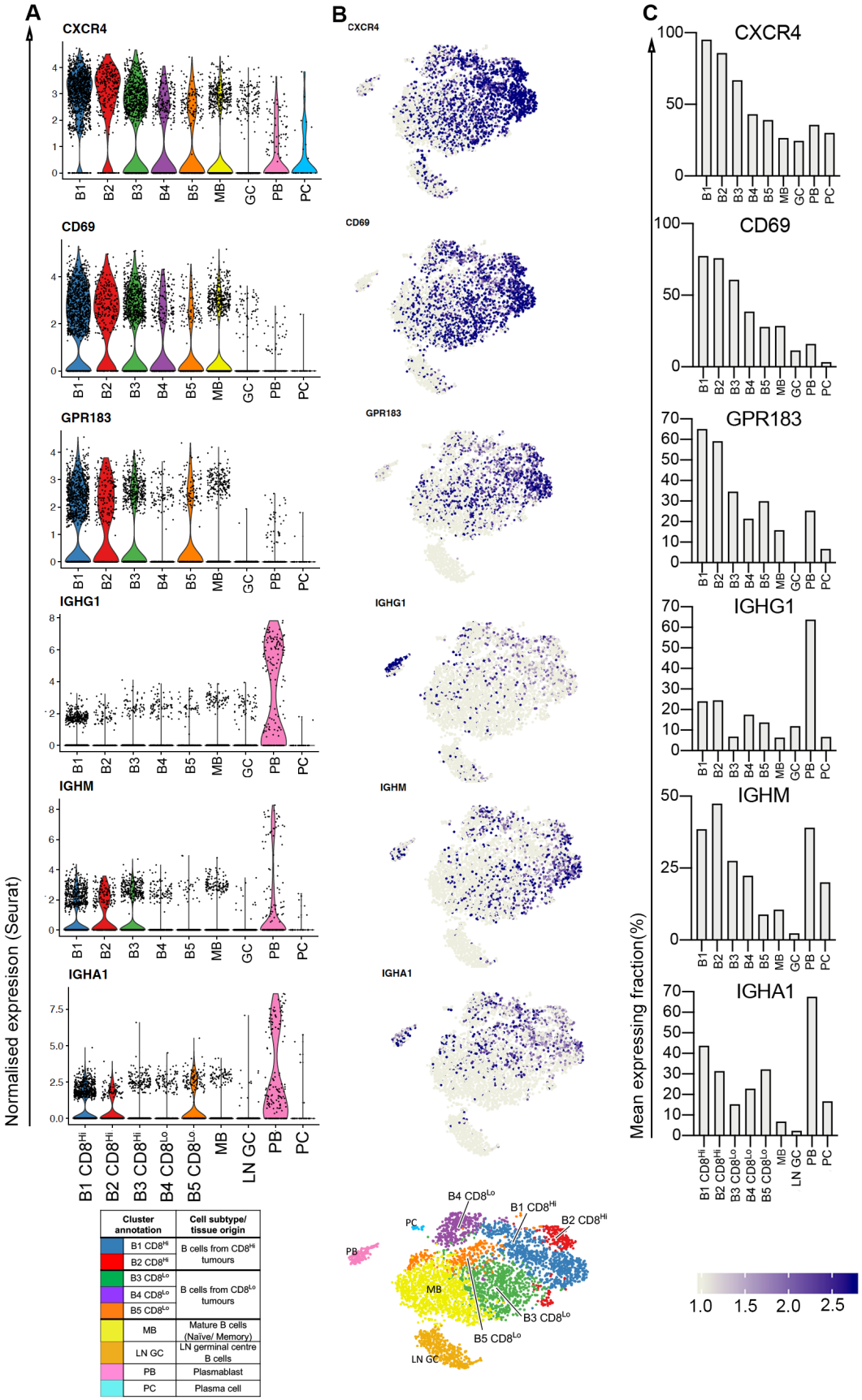


Figure 63. Highly expressed markers of activation and isotype switching.

(A) Violin plots of Seurat normalised expression for the activation markers CXCR4, CD69 and GPR183 and the IGH genes IGHG1, IGHA1 and IGHM as an indication of B cell class switching across the clusters. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes.

4.3.3.5 B cells from CD8^{Hi} and CD8^{Lo} tumours display distinct characteristics

4.3.3.5.1 Gene ontology analysis of clusters enriched in CD8^{Hi} tumours (B1 and B2)

Since B cells from CD8^{Hi} tumours formed the major part of B1 CD8^{Hi} (Blue) and B2 CD8^{Hi} (Red), the differentially expressed genes in each cluster were subjected to gene ontology analysis to identify biological processes and gene families. The gene families in B1 CD8^{Hi} consisted of a variety of cluster of differentiation (CD) molecules and tumour necrosis super family (TNFS) genes, MHC-I and II, Actins and phosphatases that dephosphorylate MAP kinase; the terms and genes are summarised in Appendix D.3.4. Significant gene ontology terms associated with biological processes (FDR<0.05) were subsequently reduced using REVIGO (Appendix D.3.4). Immunological GO terms consisted of interferon-gamma-mediated signalling, response to cytokine, cell activation and biological adhesion, summarised in Appendix D.3.4.

The same analysis rationale was applied to B2 CD8^{Hi}, where significant (FDR<0.05) gene families and gene ontologies (reduced by REVIGO) are summarised in Appendix D.3.5. Biological processes consisted of a response to type I interferon, response to IFNG, regulation of immune response and presentation of exogenous peptides. B cell-specific terms were also identified for B cell proliferation, BCR signalling, B cell differentiation and B cell apoptotic process.

4.3.3.6 Distinctive adaptive immune response features in B cells from CD8^{Hi} tumours

Differentially expressed genes that linked to an adaptive immune response in B1 CD8^{Hi} and B2 CD8^{Hi} were evaluated in greater detail. The gene functions in B cells were assessed using the literature, presented in Table 28 and Table 29 respectively. As previously stated, some overlapping genes (e.g. CXCR4, CD69 and GPR183) from the MAST analysis (cluster of interest compared to all remaining cells) were differentially expressed in several clusters (B1 CD8^{Hi}, B2 CD8^{Hi} and B3 CD8^{Lo}) (section 4.3.3.4). However, these genes displayed the highest average expression in cluster B1 CD8^{Hi} and are associated with B cell activation and GC interactions (CXCR4, CD69, GPR183, MHC1-I and MHC-II, IRF8 and TNFRSF13B), as well as features of germinal centre formation (LTB and ACTG1). The expression of TNFRSF13B (TACI), IRF8, ZFP36L2 and CD48 are displayed as violin plots, tSNE plots and mean expressing fraction (%), where the average expression is higher in B1 CD8^{Hi}, with a reduced number of cells expressing these features in B2 CD8^{Hi} (Figure 64 A-C). The average expression of these targets was reduced but not absent in the remaining clusters B3-B5 and MB. Differentially expressed genes in B cell cluster B2 CD8^{Hi} displayed features of an on-going germinal centre reaction (CD40, CD53, TNFSF10 and TNFRSF14), as well as increased negative regulators of the B cell response via negative feedback (FCGR2B, FCRL5, LGALS9). The genes SRGN and LGALS9 (16% of cells) were elevated in B2 CD8^{Hi} and may represent interesting targets that can modulate the functionality of T cells in the tumour

microenvironment. A key finding in cluster B2 CD8^{Hi} were the interferon stimulated genes (ISGs), which covered a large number of the differentially expressed genes (n=30), and were part of the interferon activation gene ontology (biological processes). The ISGs displayed high average expression across the cells in the B2 CD8^{Hi} (e.g. IFI44L, STAT1 and IRF7); Violin plots, tSNE plots and expressing fraction are shown in Figure 65 A-C for the genes IFI44L, STAT1, FCRL5, IL2RG, CD40 and SRGN. A negative regulator of the B cell response FCRL5 was also elevated in both B2 CD8^{Hi} and PB cluster, with the common gamma chain for the interferon receptor (IL2RG) displaying high average expression across B1 CD8^{Hi}, B2 CD8^{Hi} and PB. The proteoglycan Serglycin was expressed in several clusters with the highest average expression in B1 CD8^{Hi}, B2 CD8^{Hi} and PB.

B1 CD8^{Hi} (B cells from CD8^{Hi} tumours)

Gene	Av. FC	Summary	Ref.
CXCR4	0.94	CXCR4 is the receptor for CXCL12 and functions to locate GC B cells into the dark and light zones, it is also linked with cell cycle and coincides with BCR signalling in the GC followed by AID activity.	(Weber, 2018)
MHC-I and MHC-II	0.78/ 0.52	MHC-I and MHC-II (HLA-A, HLA-C, HLA-DPA1 , HLA-DPB1, HLA-DQA1, HLA-DQB1, HLADRA, HLA-DRB1, HLAE) represent activation status and the ability to present antigen to both CD4+ and CD8+ T cells.	(Katikaneni and Jin, 2019, Yuseff et al., 2013)
DUSP1 and DUSP2	0.78/ 0.35	DUSP1 and DUSP2 (Dual Specificity Phosphatase 1/2) are linked to pro-apoptotic features and are implicated as molecular controls to attenuate immune effector functions.	(Shen et al., 2017)
GPR183	0.73	GPR183 (G Protein-Coupled Receptor 183; EBI2, Epstein-Barr virus-induced gene 2) increases during the BCR activation and CD40 co-stimulation. Essential for migration of B cells to extrafollicular sites and induction of plasmablasts. Downregulation promotes B cell migration to the centre of follicles and GC formation.	(Gatto et al., 2009)
CD69	0.68	CD69 is an early activation marker, stopping egress form lymphoid tissues by inhibiting S1P1, it leads to a more stable CD4+ Tfh cell interaction.	(Shiow et al., 2006, Ise et al., 2018)
ACTG1	0.58	ACTG1 (actin gamma 1) plays a role in the immunological synapse and activation of B cells.	(Burbage and Keppler, 2018)
KLF2	0.5	KLF2 (Krupple like factor 2) KLF2 controls B cell localization its absence leads to diminished FO B cells and plasma cells.	(Winkelmann et al., 2011)
ZFP36L2	0.47	ZFP36L2 is a transcription factor that is critical for re-establishing quiescence and maintenance of genomic integrity during B cell development and SHM and CSR.	(Galloway et al., 2016)
CD48	0.44	CD48 (part of SLAM; signalling lymphocyte activation molecules) is required for the effective differentiation of B cells and T-dependent Ig production, Nk cells expressing CD48 can activate B cells leading to T-independent Ig.	(Hoffmann et al., 1998, Yuan et al., 2013)
SAMSN1	0.4	SAMSN1 Negative regulator of B cell activation that is elevated in lymphoid B cells, co-expressed with CD27.	(Chokeshai-u-saha et al., 2012)
CD37	0.39	CD37 (TSAPN26; Tetraspanin 26) forms complexes with MHC-II and is important for T and B cell interactions leading Ig production.	(van Spriel, 2011)
SELL	0.35	SELL (CD62L) expressed on Naïve lymphocytes and is subsequently lost from the cells surface upon BCR activation.	(Morrison et al., 2010)
LTB	0.34	LTB (lymphotoxin beta) plays a role in B cells by tissue remodelling and inducing the formation of TLS.	(Shen and Fillatreau, 2015)
IRF8	0.3	IRF8 is required for formation of follicular and GC B cells, loss leads to impaired Ig production. IRF8 regulates the BACH2 and FCER2 that maintain the FO B cell phenotype, it also regulates BCL6 a fundamental part of GC B cells.	(Wang et al., 2019)
TNFRSF13B	0.27	TNFRSF13B (Transmembrane activator and calcium-modulator and cytophilin ligand interactor; TACI) is expressed in switched memory B cells and is upregulated upon activation via BCR and TLR.	(Castigli et al., 2005, Smulski et al., 2017)
SOCS1	0.26	SOCS1 (Suppressor Of Cytokine Signalling 1) inhibits the JAK-STAT1 pathway and is induced by cytokines (IFNG) and acts in a negative feedback loop to attenuate cytokine signalling.	(Zhou et al., 2007)

Table 28. Evaluation of genes identified by gene ontology analysis in B1 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours).

The targets identified from the gene ontology analysis in B1 CD8^{Hi} were investigated further to assign the potential functional roles in B cells. The fold change (log2), function and supporting evidence are outlined.

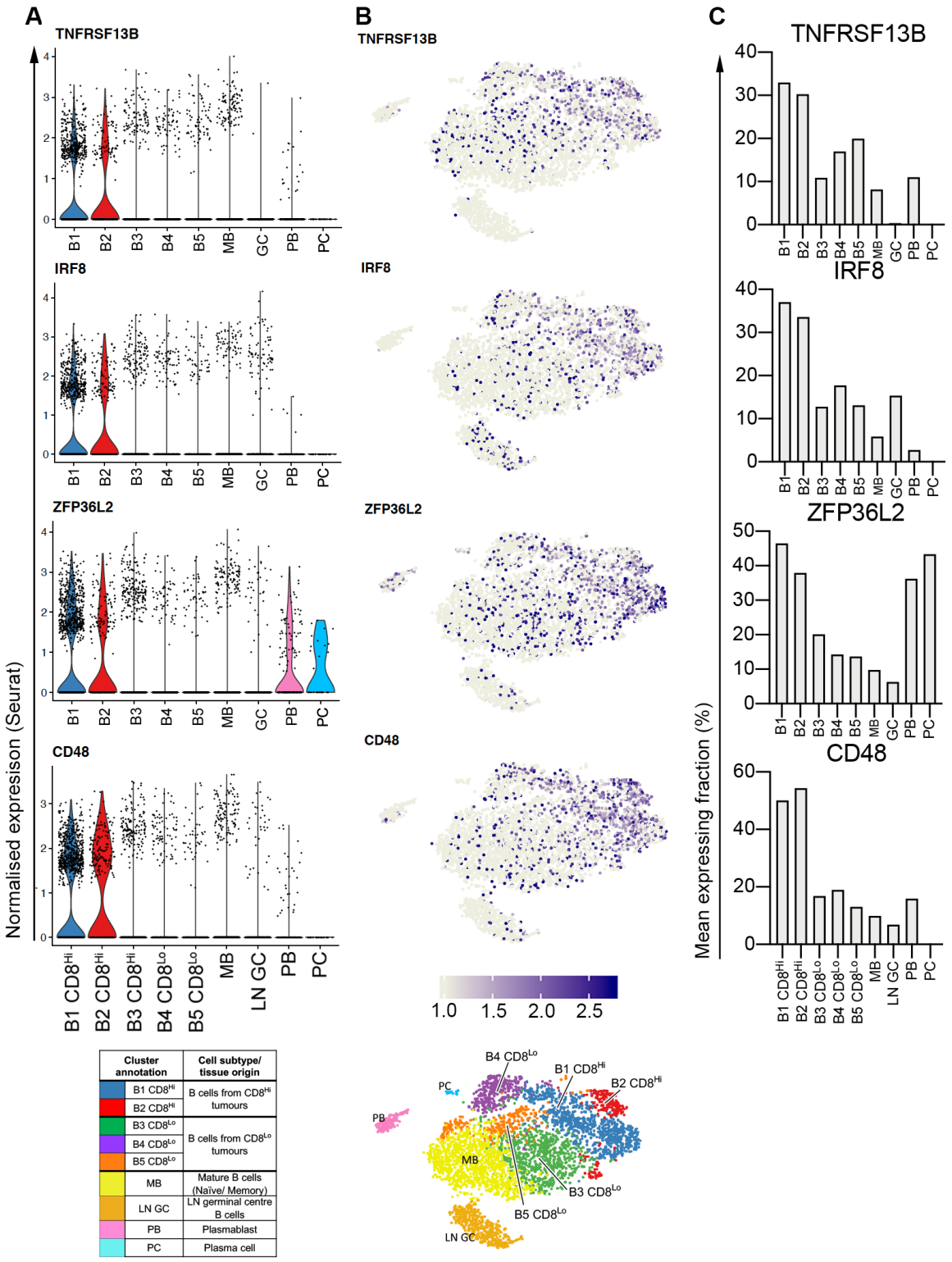


Figure 64. Expression of differentially expressed B cell features identified in B1 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours).

(A) Violin plots of Seurat normalised expression for the genes TNSRSF13B, IRF8, ZFP36L2 and CD48 across the clusters. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. An activation profile exists in C1 that indicates germinal centre and memory B cells, this was also observed in C6 but to a lesser extent.

B2 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours)

Gene	Av. FC	Summary	Ref.
ISGs	1.48/ 1.46/ 1.1	Interferon stimulated genes (ISGs) that result from a response to IFN binding cell surface receptors, it leads to signalling via JAK:STAT1 and ISG expression. ISGs=PSMB8, PSMB9, IRF2, IRF7, ISG20, PSMD13, PSME1, PSME2, IRF9, IFITM1, IFITM2,IFI44L, IFI16, IFI35, IFIT2, IFIT3, MT2A, ISG15, IFI6, MX1, MX2.	(Schoggins and Rice, 2011) (Schneider et al., 2014) (Vazquez et al., 2015) (Jackson et al., 2015)
STAT1	0.97	STAT1 signalling along with INF signalling are fundamental to the formation germinal centres.	(Domeier et al., 2016)
CXCR4	0.7	See table for C1	
FCRL5	0.62	FCRL5 functions as a regulator of the B cell response by binding to the Fc region of IgG, forming a negative feedback loop. It is expressed on activated B cells and memory B cells.	(Wilson et al., 2012, Kim et al., 2019a)
TNFSF10	0.59	TNFSF10 (TRAIL) induces apoptosis upon ligation, BCR activation and CD40 co-stimulation abrogates this process. TRAIL therefore induces apoptosis in poorly stimulated cells and helps maintain B cell homeostasis.	(Guerreiro-Cacais et al., 2010)
CD69	0.51	See table for C1	
GPR183	0.48	See table for C1	
SAMSN1	0.48	See table for C1	
CD164	0.48	CD164 expression modulates the proliferation, adhesion and migration of haemopoietic cells, it interacts with CXCR4 and enhances cells motility.	(Forde et al., 2007, Huang et al., 2013a)
SOCS3	0.47	SOCS3 (Suppressor Of Cytokine Signalling 3) is a key regulator of STAT3-dependent cytokine responses, It's expression is linked to IL10 production.	(Stanic et al., 2015, Jones et al., 2011)
LGALS9	0.4	LGALS9 expressed on activated B cells, binding to Gal-9 leads to suppression on of BCR in memory B cells. Also has the potential to regulate the functionality of T cells by binding to HAVCR2 (TIM3) inhibiting their survival and function.	(Zhu et al., 2005, Giovannone et al., 2018, Cao et al., 2018)
SRGN	0.38	SRGN (Proteoglycan serglycin) forms a macromolecular complex with granzymes and perforin. Extracellularly secreted where it can enhance immunity. Also implicated in epithelial mesenchymal transition and promotion of cancer cell invasion and metastasis.	(Kolset and Pejler, 2011, Korpetinou et al., 2014, Xu et al., 2018, Zhang et al., 2017) (Li et al., 2011)
CCR7	0.33	CCR7 is upregulated upon activation via BCR, makes B cells more responsiveness to CCL21 maintaining them at the immunological synapse.	(Pereira et al., 2010)
FCGR2B	0.31	FCGR2B a receptor for IgG, it blocks B cell activation and induces apoptosis via a negative feedback loop.	(Jhou et al., 2017)
CD53	0.31	CD53 (TSPAN25) complexes with integrins and promotes BCR dependent activation.	(Zuidscherwoude et al., 2017)
IL2RG	0.31	IL2RG (Interleukin 2 Receptor Subunit Gamma) is a signaling component of many interleukin receptors, including interleukin -2, -4, -7 and -21, leading to cytokine sensing and B cell differentiation.	(Berglund et al., 2013)
CD40	0.3	CD40 is expressed on B cells and APCs, the binding of antigen by T cells presented by B cells requires co-stimulation via CD40 and CD40LG expressed CD4 cells for effective activation of both cell types.	(Elgueta et al., 2009)
TNFRSF14	0.29	TNFRSF14 (herpesvirus entry mediator; HVEM) interacts with BTLA on T cells (CD4+ Tfh) and restrains the GC response by impeding the TCR signalling in Tfh cells leading to lower CD40LG co-stimulation.	(Mintz et al., 2019)

Table 29. Evaluation of genes identified by gene ontology analysis in B2 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours).

The targets identified from the gene ontology analysis in B2 CD8^{Hi} were investigated further to assign the potential functional roles in B cells. The fold change (log2), function and supporting evidence are outlined.

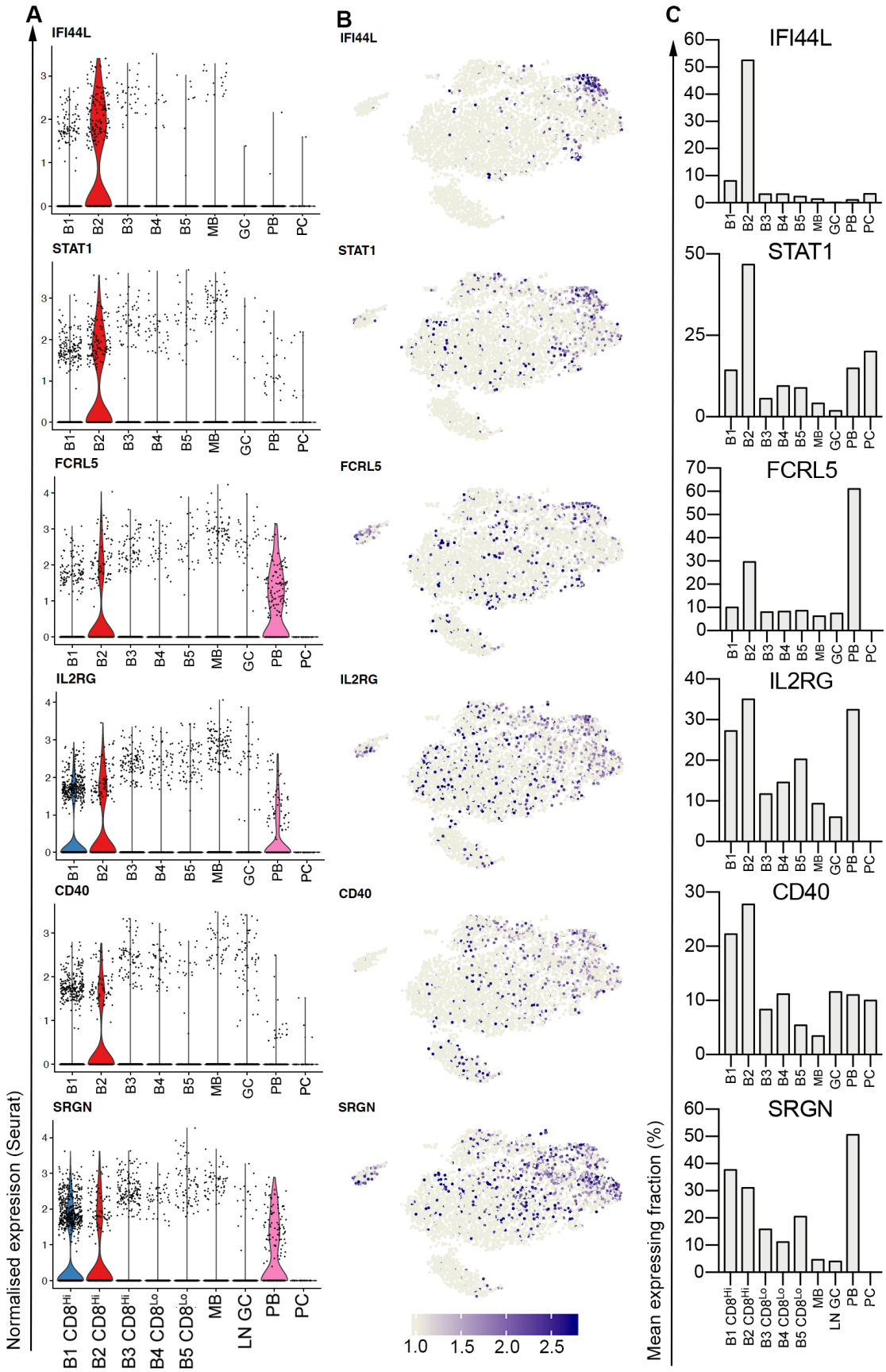


Figure 65. Expression of differentially expressed B cell features identified in B2 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours).

Figure 65 legend.

Cluster annotation		Cell subtype/ tissue origin
	B1 CD8^{Hi}	B cells from CD8^{Hi} tumours
	B2 CD8^{Hi}	
	B3 CD8^{Lo}	B cells from CD8^{Lo} tumours
	B4 CD8^{Lo}	
	B5 CD8^{Lo}	
	MB	Mature B cells (Naïve/ Memory)
	LN GC	LN germinal centre B cells
	PB	Plasmablast
	PC	Plasma cell

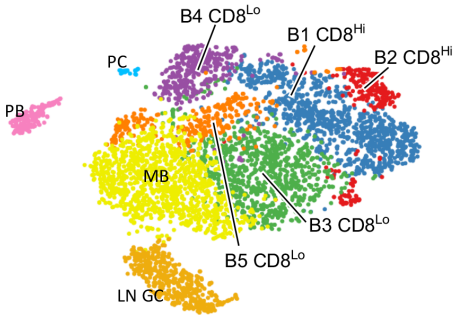


Figure 65. Expression of differentially expressed B cell features identified in B2 CD8^{Hi} cluster (B cells from CD8^{Hi} tumours).

(A) Violin plots of Seurat normalised expression for the genes IFI44L, STAT1, FCRL5, IL2RG, CD40 and SRGN across the clusters. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. Some of the featured genes in the B2 CD8^{Hi} B cell cluster highlight that an interferon response signature was present (IFI44L and STAT1), including an interferon receptor (IL2RG), along with CD40 indicating co-stimulation and response to potential environmental cytokines. Elevated SRGN was also observed and may be an interesting immunomodulatory target.

4.3.3.7 Gene ontology analysis of B cells enriched in CD8^{Lo} tumours (B3 CD8^{Lo}, B4 CD8^{Lo} and B5 CD8^{Lo})

Differentially expressed genes in clusters B3 CD8^{Lo}, B4 CD8^{Lo} and B5 CD8^{Lo} (enriched in CD8^{Lo} tumours) were subjected to gene ontology analysis of biological process and gene families, summarised in Appendix D.3.6 to D.3.8 respectively. The analysis of B3 CD8^{Lo} identified biological processes associated with CD4+ T cell activation, regulation of cell adhesion and regulation of leukocyte activation; IGHD, a marker of naïve cells, also displayed higher average expression in B3 CD8^{Lo} (Appendix D.3.6). Gene ontologies associated with B4 CD8^{Lo} were involved in cellular response to TNF and cellular response to stress/ response to unfolded protein. The gene families identified in B4 CD8^{Lo} were related to heat shock proteins (Appendix D.3.7). The gene families in cluster B5 CD8^{Lo} were linked with calcium binding, along with biological processes indicating regulation of cell adhesion, cell growth, protein localisation and immune response-activating cell surface receptor signaling pathways (Appendix D.3.8).

4.3.3.8 Distinctive adaptive immune response features in B cells from CD8^{Lo} tumours

The differentially expressed genes identified in clusters B3 CD8^{Lo}, B4 CD8^{Lo} and B5 CD8^{Lo} were explored in more detail relative to the literature and are summarised Table 30 to Table 32. The key features in B3 CD8^{Lo} were those of an enrichment of naïve mature B cell genes (IGHD and FOXP1). The genes NR4A2, FOXP1, CD83 and IGHD displayed higher average expression in B3 CD8^{Lo}, although expression was also observed in B1 CD8^{Hi} and B2 CD8^{Hi} Figure 66 A-C). In Cluster B4 CD8^{Lo}, heat shock proteins (HSPs) dominated the signal and were expressed at significantly

higher levels in the B cells from $CD8^{Lo}$ tumours relative to $CD8^{Hi}$ tumours. HSP expression (DNAJB1, HSPA1A and HSPB1) was also present in the PB and PC clusters, although at lower levels than in cluster B4 $CD8^{Lo}$ (Figure 67 A-C). These function to facilitate the correct folding of proteins, they are also induced by cellular stress and play a role in pro- and anti-inflammatory responses. A signature of B cell activation was detected in B5 $CD8^{Lo}$ with the expression of Ca^{2+} channel proteins (S100A4 and S100A10) and the activation marker CD69 (Figure 68 A-C). Elevated average expression of LGALS1 was also observed in B5 $CD8^{Lo}$, this has been linked to regulation of B cells and can also have negative effects on T cell functionality. LGALS1 was also identified at elevated levels in the PB and PC clusters, indicating its sustained expression through different B cell differentiation states.

B3 $CD8^{Lo}$ cluster (B cells from $CD8^{Lo}$ tumours)

Gene	Av. FC	Summary	Ref.
NR4A2	0.65	NR4A2 (Nuclear Receptor Subfamily 4 Group A Member 2) functions as a transcription factor in early B cell development, its absence causes disruption to B cell lymphogenesis.	(Miyai et al., 2018)
IGHD	0.65	IGHD expression is a feature of mature B cells along with IGHM. Upon activation with antigen B cells undergo loss of IGHD and with co-stimulation class switching. Surface IgD has also been implicated in tolerance where it inhibits responses to potential autoantigens.	(Gutzeit et al., 2018, Quach et al., 2011)
CD83	0.55	CD83 is a membrane Ig receptor expressed on recently activated B cells and in GC B cells, aberrations in CD83 can lead to altered antibody production and GC responses.	(Krzyzak et al., 2016)
CD69	0.38	CD69 broad activation marker highlighted previously in C1 table.	(Shiow et al., 2006, Ise et al., 2018)
CD44	0.278	CD44 involved in cell adhesions and ECM interactions as well as interacting with cytokines and growth factors in the tumour microenvironment.	(Morath et al., 2016)
FOXP1	0.25	Essential transcription factor in early B cell development and expressed on mature B cells.	(Patzelt et al., 2018) (Hu et al., 2006)

Table 30. Evaluation of genes identified by gene ontology analysis in B3 $CD8^{Lo}$ cluster (B cells from $CD8^{Lo}$ tumours).

The targets identified from the gene ontology analysis in B3 $CD8^{Lo}$ were investigated further to assign the potential functional roles in B cells. The fold change (log2), function and supporting evidence are outlined.

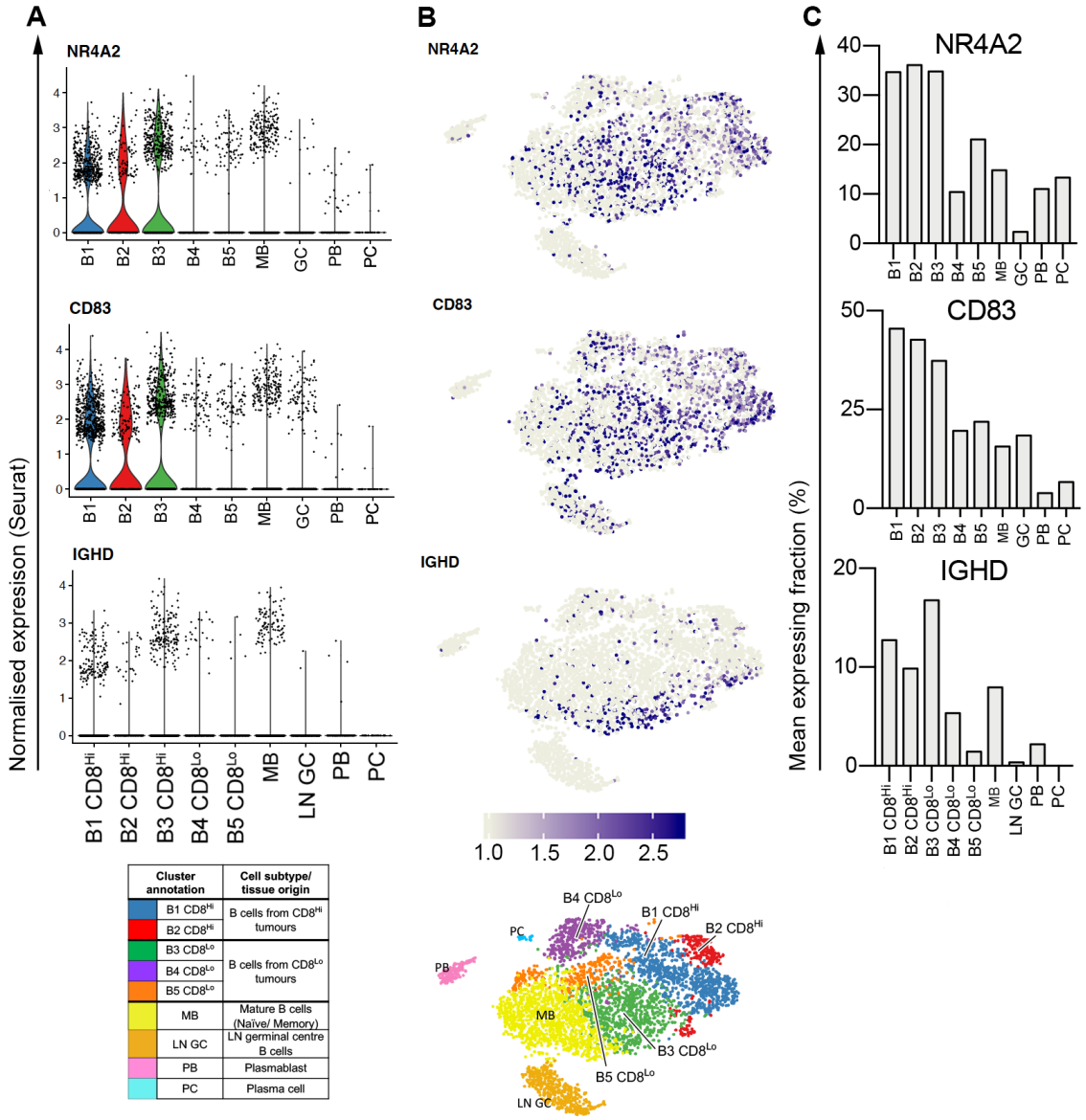


Figure 66. Expression of differentially expressed B cell features identified in cluster B3 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours).

(A) Violin plots of Seurat normalised expression for the genes NR4A2, FOXP1, CD83 and IGHD across the clusters. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. The expression IGHD was higher in B3 CD8^{Lo} and indicates an enrichment of naïve B cells.

B4 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours)

Gene	Av. FC	Summary	Ref.
ID3	0.97	ID3 (Inhibitor Of DNA Binding 3) is expressed in follicular B cells and then declines in GC B cells, it is proposed that the decline of ID3 leads to effective BCR activation and subsequent signalling.	(Chen et al., 2016)
CACYBP	1	CACYBP (Calcyclin Binding Protein) functions to degrade Beta-catenin, a regulator of cell communication and associated with differentiation of cell (CD4+ T cells).	(Fukushima et al., 2006)
GADD45B	0.69	GADD45B (Growth Arrest And DNA Damage Inducible Beta) Is upregulated by CD40 activation and functions to protect B cells from FAS-induced programmed cells death. It is found in GC B cells and is crucial pro-survival mediator that blocks FAS cytotoxicity.	(Zazzeroni et al., 2003)
HSP40	2.08	HSP40 may play a role in the production of IL10 and inhibition on T cell replication.	(Tukaj et al., 2010)
HSP70	2.05	HSP70 can directly regulate immune responses via IFN-γ and TNF-α production as well as downregulating CD86 and MHC class II.	(Milani et al., 2002)
HSP90	1.46	HSP90 is involved intra and extra cellular antigen presentation via MHC-I and -II as well as modulating the IFN response.	(Mbofung et al., 2017, Graner, 2016)
HSPs	2.08/ 2.05/ 0.91	Heat shock proteins (HSPs)/ cellular response to stress. GADD45B, DNAJB1, HSPH1, PRDX1, DNAJB6, PPP1R15A ,CHORDC1,DNAJA1, HSPA1A, HSPA1B, PTGES3, HSPA8 ,HSPB1, PMAIP1, SOD1, HSP 90AA1, HSP90AB1.	(Zininga et al., 2018)

Table 31. Evaluation of genes identified by gene ontology analysis in B4 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours).

The targets identified from the gene ontology analysis in B4 CD8^{Lo} were investigated further to assign the potential functional roles in B cells. The fold change (log2), function and supporting evidence are outlined.

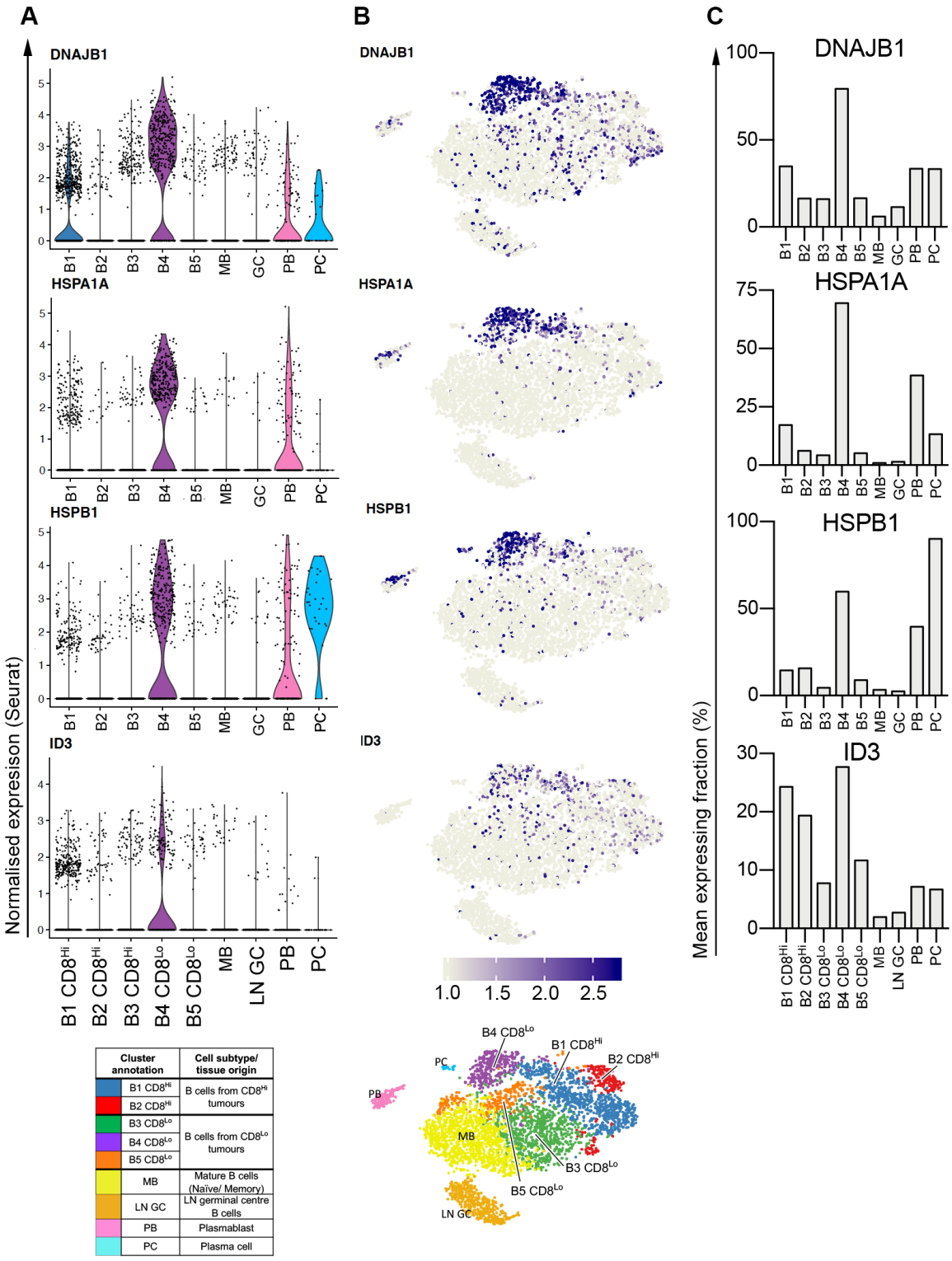


Figure 67. Expression of differentially expressed B cell features identified in B cells from B4 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours).

(A) Violin plots of Seurat normalised expression for the genes DNAJB1, HSPA1A, HSPB1 and ID3. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. Expression of heat shock proteins dominated the differentially expressed genes in B4 CD8^{Lo}, they are linked to cellular and endoplasmic reticulum stress.

B5 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours)

Gene	Av. FC	Summary	Ref.
CRIP1	1.48	CRIP1 (Cysteine-rich intestinal protein 1) is a zinc finger protein that has been used as a biomarker for cancer, it promotes cell migration and invasion, its expression also increases Beta catenin.	(Ludyga et al., 2013)
ITGB1	1.03	ITGB1 Integrin beta 1 supports the interactions between GC B cells and FDC cells for efficient GC B cell responses .	(Wang et al., 2014)
S100	1.02	S100 are calcium Ca2+ binding proteins that regulate intracellular Ca2+ levels (S100A10, S100A11 and S100A4). In B cells they function to control the influx of Ca2+ after activation via the BCR. This shapes B cell development through tolerance mechanisms that relate to excessive Ca2+ influx that results after chronic stimulation, ultimately leading to apoptosis.	(Baba and Kurosaki, 2016) (Hemon et al., 2017)
LGALS1	0.67	LGALS1 (Galectin 1) is the partner to TIM1 (HAVCR1) and has been shown to be part of regulatory B cell subtype with the ability to regulate and induce apoptosis in T cells. It functions in B cells to regulate homeostasis and promote cell survival in plasma cells.	(Zuniga et al., 2001, Alhabbab et al., 2018) (Anginot et al., 2013)

Table 32. Evaluation of genes identified by gene ontology analysis of B cells in B5 CD8^{Lo} cluster (B cells from CD8^{Lo} tumours).

The targets identified from the gene ontology analysis in B5 CD8^{Lo} were investigated further to assign the potential functional roles in B cells. The gene fold change (log2), function and supporting evidence are outlined in the table.

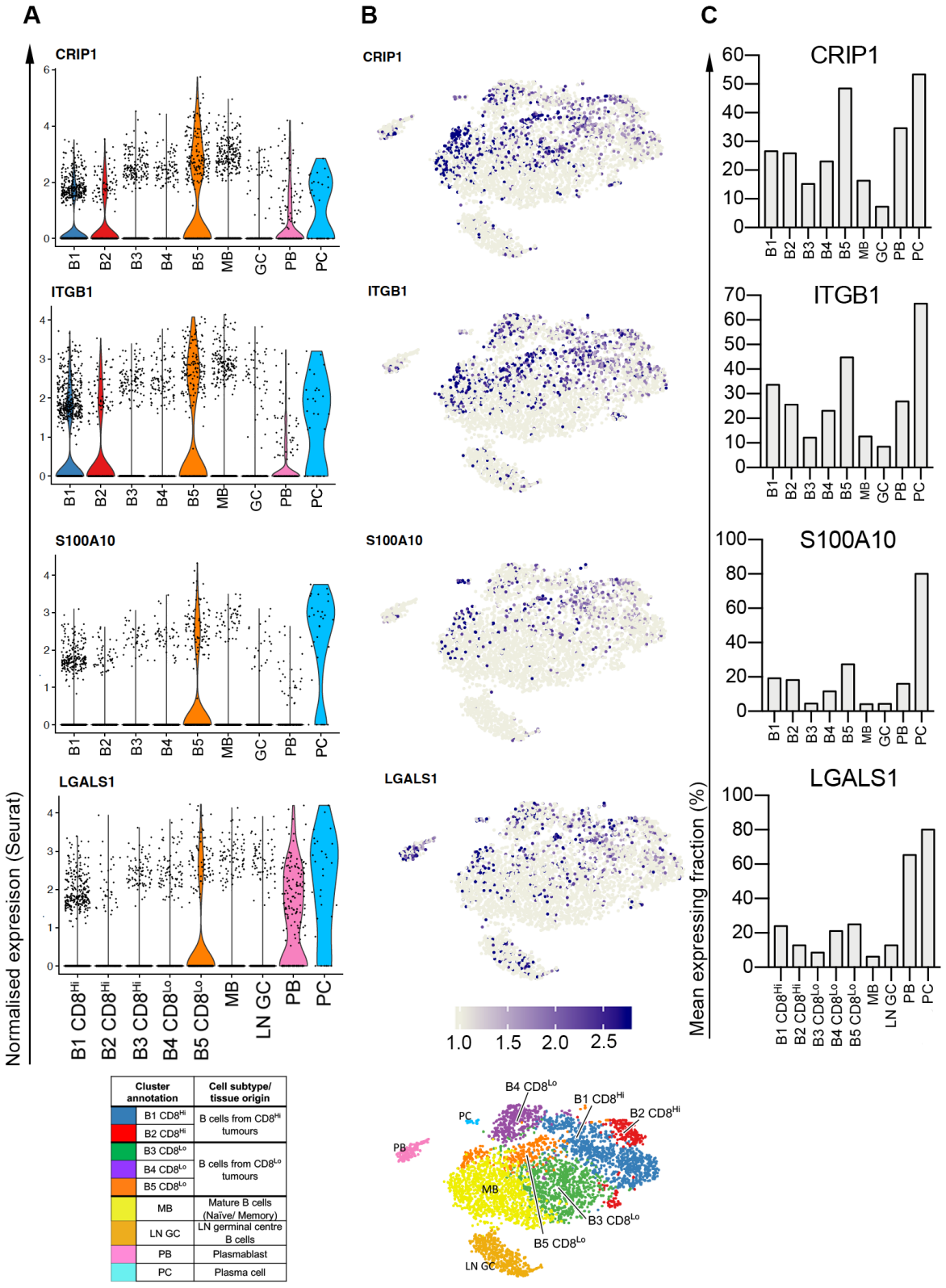


Figure 68. Expression of differentially expressed B cell features identified in cluster B5 CD8^{Lo} (B cells from CD8^{Lo} tumours).

(A) Violin plots of Seurat normalised expression for the genes CRIP1, ITGB1, S100A10 and LGALS1. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. Ca^{2+} channel expression suggests cell activation via the BCR and interactions with follicular dendritic cells via ITGB1. LGALS1 represent an interesting target that regulates B cell activation but also has the potential to negatively influence T cells.

4.3.3.9 Characteristics of B cells in the remaining clusters (LN GC, MB, PB and PC)

The remaining clusters that did not show a specific enrichment in either $CD8^{Hi}$ or $CD8^{Lo}$ tumours were the lymph node germinal centre B cells (LN GC), mature/ naïve B cells (MB), plasmablasts (PB) and plasma cells (PC). As previously outlined, differential gene testing was performed (MAST) followed by gene ontology analysis for clusters LN GC, MB, PB and PC. GO terms were subsequently reduced using REVIGO (Appendix D.3), where highly overlapping GO terms are removed to leave the core biological processes, any B cell related terms were also retained (summarised Appendix D.3.1 to D.3.3).

The LN GC cluster (Lymph node germinal centre B cells) were almost entirely made up of B cells derived from the lymph node tissue. The GO analysis of LN GC cluster displayed features of B cell activation, regulation of type 2 immune responses, BCR signalling pathway and positive regulation of B cell proliferation (Appendix D.3.1). Lymph node-derived B cells in the LN GC cluster also displayed a distinct germinal centre B cell expression profile; key to this was expression of BACH2 and BCL7A (Figure 69 A-C). The average expression of BCL6 was also highest in the LN GC cluster, although it was expressed at a low level overall and in a small fraction of the cells (Figure 69 A-C). The key genes identified in LN GC and supporting literature are shown in Table 33.

Mature/ naïve B cells (MB), as previously described (section 4.4.3.1), were defined as either quiescent memory or follicular B cells based on expression of CD27, IGHD and also evidence of isotype switching (IGHA and IGHG expression). The MB cluster displayed the fewest differential genes, likely a result of shared gene features across the remaining cells. Plasmablasts were identified in the PB cluster by the expression of key transcription factors (XBP1), activation markers (CD27), the cell markers SDC1 (CD138), CD38 and high levels of IGH expression (IGHG, IGHA and IGHM). The gene ontology identified response to unfolded protein, Ig-mediated immune responses and response to interferon (Appendix D.3.2).

The B cells derived from the PC cluster were few in number (n=30 cells) and represented the cluster with the most differential genes relative to the other cells (n=1037), the key identifying features from the gene ontology analysis were cell adhesion, localisation of cell and tissue migration (Appendix D.3.3). The cell type allocation identified this cluster as most likely plasma cells using SDC1 expression as a key marker along with tissue homing features.

Lymph node germinal centre B cells (LN GC)

Gene	Av. FC	Summary	Ref.
BACH2	1.59	BACH2 is integral to GC B cell formation by repressing the formation of Plasma cells and inhibiting the transcription factor BLIMP1 (required from PC formation). It also function as an essential part of BCR mediated cell proliferation.	(Muto et al., 2010) (Miura et al., 2018)
BCL7A	1.99	BCL7A (BAF Chromatin Remodelling Complex Subunit) has been termed a tumour suppressor gene and is linked to several lymphomas, it is also co-expressed with AID(activation induced cytidine deaminase) playing a role in isotype class switching.	(Yu et al., 2017)
CD22	1.14	CD22 functions as an inhibitory receptor reducing Ca2+ and limiting BCR signalling.	(Muller and Nitschke, 2014)
RGS13	1.83	RGS13 (Regulator Of G Protein Signalling 13) function to regulate the Germinal centre B cell response and size by modulating chemokine induced cell migration.	(Shi et al., 2002)
IL4R	0.95	Stimulation via IL4R and CD40 leads to effective BCR signalling and increased TACI and BCMA associated cell survival , IL4 is also linked to a TH2 response.	(Ferrer et al., 2014)

Table 33. Evaluation of genes identified by gene ontology analysis in the LN germinal centre B cells (LN GC) cluster.

The targets identified from the gene ontology analysis in LN GC were investigated further to assign the potential functional roles in B cells. The fold change (log2), function and supporting evidence are outlined.

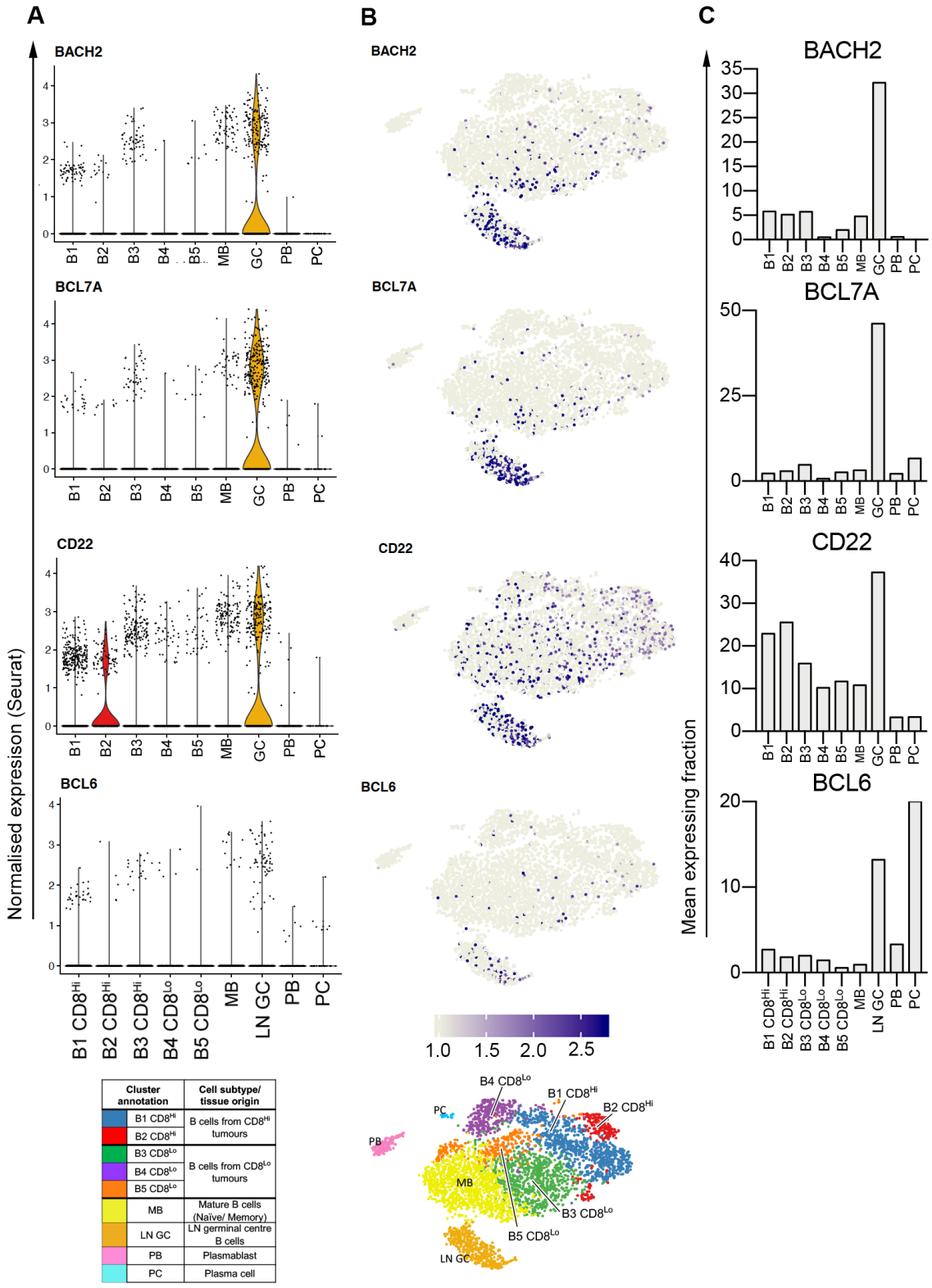


Figure 69. Expression of differentially expressed B cell features identified in the Lymph node germinal centre B cell (LN GC) cluster.

(A) Violin plots of Seurat normalised expression for the genes BACH2, BCL7A, CD22 and BCL6. (B) Seurat normalised expression overlaid across the tSNE plots and (C) mean expressing fraction (%) for the same genes. The B cells in LN GC were predominantly from the lymph node and display a germinal centre expression pattern with elevated BACH2 and BLC6.

4.3.4 Discussion

The role that B cells play in anti-tumour immunity consists of paradoxical findings that are yet to be resolved. On the one hand, they are reported to be immunosuppressive in the form of B_{reg} cells (Shalapour et al., 2015, Shalapour et al., 2017), where they have a negative impact on CD8+ T cell tumour immunity. On the other hand, a positive link to survival has been associated with high numbers of B cells and TLS in tumours, where B cells are also co-localised with CD8+ cells (Dieu-Nosjean et al., 2016, Germain et al., 2014).

Evaluation of B cells at a single cell level in NSCLC and non-involved tissues revealed distinct clusters that were linked to the density of the CD8+ T cell infiltrate in the originating tumour ($CD8^{Hi}$ and $CD8^{Lo}$ tumours). The present study aimed to investigate B cells from tumours that were $CD8^{Hi}$ and $CD8^{Lo}$, with the CD8+ T cell density being a known prognostic indicator of good outcome (Galon et al., 2006, Ward et al., 2014b). The identified clusters link to known B cell subsets, but also highlight differences in the overall function and activation profile of B cells that reflect the tissue from which they were derived ($CD8^{Hi}$ and $CD8^{Lo}$ tumours, non-involved lung (NIL) and Lymph node (LN). The B cell clusters (n=9) were categorised as enriched in $CD8^{Hi}$ tumours (clusters B1 and B2), $CD8^{Lo}$ tumours (clusters B3, B4 and B5) or annotated as B cell subsets (MB=mature/ naïve; LN GC=lymph node germinal centre; PB= plasmablast and PC= plasma cell).

A large proportion (70%) of B cells in NSCLC have previously been identified as having a memory/ antigen experienced phenotype (IGHD-CD38+/-) (Germain et al., 2014). The Single Cell data presented here was also enriched for cells that fit the profile of antigen-experienced activated/ memory cells. These were identified in clusters B1-B5, where differences associated with the CD8 density in the originating tumour emerged; these cells formed the focal point of the analysis.

B cells from $CD8^{Hi}$ tumours were enriched in clusters B1 (Blue) and B2 (Red), cell subtypes in B1 and B2 highlighted features of naïve B cells (IGHD+IGHM+), activated B cells (CD27) and evidence of isotype switching (IGHG and IGHA). Differentially expressed genes associated with B1 and to a lesser extent B2, were linked to an activation profile via BCR signalling and cellular migration (CXCR4, CD69, GPR183, CD48) (Weber, 2018, Gatto et al., 2009, Hoffmann et al., 1998, Yuan et al., 2013,(Castigli et al., 2005, Smulski et al., 2017, Weber, 2018, Gatto et al., 2009, Hoffmann et al., 1998, Yuan et al., 2013).

Both clusters B1 and B2 displayed the highest average expression of TACI (Transmembrane activator and Calcium-modulator and cytophilin ligand interactor - TNFRSF13B), a marker of switched memory cells, and an indicator of reactivated memory B cells within the tumour (Castigli et al., 2005, Smulski et al., 2017). TACI also plays a key role in the germinal centre during

differentiation into plasmablasts. It is a key part of the B cell and T cell interface in the light zone, its loss results in reduced plasmablast formation (Zhang et al., 2018). The ligation of TACI by APRIL has been shown to inhibit B cell proliferation and lead to plasma cell differentiation (Mackay and Schneider, 2008). The signature emerging from B1 was that of GC B cells and formation of TLS identified by expression of IRF8 and LTB (lymphotoxin β) (Wang et al., 2019, Shen and Fillatreau, 2015, Burbage and Keppler, 2018). All $CD8^{Hi}$ tumours displayed significantly more TLS within the tumour and at the tumour edge compared to the $CD8^{Lo}$ tumours. Other highly expressed features in B1 were MHC-I and -II genes; when coupled with features of T and B cell interactions (CD37, CD48) this highlights that antigen presentation to CD8+ and CD4+ T cells was possible and likely on-going. The role of antigen presentation by TIBs in lung cancer has been shown to have important effects on the functionality of CD4+ TIL. The presence of activated B cells leads to an effector T- cell response (IFNγ+), with unresponsive B cells (exhausted) leading to formation of Tregs (FOXP3+) (Bruno et al., 2017). The expression of MHC -I and -II genes was also elevated in B cells from $CD8^{Hi}$ tumours relative to the $CD8^{Lo}$ (Katikaneni and Jin, 2019, Yuseff et al., 2013).

A striking gene signature in cluster B2 (enriched in $CD8^{Hi}$ tumours) was that of the interferon stimulated genes (ISGs) that result from a response to IFN binding its cognate cell surface receptors. This leads to signalling via JAK:STAT1 and a wide range of ISG expression (Schneider et al., 2014). The average expression of IRF7 (interferon regulatory Factor 7) was also elevated; IRF7 is the gene regulator of Interferon alpha and beta (Honda et al., 2005). Conversely the expression of interferon genes linked to IFNγ signalling also exists (IFIL4, MX1/2 and ISG15)(Schneider et al., 2014).

Taken together, this highlights that a heterogenous stimulus is taking place leading to the expression of a wide range of interferon stimulated genes in the B cells in cluster B2 $CD8^{Hi}$. This indicates that the B cells in $CD8^{Hi}$ tumours are co-localised with antigen-specific CD4+ and CD8+ T cells and an on-going immune response, with the subsequent release of Interferon and a consequent ISG signature in the B cells. Some other key indicators of the interaction of B cells with T cells in the B2 cluster were the elevated expression of CD40, CCR7 and TNFRSF14 (herpesvirus entry mediator; HVEM) (Pereira et al., 2010, Elgueta et al., 2009, Domeier et al., 2016). Negative regulators were also identified that function to restrain the B cell response after the initial activation phase (FCGR2B, FCRL5, TNFSF10, TRAIL), SOCS3 (Guerreiro-Cacais et al., 2010), (Jhou et al., 2017), (Wilson et al., 2012, Jones et al., 2011). This feedback mechanism is also important for the selection of high affinity B cell clones in the GC, where only high affinity antibody clones are selected, this ultimately leads to a reduction in the GC reaction (Zhang et al., 2013). The formation of TLS in tumours has been shown to confer a favourable prognosis, the transcriptomic features here are indicative of TLS formation and an on-going interaction between B cells and CD4+ T cells providing co-stimulation. Histologically the $CD8^{Hi}$ tumours displayed

significantly (p=0.0024) higher numbers of B cells and dense aggregates of B cells (TLS) when compared to the $CD8^{Lo}$ tumours. This is supportive of a different B cell response and activation profile, with features of an on-going GC reaction and the co-existence of antigen specific T and B cells in the tumour. The extensive interferon signature that was observed in B2 (enriched in $CD8^{Hi}$ tumours) may be a reflection of a protective immune response to the tumour (Germain et al., 2014, Kroeger et al., 2016). From the cluster comparisons, LGALS9 emerged as differentially expressed in B2 (expressed in 16% of cells). LGALS9 is the ligand to TIM3 (HAVCR2) and is a potent regulator of CD8+ T cell responses, especially in highly active and functional CD8+ CTLs that express TIM3 (Zhu et al., 2005, Giovannone et al., 2018, Cao et al., 2018, Clarke et al., 2019). Although LGALS9 is only expressed in a small fraction of the cells it raises the question of whether B cells expressing LGALS9 have the potential to directly inhibit T-cells via TIM3. The cluster analysis also identified high average expression of a proteoglycan SRGN (Serglycin) in B2 (also in PB and to a lesser extent in B1). This forms molecular complexes with granzymes and perforin within the cell, it also exists in an extracellular secreted form. It has been implicated in promoting epithelial mesenchymal transition as well as prompting cancer cell invasion and metastasis, with higher expression linked to a worse prognosis in breast and colorectal cancer (Kolset and Pejler, 2011, Korpetinou et al., 2014, Xu et al., 2018, Zhang et al., 2017, Li et al., 2011). It is likely that even in tumours with a good prognosis ($CD8^{Hi}$) and high number of favourable TLS, B cells are present that exhibit features of direct negative regulation on T cells. Modulating/ blocking these features in an already good prognostic patient group may allow the antitumor response to be even more potent, LGALS9 and SRGN represented interesting targets that could be used for potential immunomodulation. The B cells that were identified in the $CD8^{Hi}$ tumours present an interesting group of cells that are playing a key role in the adaptive immune response. Displaying features of antigen presentation and interacting with other cells of the immune system, in particular expression of co-stimulatory molecules. Further work in this area is required to elucidate the exact nature of the positive and negative features expressed on B cells, and how they may be manipulated to improve antitumor responses.

The $CD8^{Lo}$ tumours are known to confer a poor prognosis (Galon et al., 2006, Ward et al., 2014b). The B cells from these tumours displayed a marked difference in clustering relative to the $CD8^{Hi}$ tumours, and were enriched in B3 (Green), B4 (Purple) and B5 (Orange). Naïve B cells (IGHD+) did not form a discrete B cell cluster in this cohort, IGHD was differentially expressed relative to the remaining cells in B3 (Green). However, expression of CXCR4 and CD69 in the same cluster also indicated a process of activation (Gutzeit et al., 2018, Quach et al., 2011). FOXP1 was differentially expressed in B3 cluster (also identified in MB cluster) and is found throughout B cell development, especially in mature naïve B cells (Patzelt et al., 2018, Hu et al., 2006). Taken together these two

pieces of evidence, along with lower CD27, mean they are likely cells undergoing activation, but also contain a higher proportion of naïve cells. B cells in cluster B3 were comprised mostly of cells from $CD8^{Lo}$ tumours and Lymph node, whereas the B cells from $CD8^{Hi}$ tumours were mainly absent. The B cells forming cluster B5 (Orange) were derived from LN, NIL and $CD8^{Lo}$ TIBs (least abundant in TIBS from $CD8^{Hi}$ tumours). The curated cell type allocation showed isotype switching and expression of IRF8 and SPIB, indicating activation and memory cell formation (Wang et al., 2019). A number of Calcium (Ca^{2+}) binding proteins that are associated with BCR activation and the regulation of Ca^{2+} influx were also differentially expressed in cluster B5 (Baba and Kurosaki, 2016, Hemon et al., 2017). ITGB1 (Integrin beta 1) was also elevated and plays a role in the interaction between GC B cells and FDC cells at the immunological synapse (Wang et al., 2014). An interesting target that emerged from the differential gene analysis was LGALS1 (Galectin 1), which binds β-galactosides. It is part of normal B cell homeostasis and is integral to BCR signaling and activation (Tsai et al., 2014). The expression of LGALS1 in B cells can induce a regulatory function, where it is required for IL10 production (Alhabbab et al., 2018). Immunomodulatory effects of LGALS1 have also been reported, but have not been directly associated with tumour infiltrating B cells. These effects include the induction of T-cell apoptosis and a reduction in cytokine production, where TCR signaling and IL2 production are modulated by LGALS1 (Zuniga et al., 2001) (Anginot et al., 2013, Perillo et al., 1995, Chung et al., 2000). The immune-inhibitory action of LGALS1 is diverse, affecting expansion of Tregs, T cell apoptosis, TCR activation, macrophage deactivation and a shift towards Th2 cytokine production (Cedeno-Laurent et al., 2012, Mendez-Huergo et al., 2017). Small molecule inhibitors of LGALS1 are being used in phase 1 clinical trials of OTX008 (NCT01724320), this opens the possibility to block its negative effects on T cells (Sundblad et al., 2017, Zheng et al., 2019, Paz et al., 2018). LGALS1 represents a novel marker identified in B cells that could be used to modulate the immune response alongside other immunomodulatory therapies (anit-PD1 and anti-CTLA4).

TIBs in $CD8^{Lo}$ tumours were also enriched in B4 (Purple) where a number of heat shock proteins (HSPs) were expressed. These function to facilitate the correct folding of proteins, they are also induced by cellular stress (Zininga et al., 2018). At an immunological level, HSPs play a role in both pro- and anti-inflammatory responses. HSP90 is involved in intra- and extra-cellular antigen presentation via MHC-I and -II as well as modulating the IFN response. HSP70 can directly regulate immune responses via IFN-γ and TNF-α production, as well as downregulating CD86 and MHC class II (Milani et al., 2002). The unfolded protein response (UPR) is a normal process of cells that are undergoing endoplasmic reticulum (ER) stress, and a key feature of cells preparing to become antibody secreting cells and is driven by the expression of XBP1 (highly expressed in plasmablasts (PB) and plasma cells (PC), which is upregulated upon ER stress. It should also be stated that XBP1 is expressed in cells that are not under ER stress, and is a requirement for effective differentiation

into plasmablasts via BCR signaling (Hetz, 2012, Hu et al., 2009). Evidence indicated isotype switched B cells in B5 (Purple) were also present, with IGHA and IGHG expression along with the activation marker CD27. This and the expression of cellular response to ER stress would indicate that these cells are undergoing differentiation into antibody secreting cells. Alternatively, the HSPs and Unfolded Protein response may be a combination of preparing to secrete antibody, and a cellular response to environmental stress. Cellular stress can be induced by conditions in the tissue microenvironment such as hypoxia and cellular damage, which are all hallmarks of cancer (Hanahan and Weinberg, 2011). The $CD8^{Lo}$ tumours present a picture of B cells with less activation, reduced interaction with other immune cells and display far fewer TLS. This fits with the tumours likely exhibiting a low level of immunogenicity (fewer antigens), leading to lower overall immune infiltrates.

Known B cell subsets (PC, PB, MB and LN GC) were identified using B cell lineage markers. The plasmablasts (PB) and plasma cells (PC) were identified by key transcriptional regulators (MZB1, XBP1, PRMD1), cell markers SDC1 (CD138) (Andreani et al., 2018) (Basso and Dalla-Favera, 2015, Nutt et al., 2015). The plasmablasts (PB) displayed the highest expression of IGHM, IGHA and IGHG genes linked to antibody secreting cells. The B cells in this cluster were derived evenly from all of the tissue types, suggestive of B cell activation in the $CD8^{Hi}$ and $CD8^{Lo}$ tumour, non-involved lung (NIL) and Lymph node (LN) B cells. The plasmablasts (PB cluster), especially the IGHA expressing cells, were those that have previously been reported as potentially immunosuppressive. The data here was unable to resolve features linked to IL10 expression or PDL1 expression that have been reported in the literature (Shalapour et al., 2015, Shalapour et al., 2017). Another B cell type not identified were the GZMB+ B cells that can display both immunosuppressive and cytotoxic effects in tumours (Lindner et al., 2013, Hagn et al., 2012). This is likely a limitation of the current single cell technology where the expression of IGH RNA transcripts dominate the available RNA-Seq capacity, this ultimately leads to an inability to capture transcripts expressed at lower levels in this cell type. Plasma cells (PC cluster) were few in number and displayed features of cellular adhesion and localisation, expressing SDC1, EPCAM and integrins (Tarte et al., 2003). The B cells in cluster MB (mature/ naïve B cells) displayed very few distinguishing features relative to the other clusters; this likely indicates that they share features with the remaining B cells in the comparison. Assessment of transcription factors identified SPIB, PAX5 and EBF1 in MB B cells, in addition to this there was also evidence of isotype switched IGHA, IGHG and non-switched IGHM B cells (Basso and Dalla-Favera, 2015, Nutt et al., 2015). Taken together this indicates that they likely represent quiescent memory/ follicular B cells that are enriched in the lymph node, non-involved lung tissue and to some extent, the B cells in $CD8^{Lo}$ tumours. Interestingly, the MB cluster was all but absent from the TIBs in $CD8^{Hi}$ tumours. A fraction of B cells isolated from the lymph node formed a unique cluster (LN GC) with the

hallmarks of a classical germinal centre reaction, expressing BACH2, BCL6, IL4R and CD40. This would argue that the GC signature observed in the remaining clusters is not identical to that of a GC B cells that resides in a secondary lymph node (Muto et al., 2010, Basso and Dalla-Favera, 2010, Miura et al., 2018, Ferrer et al., 2014).

TIBs derived from $CD8^{Hi}$ and $CD8^{Lo}$ tumours did not cluster based on a specific B cell lineage feature. They did however, cluster separately with an enrichment of cells from $CD8^{Hi}$ tumours in B1 (Blue) and B2 (Red), with $CD8^{Lo}$ tumours enriched for B cells in B3 (Green), B4 (Purple) and B5 (Orange). Previous single cell studies in NSCLC have highlighted the B cells as plasmablast/plasma cells, activated/memory and BALT (bronchus associated lymphoid) B cells (Table 25). The data presented here has identified the same cell types but also identifies differences in the B cells based on whether they were residing in a $CD8^{Hi}$ or $CD8^{Lo}$ tumour. This suggests that within the previously reported activated/memory and BALT clusters, B cells were also present that differ by function and response to environmental cues, in particular CD8 density.
The activation and differentiation of the cells in $CD8^{Hi}$ and $CD8^{Lo}$ tumours does not appear to follow the same path, in $CD8^{Hi}$ tumours the cells co-express markers in clusters B1 and B2, whereas in $CD8^{Lo}$ tumours co-expression occurs in B3, B4 and B5 clusters ultimately leading to activated B cells. One way of investigating this further will be to perform unsupervised pseudotime cell state trajectory analysis (Trapnell et al., 2014), this allows a cell's state to be plotted on pseudotime based on its transcriptional profile. This is currently underway and will aim to resolve the identified differential activation and differentiation profile of the B cells in different prognostic groups ($CD8^{Hi}$ and $CD8^{Lo}$ tumours). Where it is hypothesised that the cells in $CD8^{Hi}$ and $CD8^{Lo}$ tumours are activated under different conditions, this leads to a distinct transcriptional profile that is reflective of the co-stimulatory and cytokine conditions in the tissue microenvironment. The cells would be predicated to diverge during activation and differentiation, this then leads to different cell population characteristics.

4.3.5 Conclusion

The data presented here has revealed that B cells isolated from tumour and control tissues generate distinct clusters. These were biologically relevant and confirmed what had been reported in the literature previously relating to B cell subtypes. B cells from two prognostically different groups ($CD8^{Hi}$ and $CD8^{Lo}$ tumours) represented a unique analysis feature that allowed specific B cell characteristics in these settings to be investigated. Analysis of the transcriptomic features of B cells enriched in $CD8^{Hi}$ tumours showed a distinct activation and differentiation profile that was not identified in the B cells from $CD8^{Lo}$ tumours, lymph node or non-involved lung

B cells. In addition, potential targets of immunomodulation were also revealed that were expressed in the different prognostic groups (SRGN, LGALS9 and LGALS1). Histological assessment of these tumours identified a stark difference in the location and organisation of B cells in tumours that were enriched for CD8+ T cells compared to tumours with sparser infiltrates. The data has revealed that the B cells within, and outside these aggregates have different transcriptional characteristics depending on the overall density of the CD8+ T cell infiltrate. This has potential applications for ways to improve immunotherapy by modulating the negative effect of B cells during the adaptive immune response in different prognostic groups ($CD8^{Hi}$ and $CD8^{Lo}$ tumours).

Chapter 5: Concluding statement

The work outlined in this thesis investigates tumour infiltrating lymphocytes (TIL) in different biological settings. In particular it addresses the differences that exist in these immune cells, when they reside in TIL^{Hi} and TIL^{Low} tumours. The role of immune cells in tumour biology has gained an immense amount of scientific attention, this is due to patient outcomes being linked to immune cell density. Building on this, the blockade of immune regulators (e.g. PD1) has transformed cancer treatment and capitalised on the existing immune response to tumours. However, despite these advances response rates to immunotherapies stand at 20-30%, the reasons for which are still being elucidated. The work here evaluates the tumour as a whole and the stability of the immune signatures, the stratification of patients, and how the immune response may differ in virally induced tumours (HPV(+) and (-)). Following this the immune cells were investigated at an individual population level, evaluating CD8 T-cells and B cells in TIL^{Hi} and TIL^{Low} tumours with the aim of identifying unique features that may improve our understanding of tumour immunology.

The immunological stability of a tumour, both spatially and temporally, has important implications for monitoring of immunomodulatory therapy in clinical trials, as well as during routine treatment regimens. Multiple tumour (HNSCC) replicates from the same patient, separated spatially and temporally, were taken between diagnostic biopsy and surgical resection. The analysis revealed a high level of consistency at the global gene expression level across tumour replicates both spatially and temporally. Further interrogation of immune gene expression and CD8 cell numbers by immunohistochemistry also showed stable immune cell features across spatially and temporally distinct tumour replicates. This data has increased our confidence in the stability of small biopsies taken for “window of opportunity studies”, where immunological readouts are a key focus (Wood et al., 2017). This key piece of work was used as a basis for the evaluation of the PI3Kδ inhibitor of in HNSCC (AMG319 study in HPV positive and negative HNSCC: NCT02540928) (Ali et al., 2014). Knowing that the immune signal was stable between replicates across time and space, enabled confidence in assessing change caused by the therapeutic intervention. It also means that tumour biopsies taken for patient stratification and disease monitoring, are likely to be representative of the tumour as a whole.

The survival difference that exists between HPV(+) and (-) tumours is well documented in the literature. This has been categorised further with the use of TIL status, where high/ moderate TIL confer a substantial survival benefit. However, differences still exist between HPV(+) and (-) TIL rich tumours. In order to better understand the survival difference between them, the analysis

focused on the comparison of HPV(+) and (-) immune rich tumours (TIL high/ Moderate). Differential gene expression analysis was used to compare the two groups. This led to the finding that TIL rich HPV(+) tumours were significantly more enriched for immune genes than HPV(-) tumours. As a result, the immune cell markers detected by IHC were enumerated and correlated to the RNA-Seq data, confirming the finding. At this point the numerical differences appeared to be the key driver in the survival difference, however, it was still possible that a specific immunological signature existed in the HPV(+) relative to the (-). The analysis was then repeated with an Immune cell correction; this used the immune gene expression as a confounding variable in the differential gene analysis. The analysis yielded vastly different results, with tumour-driving gene ontologies replacing the significant immunological signature identified in the HPV(+) cases, highlighting the loss of the differential immune gene expression signature. This would indicate that at the bulk tumour level, the immune characteristics in TIL rich HPV(+) and (-) tumours are comparable, and that the presence of a persistent immunogenic TAA leads to a higher density of immune cells. However, during the analysis, a signature of B cell-related genes remained that was a distinguishing feature between TIL rich HPV(+) and (-)tumours (Wood et al., 2016b). The B cells that are enriched in HPV(+) tumours are being assessed in more detail in a separate project using flow cytometry, IHC, ELISA and ELISPOT analyses.

Analysing the tumour as a whole enabled the global immune context to be investigated, it also highlighted that immune density could be assigned based on a very limited number of immune genes. This becomes important due to the potential of applying rapid gene expression assays, this the enables the grouping of patients into categories that may/ or may not respond to treatments. The ability to deliver this in clinical trial and/ or diagnostics settings, where drug targets could be tailored to specific tumours is highly attractive and could improve immunotherapy response rates.

Following on from the findings of global immunological similarities at the whole tumour level in HPV(+) and (-) tumours, a greater level of detail and insight into the immune cell populations and how they differ between patient groups would be achieved by transcriptomic analysis of specific immune cells. The aim was to sort CD4, CD8, B cell and Macrophages from HNSCC, NSCLC and non-involved background lung and carry out high resolution transcriptomic analysis of the cells using proprietary RNA-Seq techniques developed in house in La Jolla institute for allergy and disease. The experimental procedures have been optimised for the dissociation of tumours, subsequent flow cytometric analysis and cell sorting using a simple immunophenotyping panel for isolation of the purified CD4, CD8, B cell and Macrophage populations for RNA-Seq analysis. The flow sorting procedure has also been applied to the collection of immune cell populations in tertiary assays on the AMG-319 and HARE40 clinical trials looking at immunomodulatory

treatments in HNSCC. The immune cell percentages, as determined by flow cytometry, map well onto what has previously been reported in the literature (HNSCC and NSCLC).

Transcriptomic analysis of CD8+ T cells in HNSCC and NSCLC revealed a distinct tissue-resident memory cell features in tumours relative to non-involved lung. This was confirmed by flow cytometric analysis of NSCLC, non-involved lung tissue and blood. This identified that CD8+CD103+ T cells were significantly enriched in tumours relative to the control tissue, and that they also expressed negative regulators of the immune response (PD1). Following on from this the transcriptome of CD8+ T cell in $CD8^{Hi}$ and $CD8^{Lo}$ tumours was investigated by differential gene testing. The gene signature that emerged was one of activated cytolytic CD8+T_{RMs} in the $CD8^{Hi}$ tumours relative to the $CD8^{Lo}$. The positive association of CD8+CD103+ cells with superior cytolytic functionality meant that the density of CD103 in tumours would be predicted to confer a survival benefit in NSCLC. A retrospective cohort of NSCLC and HNSCC were assessed for survival in relation to the density of CD103+ and CD8+ cells. The CD103+ cell density displayed a significant improvement in survival in NSCLC and the same trend in HNSCC relative to the known good outcome marker CD8 (Ganesan, Clarke, Wood et al., 2017). The identification of numerous CD8+T_{RMs} in immune rich tumours is an important consideration for immunotherapy and disease monitoring. These cells are markers of a protective immune response, they are also released during the blockade of immune checkpoints (Clarke et al., 2019). In-depth analysis of CD8+T_{RMs} identified HACVR2/ TIM3 as a prominent feature, enabling the CD8+T_{RMs} to be grouped into highly active CD8+T_{RMs} TIM3+ (Clarke et al., 2019). This has potential uses in selecting tumour reactive cells for adoptive cell therapies, and also as a marker of immunotherapy response.

The transcriptomic data from CD8+ TILS isolated from HNSCC was used to compare CD8 T cells from HPV(+) and (-) cases. The data confirmed that there was not a significant association with a viral signature in CD8+T cells from HPV(+) tumours. This corroborated the findings in the bulk tumour analysis where the CD8+ T cell characteristics appeared equivalent between HPV(+) and (-) tumours. This indicated that it was the density and number of cells that conferred the substantial survival benefit in HPV(+) tumours relative to HPV(-), rather than a qualitative difference. The consistent features identified in CD8 cells between NSCLC, HNSCC HPV(+) and (-) tumours was a key finding (Cillo et al., 2020) also identified similarities in CD8 cells in HPV(+) and (-) HNSCC. It shows that the CD8 cells in different tumours are highly similar and that the observed functional differences are a reflection of the tumour itself. Where highly immunogenic tumours drive a large number of cytolytic CD8 cells into the tumour. Conversely, poorly immunogenic tumours have fewer/ less activated infiltrates. It also demonstrates that

immunotherapies targeting immune cells in one cancer type will likely be successful across many cancers, providing a dense immune infiltrate is present.

Analysis of B cells (single cell) isolated from $CD8^{Hi}$ and $CD8^{Lo}$ tumours showed that the transcriptomic features and clustering were heavily influenced by the tissue from which they originated. The transcriptional features indicate a different activation profile, with B cells in $CD8^{Hi}$ tumours exhibiting a response to interferon and co-stimulatory signals that are all indicative of an antigen driven immune reaction. Conversely the B cells isolated form $CD8^{Lo}$ tumours displayed a more naïve profile and activated features that were lacking co-stimulatory signals. The analysis so far has identified several unique regulatory features that could also be potential targets for immunomodulation. This work is on-going and requires further analysis to map the activation of the B cells using cell state trajectories. IHC and flow cytometry will also be required to confirm the presence of the unique B cell features in $CD8^{Hi}$ and $CD8^{Lo}$ tumours.
The importance of tumour infiltrating B cells and their role in supporting the anti-tumour response has recently moved to the forefront of tumour immunology. The data here shows that B cells exhibit features of cellular co-ordination and promotion of both CD4 and CD8 T-cells. Several studies have recently identified TLS and the cells that reside within them to be crucial for the response to immune check point blockade. These studies also highlight the Tfh cells that interreact with B cells as key cells (Helmink et al., 2020, Hollern et al., 2019, Cillo et al., 2020) . The balance of suppressive and activating B cells in the tumour and their manipulation, has the potential to further enhance the antitumor effects of the immune system.

A general overview of the findings are presented in Figure 70, here the tumours have been divided into $Immunogenic^{High}$ and $Immunogenic^{Low}$. An excellent example of these two tumour types are the HPV(+) and HPV(-) tumours, where highly abundant HPV antigens drive a prominent immune response, this also happens in other tumours that express tumour associated antigens. Many studies (Sahin et al., 2017, Ott et al., 2017) have identified the mutational (neoantigens) burden as a key factor in the overall survival, and also response to immunotherapies. These tumours ($Immunogenic^{High}$) have dense immune infiltrates consisting of cytolytic $CD8+T_{RMs}$, B cells that aggregate into lymphoid structures and interact with Tfh cells. This ultimately confers a survival benefit, it also points to the patient selection group that are most likely to benefit from immunotherapies. Tumours with lower immune infiltrates fit into immune excluded (see section 1.3.3) and the $Immunogenic^{Low}$ group, where the lack of abundant antigen leads to fewer expanding $CD8+T_{RMs}$, less B cells and less TLS. This patient group has poorer survival and are less likely to respond to immunotherapies. The immune cell functionality and overall density in tumours is shaped by the tumour itself ($Immunogenic^{High}$ and $Immunogenic^{Low}$ tumours), this

means that to improve the patient outlook during immunotherapy treatments, they need to be uniquely tailored to each patient and their tumour.

Figure 70. Overview of immune infiltrates in immune 'Hot' and 'cold' tumours.
A schematic representation tumours of with high numbers of immune infiltrate ('Immune Hot'), and tumours with sparse immune infiltrates ('Immune Cold'). (A) Tumours that have abundant antigens (e.g. HPV(+)/ high number of neoantigens) exhibit dense immune infiltrates. This consists of cytolytic CD8 T_{RM} that improve patients survival and B cells that form tertiary lymphoid structures with evidence of antigen presentation, class switching and T-cell help. (B) Tumours that are poorly immunogenic (HPV(-)/ low mutational burden) and/ or display stromal barrier features (not shown), have fewer immune infiltrates ('Immune Cold'). Patient survival is lower with less CD8 T_{RM} cells and less activated B cells, significantly fewer tertiary lymphoid structures and display less interaction with T-cells.

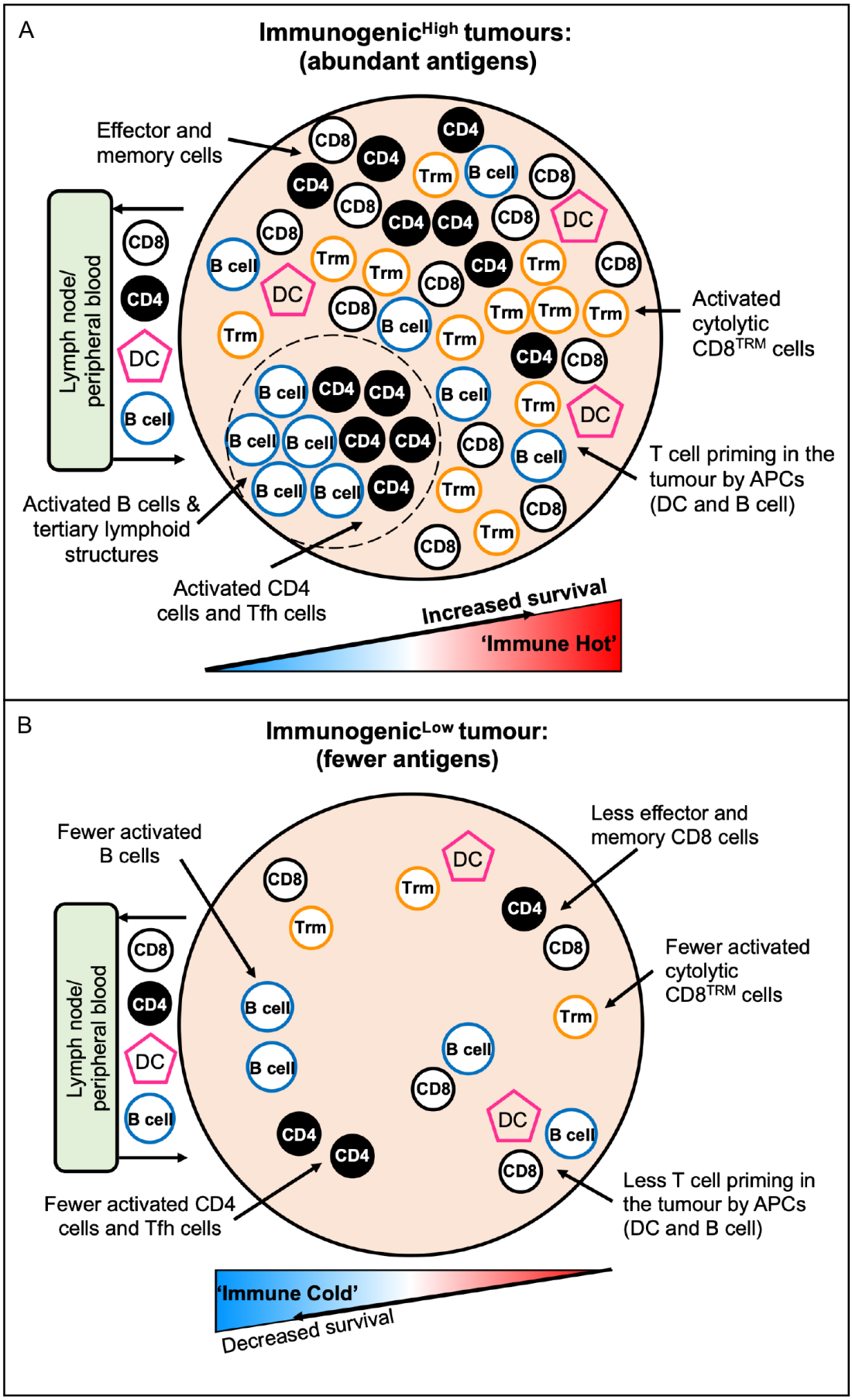


Figure 70. Overview of immune infiltrates in immune 'Hot' and 'cold' tumours.

Appendices

Appendix A

A.1 Patient demographics

A.1.1 Clinical cohort for the analysis of tumour replicates across time and space (data set TRTS)

Patient ID	Age (years)	No. of biopsies	Diagnostic biopsy	Surgical resection	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	HPV Status	No. of CD8+ cells (av. HPF)	Sample type
1	68	3	3	0	T3	N2b	M0	Non Smoker	Positive	6	Single timepoint
2	60	2	0	2	T4	N0	M0	N/A	Negative	10	Single timepoint
3	N/A	3	1	2	T3	N1	M0	N/A	Positive	24	Two timepoint
4	72	2	2	0	T2	N2c	M0	Ex-Smoker	Negative	16	Single timepoint
5	61	2	1	1	T1	N2b	M0	N/A	Negative	3	Two timepoints
6	76	2	1	1	T2	N0	M0	Non smoker	Negative	N/A	Two timepoints
7	50	2	1	1	T2	N1	M0	Non smoker	Positive	8	Two timepoints
8	64	2	0	2	T4	N2b	M0	Smoker	Negative	10	Single timepoint
9	48	2	1	1	T2	N2b	M0	Smoker	Positive	38	Two timepoints
10	63	2	1	1	T4	N2c	M0	Smoker	Positive	21	Two timepoints
11	61	3	0	3	T1	N0	M0	Smoker	Negative	N/A	Single timepoint
12	63	3	0	3	T4	N2b	M0	Smoker	Negative	22	Single timepoint
13	57	3	1	2	T1	N2b	M0	Smoker	Positive	N/A	Two timepoints
14	59	6	3	3	T4	N1	M0	Smoker	Negative	10	Two timepoints
15	47	3	3	0	T3	N2a	M0	Smoker	Positive	12	Single timepoint
16	62	2	0	2	T2	N1	M0	Smoker	Negative	28	Single timepoint

A.1.1.Clinical cohort for the analysis of tumour replicates across time and space (data set TRTS).

Samples collected from 16 patients between diagnostic biopsy and surgical resection (two timepoints) and also spatially by collection of multiple replicates at the same timepoint (one timepoint).

A.1.2 Clinical cohort analysing HPV(+) and HPV(-) TIL rich tumours (Dataset HPVN)

Patient ID	Age (years)	Gender	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	HPV Status	No. of CD8+ cells (av. per HPF)	No. of CD3+ cells (av. per HPF)	No.of CD20+ cells (av. per HPF)	TIL status	IO Biomark qPCR
Patient2	57	male	4	T3	N2	M0	non-smoker	POSITIVE	17.2	70.7	64	3	NA
Patient8	59	Male	3	T1	N1	M0	non-smoker	POSITIVE	NA	NA	NA	3	NA
Patient12	50	Male	2	T2	N0	M0	smoker	POSITIVE	32.3	116.8	328	3	NA
Patient14	58	Male	4	T3	N2a	M0	non-smoker	POSITIVE	21.2	83.1	69.8	3	NA
Patient16	75	Male	4	T2/3	N2b	M0	non-smoker	POSITIVE	NA	NA	NA	3	NA
Patient17	57	Male	3	T3	N0	M0	non-smoker	POSITIVE	NA	83.1	NA	3	NA
Patient19	66	Female	4	T4	N2c	M0	non-smoker	POSITIVE	41	89	50.1	3	NA
Patient20	47	male	4	T1	N2b	M0	non-smoker	POSITIVE	27.6	172	208.6	3	NA
Patient1	88	male	4	T1	N2a	M1	smoker	NEGATIVE	24.2	60.5	42	3	Yes
Patient3	70	Female	4	T4	N2c	M0	non-smoker	NEGATIVE	29.7	63.2	11.9	3	NA
Patient6	77	male	4	T4	N2c	M1	non-smoker	NEGATIVE	23.2	67.2	12.3	3	Yes
Patient31_a	84	male	1	T1	N0M0	M0	heavy smoke	NEGATIVE	11.9	67.1	8.9	3	Yes
Patient11	60	male	3	T2	N1M0	M0	non-smoker	NEGATIVE	35.9	100.5	18.1	3	NA
Patient23	58	male	4	T4	N0	M0	heavy smoke	NEGATIVE	20.5	113.9	70.1	3	NA
Patient7	47	Male	4	T3	N2	M0	non-smoker	POSITIVE	NA	NA	NA	2	NA
Patient18	53	Male	4	T2	N2b	M0	non-smoker	POSITIVE	33.3	132.7	172.9	2	Yes
Patient5	55	Male	4	T4	N2c	M0	smoker	NEGATIVE	19.6	41	23.8	2	Yes
Patient9	68	male	1	T1	N0	M0	non-smoker	NEGATIVE	24.9	41	13.3	2	NA
Patient13	52	Female	3	T3	N0	M0	heavy smoke	NEGATIVE	9.9	25.3	8.1	2	NA
Patient21	29	Female	1	T1	N0	M0	non-smoker	NEGATIVE	NA	NA	NA	2	NA
Patient25	50	male	4	T2	N2a	M0	smoker	NEGATIVE	15.5	51.5	18.8	2	NA
Patient34	63	Male	4	T2	N2c	M0	non-smoker	NEGATIVE	37.4	93.4	20.9	2	Yes
Patient39	75	Male	3	T3	N0	M0	smoker	NEGATIVE	18.2	32.9	18.3	2	NA
Patient37_A	57	Male	4	T4	N1M0	M0	non-smoker	NEGATIVE	NA	NA	NA	2	NA
Patient15	69	Male	4	T4	N2c	N/A	smoker	POSITIVE	8.8	11.7	3.1	1	NA
Patient4_C	56	male	4	T4	N1M0	M0	smoker	NEGATIVE	NA	NA	NA	1	NA
Patient10	55	Male	4	T2	N2c	N/A	heavy smoke	NEGATIVE	NA	NA	NA	1	NA
Patient22	82	Female	1	T1	N0	M0	unknown	NEGATIVE	7.5	13	3.3	1	Yes
Patient24_a	67	male	4	T4	N2b	N/A	heavy smoke	NEGATIVE	4.1	5.6	4.7	1	NA
Patient26_a	68	Female	3	T3	N0	M0	smoker	NEGATIVE	4.5	9.7	10.9	1	Yes
Patient30_a	76	male	3	T3	NxM0	M0	smoker	NEGATIVE	8.6	24.6	3.8	1	NA
Patient32	80	female	2	T2	N0	M0	non-smoker	NEGATIVE	11.2	12.6	1.9	1	Yes
Patient33	51	Male	4	T4	N2c	M0	heavy smoke	NEGATIVE	1.8	7.1	1.3	1	NA
Patient35	69	Female	4	T2	N2b	M0	non-smoker	NEGATIVE	7.8	12.5	1.1	1	NA
Patient36	66	Male	4	T4	N1	M0	smoker	NEGATIVE	8.5	13.5	0	1	Yes
Patient38	65	Male	1	T1b	N0	M0	non-smoker	NEGATIVE	NA	NA	NA	1	NA

A.1.2. Clinical cohort analysing HPV(+) and HPV(-) TIL rich tumours (Dataset HPVN)

Clinical features of the HPV(+) and HPV(-) sample cohort (n=39).

A.1.3 Clinical details of the Sorted B cells for qPCR

Patient ID	Age (years)	Gender	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	HPV Status
HN40	60	Male	3	3	1	0	Smoker	Positive
HN47	51	Male	4	2	3	0	Never	Positive
HN54	69	Male	3	2	1	0	Never	Positive
HN60	67	Male	4	4	2	0	Never	Positive
HN63	48	Male	4	2	2	0	Ex-smoker	Positive
HN76	47	Male	4	3	2	0	Smoker	Positive

A.1.2. Clinical detail of the Sorted B cells for qPCR

Clinical characteristics of the cases where B cells were isolated for evaluation with qPCR.

A.1.4 Clinical cohort of NSCLC cases with CD8 Cells sorted for transcriptomics analysis

Patient ID	Age (years)	Gender	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	Tumor histology	no. of CD8a+ cells (av. per HPF)	TIL status
NSCLC_01	87	M	IA	1A	0	0	Ex	adenocarcinoma	32.7	High
NSCLC_02	74	M	IIB	2B	1	0	Ex	squamous carcinoma	8.6	Intermediate
NSCLC_03	77	M	IA	2B	2	0	Ex	adenocarcinoma	28.2	High
NSCLC_04	67	M	IB	2A	0	0	Ex	squamous carcinoma	14.7	High
NSCLC_05	84	F	IIA	1B	1	0	Ex	adenocarcinoma	11.4	Intermediate
NSCLC_06	72	M	IA	1B	0	0	Ex	adenocarcinoma	15.6	High
NSCLC_07	74	M	IIB	3	0	0	Ex	adenocarcinoma	80.3	High
NSCLC_08	63	M	IB	2A	0	0	Ex	adenocarcinoma	21.2	High
NSCLC_09	83	M	IIA	2B	0	0	Ex	squamous carcinoma	11.4	Intermediate
NSCLC_10	64	M	IB	1A	0	0	Ex	adenocarcinoma	23.2	High
NSCLC_11	72	F	IIIA	4	0	0	Current	adenocarcinoma	9.9	Intermediate
NSCLC_12	72	F	IIA	2B	0	0	Never	adenocarcinoma	28.1	High
NSCLC_13	68	F	IIA	2A	1	0	Ex	adenocarcinoma	9.2	Intermediate
NSCLC_14	50	M	IB	2B	0	0	Current	adenocarcinoma	7.1	Low
NSCLC_15	74	M	IB	2A	0	0	Ex	adenocarcinoma	8.3	Intermediate
NSCLC_16	65	F	IA	1A	0	0	Ex	adenocarcinoma	3.0	Low
NSCLC_17	68	M	IIA	2B	0	0	Ex	squamous carcinoma	17.5	High
NSCLC_18	71	F	IIIA	3	1	0	Current	squamous carcinoma	15.0	High
NSCLC_19	68	F	IA	1A	0	0	Ex	adenocarcinoma	6.5	Low
NSCLC_20	72	F	IB	2A	0	0	Ex	adenocarcinoma	38.7	High
NSCLC_21	72	M	IV	1A	0	1B	Ex	adenocarcinoma	10.3	Intermediate
NSCLC_22	70	M	IIIA	3	2	0	Ex	adenocarcinoma	4.1	Low
NSCLC_23	51	F	IB	2A	0	0	Never	adenocarcinoma	9.6	Intermediate
NSCLC_24	77	F	IB	2A	0	0	Ex	adenocarcinoma	10.8	Intermediate
NSCLC_25	60	F	IA	1B	0	0	Ex	adenocarcinoma	10.7	Intermediate
NSCLC_26	77	F	IIA	2A	1	0	Ex	adenocarcinoma	6.3	Low
NSCLC_27	81	F	IIB	3	0	0	Ex	squamous carcinoma	10.8	Intermediate
NSCLC_28	69	F	IB	2A	0	0	Ex	adenocarcinoma	6.8	Low
NSCLC_29	73	M	IB	2A	0	0	Ex	adenocarcinoma	3.7	Low
NSCLC_30	81	F	IIIB	4	2	0	Never	adenocarcinoma	4.3	Low
NSCLC_31	76	M	IA	1B	0	0	Current	squamous carcinoma	2.7	Low
NSCLC_32	77	F	IIIA	2A	2	0	Never	adenocarcinoma	4.8	Low
NSCLC_33	67	M	IIB	3	0	0	Current	squamous carcinoma	4.4	Low
NSCLC_34	70	F	IA	1B	0	0	Ex	adenocarcinoma	12.6	Intermediate
NSCLC_35	66	M	IA	1A	0	0	Ex	adenocarcinoma	10.1	Intermediate
NSCLC_36	80	M	IB	2A	0	0	Ex	squamous carcinoma	18.6	High
NSCLC_37	81	M	IA	1A	0	0	Ex	squamous carcinoma	N/A	N/A
NSCLC_38	69	M	IB	1B	0	0	Ex	adenocarcinoma	N/A	N/A
NSCLC_39	75	M	IIIA	3	1	0	Current	squamous carcinoma	N/A	N/A
NSCLC_40	58	F	IA	1A	0	0	Current	adenocarcinoma	N/A	N/A
NSCLC_41	76	M	IB	2A	0	0	Ex	adenocarcinoma	N/A	N/A
NSCLC_42	74	M	IA	1A	0	0	Ex	adenocarcinoma	N/A	N/A
NSCLC_43	79	M	IIB	3	0	0	Ex	squamous carcinoma	N/A	N/A

A.1.4 Clinical cohort of NSCLC cases with CD8 Cells sorted for transcriptomics analysis.

Summary of clinical features from the NSCLC cohort where CD8+ T cells were isolated, these were then assessed by RNA-Seq.

A.1.5 Clinical cohort of HNSCC cases with CD8 Cells sorted for transcriptomics analysis.

Patient ID	Age (years)	Gender	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	HPV Status	no. of CD8+ cells (av. per HPF)	TIL status
HNSCC_01	82	M	III	2	1	0	Ex	Negative	25	Intermediate
HNSCC_02	55	F	IVA	4	1	0	Ex	Negative	12.5	Intermediate
HNSCC_03	94	F	IVA	3	0	0	N/A	Negative	3.1	Low
HNSCC_04	69	M	III	2	1	0	N/A	Positive	26.1	Intermediate
HNSCC_05	57	M	IVA	1	2B	0	Smoker	Negative	35.7	High
HNSCC_06	66	M	IVA	4	2B	0	Never	Positive	23.9	Intermediate
HNSCC_07	64	F	I	1	0	0	Ex	Negative	35.5	High
HNSCC_08	63	M	IVA	4	2C	0	Current	Negative	29.3	High
HNSCC_09	66	F	IVA	2	2B	0	N/A	Negative	28.5	High
HNSCC_10	86	F	IVA	4A	0	0	Never	Positive	24.1	Intermediate
HNSCC_11	70	M	IVA	4B	2B	0	N/A	Negative	10.2	Low
HNSCC_12	56	M	IVA	3	2B	0	Never	Negative	32	High
HNSCC_13	47	M	IVA	3	2A	0	Current	Positive	N/A	N/A
HNSCC_14	67	M	IVA	4A	2B	0	Never	Positive	24.8	Intermediate
HNSCC_15	74	M	III	2	1	0	N/A	Negative	11	Low
HNSCC_16	57	M	IVC	4	3	1	Current	Negative	37.2	High
HNSCC_17	60	F	IVA	4A	2B	0	Ex	Negative	1.6	Low
HNSCC_18	48	M	IVA	2	2A	0	Ex	Positive	32.5	High
HNSCC_19	60	M	III	3	1	0	Current	Positive	26.8	Intermediate
HNSCC_20	51	M	IVA	4	0	0	Ex	Positive	25.5	Intermediate
HNSCC_21	62	M	II	2	0	0	Never	Negative	28.4	High
HNSCC_22	55	M	IVA	2	2C	0	Current	Negative	31.1	High
HNSCC_23	68	M	IVA	2	2C	0	Ex	Positive	24.4	Intermediate
HNSCC_24	75	M	III	2	1	0	Ex	Negative	N/A	N/A
HNSCC_25	50	M	III	1	1	0	Never	Positive	N/A	N/A
HNSCC_26	68	M	IVA	3	2B	0	Never	Positive	2	Low
HNSCC_27	62	F	IVA	4	1	0	Current	Negative	2.4	Low
HNSCC_28	29	F	II	2	0	0	Ex	Positive	27.4	Intermediate
HNSCC_29	61	F	IVA	2	2C	0	Current	Negative	20.5	Intermediate
HNSCC_30	52	M	IVA	4	0	0	Current	Negative	1.5	Low
HNSCC_31	70	F	II	2	0	0	N/A	Negative	11.2	Low
HNSCC_32	67	F	II	2	0	0	Ex	Negative	2.2	Low
HNSCC_33	60	M	IVA	2	2C	0	Never	Positive	45	High
HNSCC_34	57	M	IVA	1	2B	0	Smoker	Negative	47.7	High
HNSCC_35	51	M	IVB	2	3	0	Never	Positive	24.4	Intermediate
HNSCC_36	71	F	III	3	0	0	Never	Positive	41.4	High
HNSCC_37	63	M	IVA	3	2B	0	Ex	Negative	24.1	Intermediate
HNSCC_38	61	M	IVA	1	2B	0	N/A	Negative	N/A	N/A
HNSCC_39	63	M	IVA	4	2B	0	Current	Negative	6.4	Low
HNSCC_40	38	M	IVA	3	2B	0	Current	Positive	5	Low
HNSCC_41	62	M	II	2	0	0	Ex	Negative	2.5	Low

A.1.5 Clinical cohort of HNSCC cases with CD8 Cells sorted for transcriptomics analysis.

Summary of clinical features from the HNSCC cohort where CD8+ T cells were isolated, these were then assessed by RNA-Seq.

A.1.6 Clinical characteristics of samples used for assessment of tissue storage methodology.

Patient ID	Gender	Tumour type	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	Tissue comparison
AC001	M	NSCLC	T3	N1	M0	Smoker	Fresh, CT, DST
AC013	M	Melanoma	T4	N0	M1	NA	Fresh, CT, DST
AC020	F	Melanoma	T2	N1	M1	Smoker	Fresh, CT, DST
AC023	F	Melanoma	T4	N2	M1	Non-Smoker	Fresh, CT, DST
AC043	M	NSCLC	T2	N1	M0	Smoker	Fresh, CT, DST
AC066	F	Melanoma	T4	N1	M0	NA	CT, DST

A.1.6. Clinical characteristics of samples used for assessment of tissue storage methodology.

Details of samples used for tissue processing method evaluation in CD8+ T cells. Tumours were processed on the day of surgery (Fresh), Cryopreserved as 2mm tissue pieces (CT) and cryopreserved as a cell suspension (DST).

A.1.7 Clinical characteristics of samples used for evaluation of CD8+ tissue resident memory cells.

Patient ID	Gender	Tumour type	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	FACS experiment
649	Female	Adenocarcinoma	1b	2a	0	0	Smoker	TRM + Ki67
681	Male	Squamous cell carcinoma	2b	3	0	0	Non-Smoker	TRM + Ki67
683	Male	Adenocarcinoma	1b	2a	0	0	Ex-Smoker	TRM + Ki67
773	Male	Adenocarcinoma	1b	2a	0	0	Smoker	TRM + Ki67
805	Male	Adenocarcinoma	1a	1b	0	0	Smoker	TRM + Ki67
822	Female	Squamous cell carcinoma	3a	3	2	0	Ex-Smoker	TRM + Ki67
835	Male	Squamous cell carcinoma	2b	3	0	0	Ex-Smoker	TRM + Ki67
836	Male	Adenocarcinoma	1a	1b	0	0	Smoker	TRM + Ki67
839	Female	Adenocarcinoma	3a	2a	2	0	Non-Smoker	TRM + Ki67
842	Male	Adenocarcinoma	2a	1b	1	0	Ex-Smoker	TRM + Ki67
614	Male	Squamous cell carcinoma	3a	3	2	0	Ex-Smoker	TRM
629	Male	Adenocarcinoma	1b	2a	0	0	Smoker	TRM
641	Male	Adenocarcinoma	3a	3	2	0	Ex-Smoker	TRM
642	Male	Adenocarcinoma	1a	1a	0	0	Ex-Smoker	TRM
621	Female	Adenocarcinoma	1b	2a	0	0	Ex-Smoker	TRM
628	Female	Adenocarcinoma	1b	2a	0	0	Ex-Smoker	TRM

A.1.7. Clinical characteristics of samples used for evaluation of CD8+ tissue resident memory cells.

Clinical summary of the cases used for evaluating CD8+ tissue resident memory cells in NSCLC by flow cytometry.

A.1.8 Clinical characteristics of samples used for assessment of CD8+PD1+ and PD1- cells.

Patient ID	Gender	Tumour type	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status
AC006	F	NSCLC	T3	N2	M1	Non-Smoker
AC020	F	Melanoma	T2	N1	M1	Smoker
AC020	F	Melanoma	T2	N1	M1	Smoker
AC023	F	Melanoma	T4	N2	M1	Non-Smoker
AC042	M	NSCLC	T4	N0	M0	Non-Smoker
AC051	F	HNSCC	T2	N0	M0	Smoker

A.1.8. Clinical characteristics of samples used for assessment of CD8+PD1+ and PD1- cell.

Clinical summary of the cases used for evaluating CD8+PD1+ and PD1- T cells by RNA-Seq.

A.1.9 Clinical characteristics of samples used for evaluation of B cells at the single cell level.

Patient ID	Gender	Tumour type	Stage	Tumor status (T)	Nodal status (N)	Metastasis status (M)	Smoking status	CD8+ TIL score	No. Tumour B cells	No. NIL B cells	No. LN B cells
TL938	Male	Adenocarcinoma	1b	2a	0	0	Smoker	High	1500	1500	/
TL925	Male	Squamous cell carcinoma	1b	2a	0	0	Smoker	High	1500	1500	/
TL873	Male	Adenocarcinoma	1a	1b	0	0	/	High	1500	1500	/
TL920	Male	Adenocarcinoma	1b	2a	0	0	Smoker	High	1500	/	/
TL853	Female	Adenocarcinoma	2a	2b	0	0	Ex-Smoker	High	1500	/	/
TL846	Female	Adenocarcinoma	1a	1a	0	0	Smoker	High	1500	/	/
TL959	Female	Adenocarcinoma	3a	2a	2	0	Ex-Smoker	Low	1500	462	/
TL956	Male	Adenocarcinoma	2a	2b	0	0	Ex-Smoker	Low	1500	1500	/
TL29.111	Male	Adenocarcinoma	1a	1a	0	0	Ex-Smoker	Low	1500	1500	/
TL939	Female	Adenocarcinoma	1b	2a	0	0	Ex-Smoker	Low	1500	/	/
TL933	Male	Adenocarcinoma	3a	3	3	0	Ex-Smoker	Low	1500	/	/
TL924	Male	Adenocarcinoma	3a	1a	2	0	Smoker	Low	1500	/	/
TL1034	Male	EBUS of LN	/	/	/	/	Smoker	NA	/	/	3000
TL1033	Female	EBUS of LN	/	/	/	/	Ex-Smoker	NA	/	/	3000
TL1023	Female	EBUS of LN	/	/	/	/	Ex-Smoker	NA	/	/	3000

NIL = non-involved lung

LN = Lymphnode

EBUS = Endobronchial ultrasound biopsy

A.1.9. Clinical characteristics of samples used for evaluation of B cells at the single cell level.

The clinical characteristics for each case and the number of cells isolated from each tissue for subsequent single cell transcriptomics using the 10x Genomics platform.

A.2 Cell dissociation and immune cell sorting SOP384 (HARE40, AMG319 and Accelerator project)

Accelerator Programme_Southampton_SOP share_ March 2017
Ottensmeier Laboratory

SOP

SOP Number: 384
Version Number: 02
Manual Number:
Division/Unit: Ottensmeier Laboratory, Cancer Sciences

Processing of tumour tissue for RNA-Seq analysis: disaggregation, FACS staining and cell sorting

1 Introduction

This SOP describes how to process a tumour sample in order to create a single cell suspension that can then be used to identify individual cell populations, by staining with specific antibodies, and then sort these populations on a FACS sorter. RNA extraction of the sorted cell populations can then be performed and used for RNAseq analysis. Since the end purpose of this SOP is to analyse the number of RNA transcripts within specific cells of a purified population, manipulation of the cells needs to be kept to a minimum in order to reduce alternations in gene expression and thus RNA transcript levels.

2 Scope

This SOP can be applied to any endpoint laboratory for the purpose of tumour disaggregation and the purification of cell populations for RNA extraction and subsequent RNA seq analysis. This SOP is also applicable to other tissue samples, such as skin biopsies with a positive DTH response.

3 Process

3.1 Equipment Materials & Tools

Item of Equipment	Supplier	Code
TC-treated culture dishes	SLS	430167
Sterile Scalpel*	Fisher	11748353
MACS SmartStrainers (70 µm)	Miltenyi	130-098-462
Polypropylene tubes*	Sarstedt	62.554.002 (15 mL) 62.547.004 (50 mL)

RNAse free screw cap tubes	Starlab	14463
5mL syringe*	Fisher	10080264
Centrifuge*	SORVALL (Kendro)	LEGEND RT (CTCENT01)
	ThermoScientific	SL40R (CTCENT02)
	Fisher Scientific	accuSpin Micro R
FACS tubes	BD Falcon	352052
5ml polystyrene round-bottom tube with cell-strainer cap	BD Falcon	352235
Adjustable pipettes*	GILSON	P1000, P200, P20
-80 C freezer*		
Class II Laminar Flow Hood*	ScanLaf	Mars Safety Class 2
Neubauer Haemacytometer	Marienfeld	
Ice machine/Ice*	ITV Ice maker	
Orbital Shaker (incubator)*	Grant-bio	Model ES-20
Flow cytometer*	BD Biosciences	FACS Canto I/ II,
		FACS Aria
Accudrop beads	BD Biosciences	345249
Flow cytometry acquisition software*	BD Biosciences	FACSDiva6.1.3

*Equivalent equipment from alternative suppliers may be substituted, all relevant identification must be recorded in the lab. book/worksheet.

Materials /Reagents/Chemicals	**Manufacturer**	**Code**
RPMI	Gibco (Life Technologies)	42401-018
Sodium pyruvate (100mM) or	Gibco (Invitrogen)	11360-039
Sodium pyruvate (100mM) or	Sigma	S8636-100ml
Sodium pyruvate (100mM)	PAA	S11-003
Penicillin, streptomycin, glutamine (x100) (PSG) or	Gibco (Invitrogen)	10378-016
Penicillin-Streptomycin (10,000 units penicillin,10 mg streptomycin/mL) and L-Glutamine (200mM)	Sigma	P4333-100ML
	Sigma	G7513

DAPI (stock 200ug/ml)	Sigma	D9542
Liberase DL 20 units/ml	Roche	05401160001
DNAse 80000 units/ml	Sigma	D4527
Red cell Lysis Buffer	Qiagen	158904
MACS Buffer pH7.2:		
PBS 1x	VWR	437117K
EDTA 2mm	Invitrogen	15575-038
BSA 0.5%	Sigma	A7906
TRIzol LS	Ambion, Life Technologies	10296028
FcR blocking reagent, human	Miltenyi	130-059-901
CD45 FITC	Biolegend	304038
CD4 PE	Biolegend	317410
CD8a PerCP-Cy5.5	Biolegend	344710
CD19 PerCP-Cy5.5	Biolegend	302230
CD20 PerCP-Cy5.5	Biolegend	302326
CD3 PE-Cy7	Biolegend	344816
HLA-DR APC	Biolegend	307610
CD14 APC-Cy7	Biolegend	325620

3.2 Procedure

3.2.1 Tissue Disaggregation

3.2.1.1 The tumour tissue should be kept on ice (4 °C) prior to processing and while in transit from surgery. If the sample will be in transit for more than 10 mins, ensure it is kept in tissue media (CRPMI, see preparation in Appendix 1) at 4 °C. If the sample has been transported in media (CRPMI) proceed to step 3.2.1.2. If the sample is not in tissue media during transit, add 20ml of media RPMI to rinse off residual peripheral blood.

3.2.1.2 Discard the media and put the tumour sample in a TC-treated culture dish and rinse residual blood away with an additional 5ml of tissue media (RPMI). Using a 2ml pastette, remove the media and discard.

3.2.1.3 Using 2 x scalpels remove any blood clots and then mince the tumour by cutting it into small pieces, continue this process until the tumour has no pieces greater than $1mm^3$. If desired, collect 1-4 $1mm^3$ tumour fragments for subsequent TIL expansion; place tumour

fragments in ~1mL of freezing medium consisting of 90% FBS and 10% DMSO and control freeze in a Nalgene box (Mr Frosty).

3.2.1.4 Place tumour fragments in a universal tube containing 1-2 ml room temperature media (RPMI); the desired volume will depend on the size of the sample, in general samples <15mm^3 will need 1ml and samples >15mm^3 will require 2ml or more.

3.2.1.5 Add Liberase DL to give a final concentration of 0.15WU/mL, e.g. add 10uL of a 100X stock per mL of RPMI. Add DNase to give a final concentration of 800 Units/mL, e.g. add 10uL of a 100X stock per mL of RPMI (e.g. add 10uL per ml of media of a x100 stock). Incubate for 15 mins (RNA expression profiling) at 37°C on an orbital shaker set to 180 rpm. Do not exceed 20 mins.

3.2.1.6 Pass the digested tumour fragment solution through a 70um strainer into a 50mL falcon tube with the aid of a 5mL syringe plunger. Wash the cells through the strainer using 2mL of cold RPMI; repeat 10 times. Top up the volume to 30mL with cold RPMI and centrifuge at 485g, 4°C for 7 minutes.

3.2.1.7 Resuspend the cell pellet in RBC lysis solution at a 1:10 ratio (1, volume of cell pellet to 9, volume of RBC lysis buffer). Incubate for 10 minutes on ice or at 4°C. Add up to 30mL of cold MACS buffer to wash. Centrifuge at 485g, 4°C for 7 minutes

3.2.1.8 Resuspend the cell pellet in 0.5-1mL of MACS buffer, depending on the size of the pellet, and count cells according to CSD/SOP/306. Ensure that the cells remain on ice or at 4°C.

3.2.1.9 It is standard practice to store a proportion of the cell suspension for future analysis. The number of cells will be dependent on the overall yield, however, 50,000 to 250,000 is recommended. Take an appropriate volume of cell suspension and add it to 750uL of Trizol LS reagent in an RNase-free tube, which has been pre-labelled according to Section 2.3, e.g., DST for disaggregated tumour in TRIzol. N.B. Adjust the volume of cells/TRIzol to give a final volume of 1mL by the addition of the appropriate volume of 1X PBS. Vortex the sample and leave to stand for 5 minutes at room temperature prior to storage at -80°C. **This will serve as the unsorted sample.**

3.2.1.10 If there is excess material to requirements or further processing is not possible at this time, some or all of the disaggregated cell suspension may be cryopreserved as detailed in CSD/SOP/269.

3.2.2 FACS Stain protocol

3.2.2.1 Transfer 50,000 cells to a FACS tube to act as an unstained control.

3.2.2.2 Adjust the concentration of the cell suspension to 10x10^6 cells/mL with MACS buffer, equivalent to 1x10^6 cells in 100uL; this may require an additional centrifugation step. A concentration of 10x10^6 cells/mL delivers a practical flow rate on the flow cytometer.

3.2.2.3 Add 10uL FcR blocking reagent per 1x10^6 cells and incubate for 10 minutes at 4°C.

3.2.2.4 Without washing, add an appropriate volume of each antibody to the cell suspension; volumes listed below are shown for the staining of $1x10^6$ cells, adjust as required):

Antibodies	**Supplier**	**Vol per 10^6 cells/mL (uL)**
CD45 FITC	Biolegend (304038)	1.5
CD4 PE	Biolegend (317410)	2.5
CD8a PerCP-Cy5.5	Biolegend (344710)	2.5
CD19 PerCP-Cy5.5	Biolegend (302230)	2
CD20 PerCP-Cy5.5	Biolegend (302326)	2
CD3 PE-Cy7	Biolegend (344816)	1.5
HLA-DR APC	Biolegend (307610)	1.25
CD14 APC-Cy7	Biolegend (325620)	3

Incubate with antibody for 30 minutes at 4°C, in the dark. N.B. The sort stream may be established during this step, see below.

3.2.2.5 Add up to 2mL of MACS buffer to wash. Centrifuge at 485g, 4°C for 5 minutes.

3.2.2.6 Discard the supernatant and resuspend the cells in an appropriate volume of MACS buffer to maintain a cell concentration of ≤10x106 cell/mL, e.g., 500uL for $5x10^6$ cells, however, further dilution may be required if the event rate on the flow cytometer is too high.

3.2.3 Cell Isolation and FACS Aria stream setup (Please refer to local guidance procedures and SOPs for detailed sort stream setup)

3.2.3.1 Ensure the FACS sorter is equipped with a 100um nozzle.
3.2.3.2 Turn on the stream and let it stabilise for ~10 mins.
3.2.3.3 Turn on the refrigeration unit and attach the chilled sort block.

3.2.3.4 **Stream setup**:

- Using BD Accudrop beads at a flow rate of 3000 events/second, set the amplitude to achieve a break off point in the 1st quarter of the stream.
- Ensure the GAP is >8 and <12 and match the Live GAP and Drop 1 values to the set values.
- Turn the sweet spot on and place a screw cap, RNase-free tube in position 1 (with the lid on) and a 5ml FACS tube in position 2 of the 4-way sort block.
- Open the sort layout and set the instrument to 4-way sorting and the purity setting.

- Begin sorting all the fluorescent Accudrop beads into the 5ml FACS tube at 1000-1500 events/second. Turn on the optical filter and adjust the Drop delay so as to achieve as close to 100% as possible on the purity setting (minimum 95%).
- Once complete, deactivate the optical filter and remove the FACS tube from the sort block.
- Place 4 screw cap tubes (with the lids on) in each position of the 4-way sort block.
- Select the test sort to view the side stream positions. Rapidly select then deselect the waste drawer option to allow a small droplet to be sorted onto the top of each screw cap tube. Adjust the side stream positions to ensure the droplet falls into the centre of the tube, and then proceed to the sort.

3.2.3.5 **Acquisition and Sort:**

- For acquisition and sort, use the gating strategy as illustrated in Figure 1 below. Briefly, for CD8+ and CD4+ T cells and B cells:
  - From the singlet gate, gate on the live cells,
  - From the live cells, gate on the $CD45^+$ cells, followed by CD3+ cells
  - From the $CD3^+$ cells, gate on the $CD4^+$ and $CD8^+$ cells
  - From the not $CD3^+$ cells, gate on the $CD19^+$ and $CD20^+$ cells

  For the macrophages and tumour cells:
  - From the macrophages singlets, gate on the live cells (Macro live)
  - From the macro live cells, gate on $CD45^-$ (tumour cells) and $CD45^+$ cells (Macro $CD45^+$)
  - From the macro $CD45^+$ cells, gate on the $CD14^+$ $HLA\text{-}DR^+$ cells

- Place 4 x 1.5ml labelled screw cap RNase free tubes, pre-filled with 750ul Trizol LS, into positions 1-4 in the sort block.
- Immediately prior to sorting, filter the cells through a 30-35um filter into a FACS tube (alternatively use a blue capped FACS tube strainer).
- Run the sample (unstained and stained) both with and without the addition of the live/dead marker DAPI:
  - Run the unstained sample with no DAPI and record the data for 10,000 events
  - Add DAPI (stock of 200ug/mL) at 2uL per $1x10^6$ cells to the unstained sample, run and record the data for 10,000 events
  - Run the stained sample with no DAPI and record the data for 10,000 events
  - Add DAPI (stock of 200ug/mL) at 2uL per $1x10^6$ cells to the unstained sample, run and record the data for 10,000 events
- Set the gating strategy by recording <50,000 events and use this to ensure that all immune cell subsets are appropriately gated for sorting.
- As an example, to sort $CD8^+$, $CD4^+$, B cells, macrophages, tumour cells, set the sort layout as follow:
  - Macrophages ($CD45^+HLA\text{-}DR^+CD14^+$)
  - CD8 ($CD45^+$ $CD3^+$ $CD8^+$)
  - CD4 ($CD45^+$ $CD3^+$ $CD4^+$)
  - Tumour ($CD45^-$), followed by B cells ($CD45^+$, NOT $CD3^+$, $CD20^+$,$CD19^+$)

- Initiate immune cell sorting and ensure that the flow rate gives <2,000 events/second and an efficiency of >80% for smaller samples; the flow rate can be increased with larger samples, but be aware that this will compromise the efficiency and yield of the sort.
- Record >500,000 events whilst sorting for analysis
- Sort in 10 minute intervals or until 50,000 cell have been sorted. At this point, stop the acquisition and sorting, remove the tube from the sort block and vortex for ~1 minute. Ensure that the number of sorted cells are recorded each time the sort is paused so that the number of sorted cells can be tracked to ensure that no more than 50,000 are sorted into each tube containing TRIzol.
- Collect up to 50,000 cells per tube, with a minimum of 1 tube per immune cell subset; multiple tubes can be collected if sufficient material exists. 50,000 cells are equivalent to 250ul, making a total volume of 1mL with 750ul of Trizol. Adjust the final volume to 1mL with PBS if less cells are collected; for example, if only 25,000 cells are available for collection, add 125ul of PBS post-sort into the collection tube so concentrations of Trizol are appropriate.
- Leave sorted cells in Trizol at RT for 5 mins prior to short term storage at 4°C. Transfer to -80°C freezer for long term storage

4 Related Documents and Attachments

Appendix 1 Preparation of RPMI

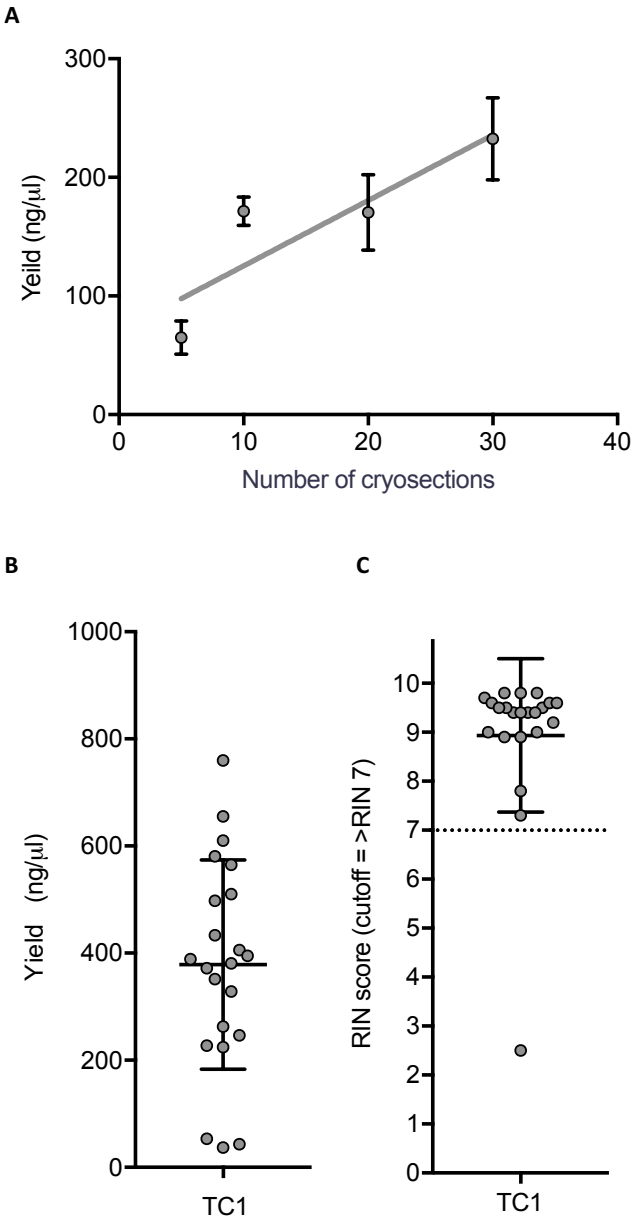
CSD/FORM/350 Tumour Sample Disaggregation Processing Sheet

SOP Number	384	Version Number	02	Page **7** of **8**
SOP Title	Processing of tumour tissue for RNA-Seq analysis: disaggregation, FACS staining and cell sorting			

A.2. Cell dissociation and immune cell sorting SOP384 (HARE40, AMG319 and Accelerator project.

Details of the standard operating procedure (SOP) written for the isolation of immune cells from tumours.

A.3 Optimising RNA isolation from frozen tumour tissue



A.3. Optimisation of RNA isolation using mouse TC1 tumour.

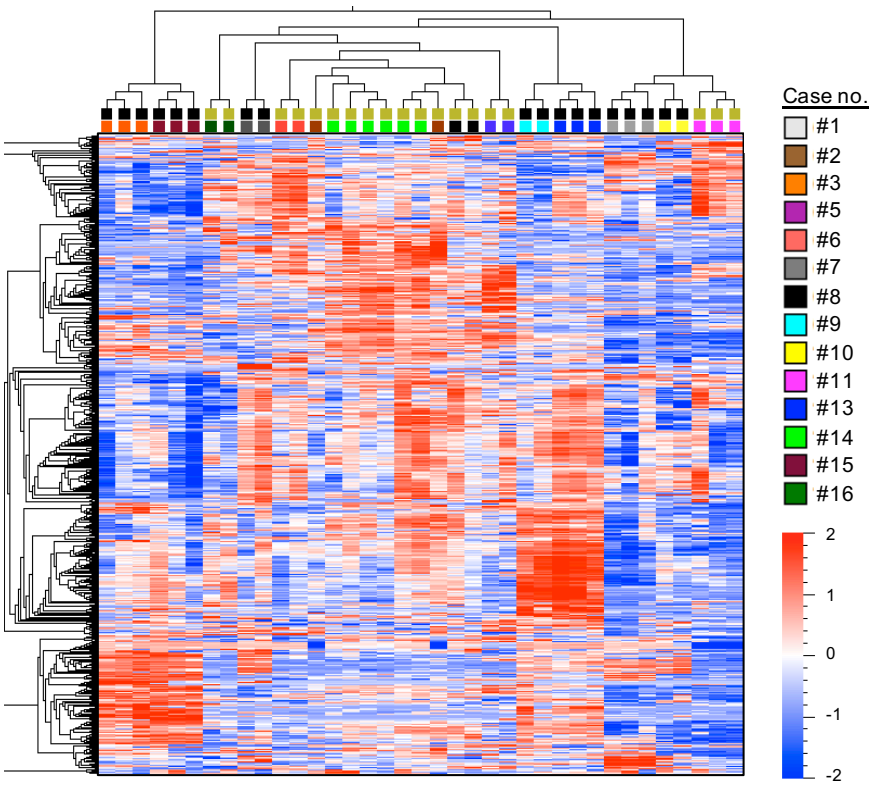
(A) Titration of tumour cryosections and RNA yield. The yield of RNA increases with a linear relationship with increasing cryosections. (B) Yield and (C) RIN (RNA integrity number) for 22 mouse TC1 tumours. The average RNA yield was 378 ng/µl with an average RIN of 8.9 highlighting excellent RNA yield and quality.

Appendix B

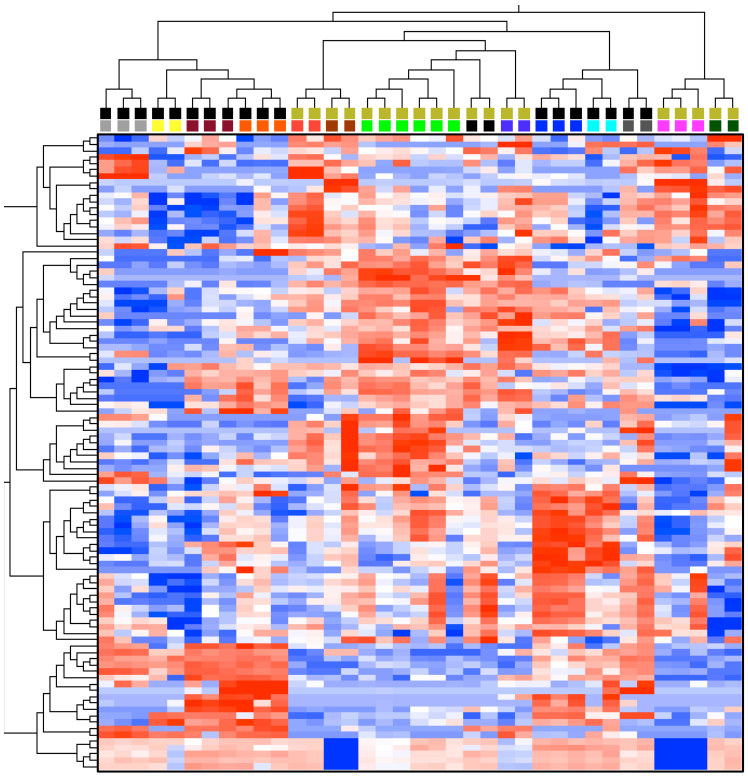
B.1 Tumour replicates

B.1.1 RNA-Seq analysis showing row wise z-scores of normalized read counts for a decreasing number of genes (variance filtered) across the tumour replicates

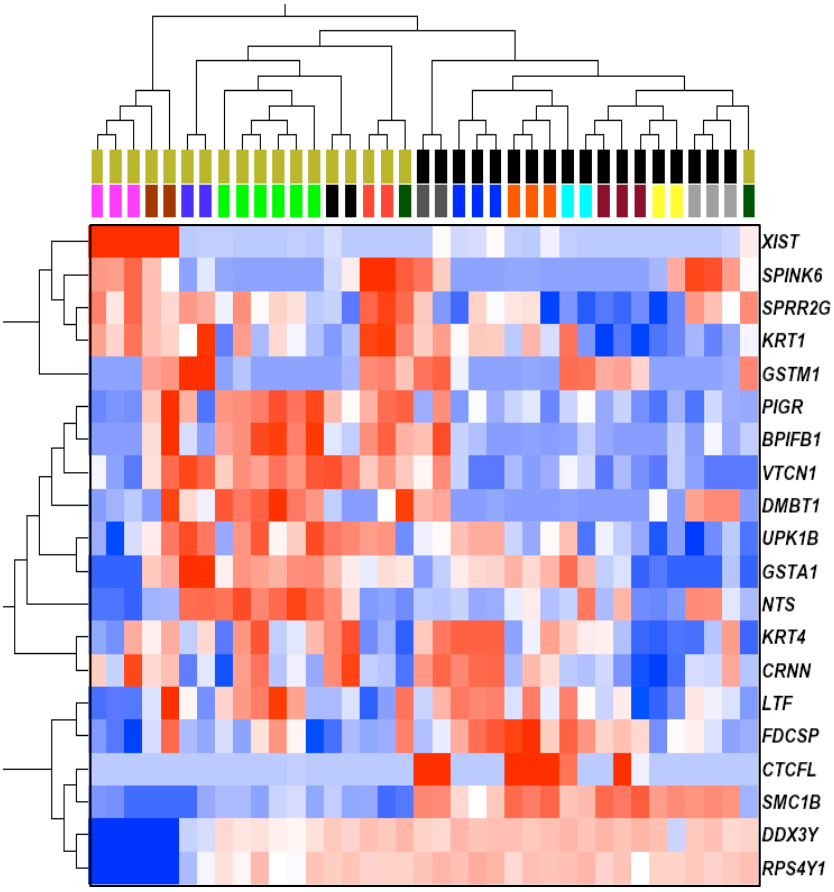
A



B



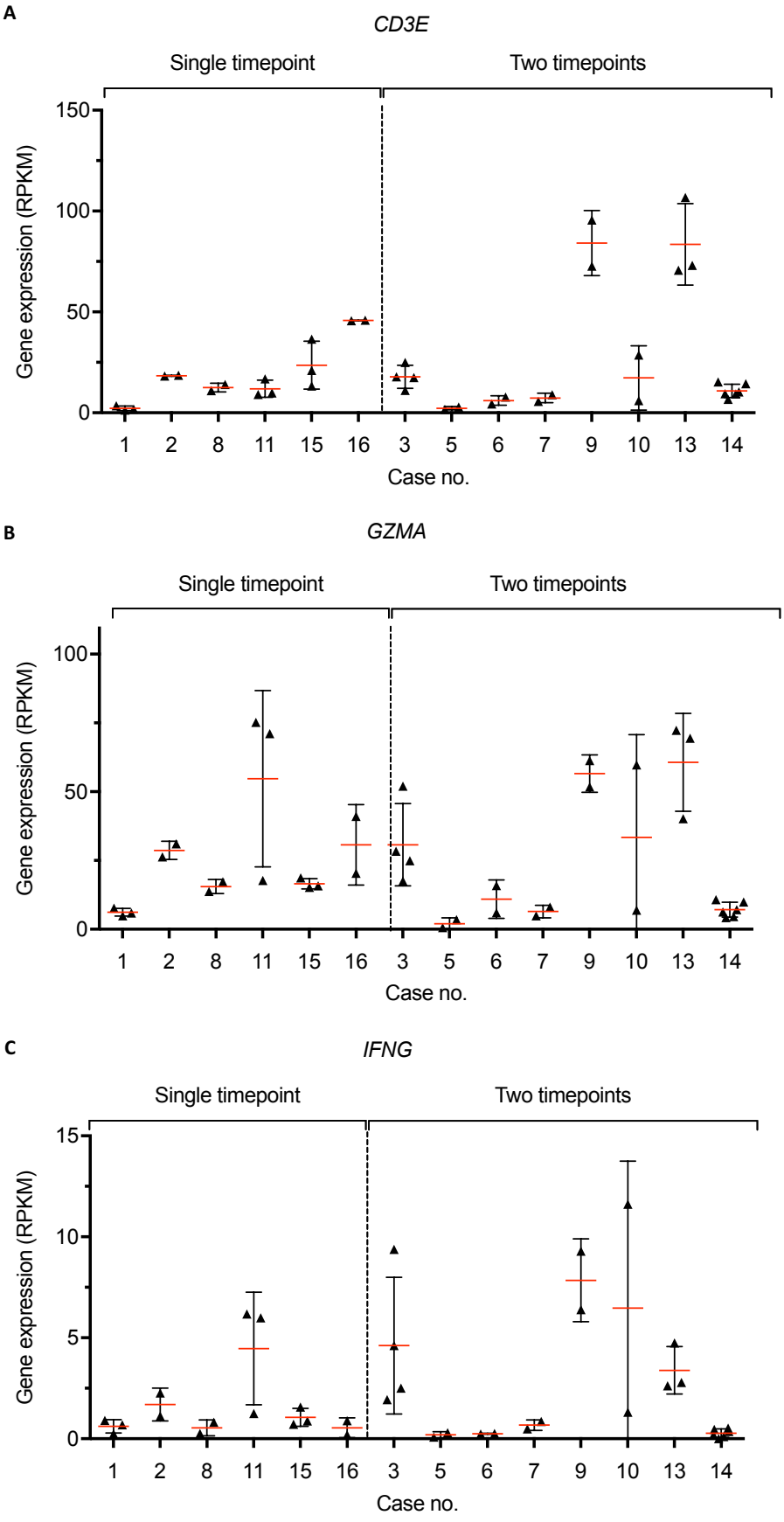
C



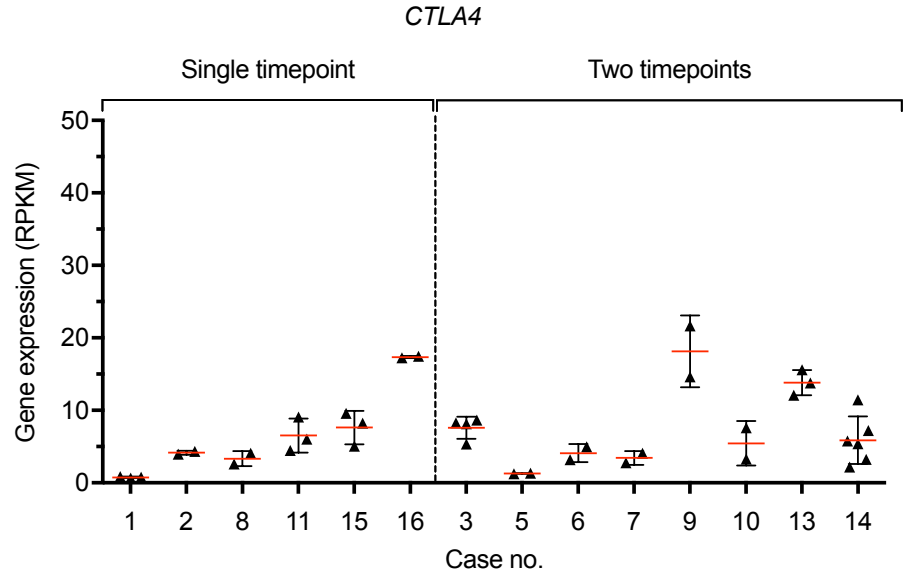
B1.1. RNA-Seq analysis showing row wise z-scores of normalized read counts for a decreasing number of genes (variance filtered) across the tumour replicates.

RNA-Seq analysis showing row wise z-scores of normalized read counts for a decreasing number of genes (variance filtered) across the tumour replicates. Patient tumour replicates are colour coded and displayed on the heatmap and principle component analysis (PCA) plot; HPV(+)=black and HPV(-)=beige; diagnostic biopsy (DB) and surgical resection (SR) replicates are annotated below. (A) 1000 genes, (B) 100 genes and (C) 20 genes. Hierarchical clustering (distance measure=Pearson's correlation metric; clustering=average linkage method) of tumour replicates displayed as a heatmap shows close clustering of related samples across the different number of genes.

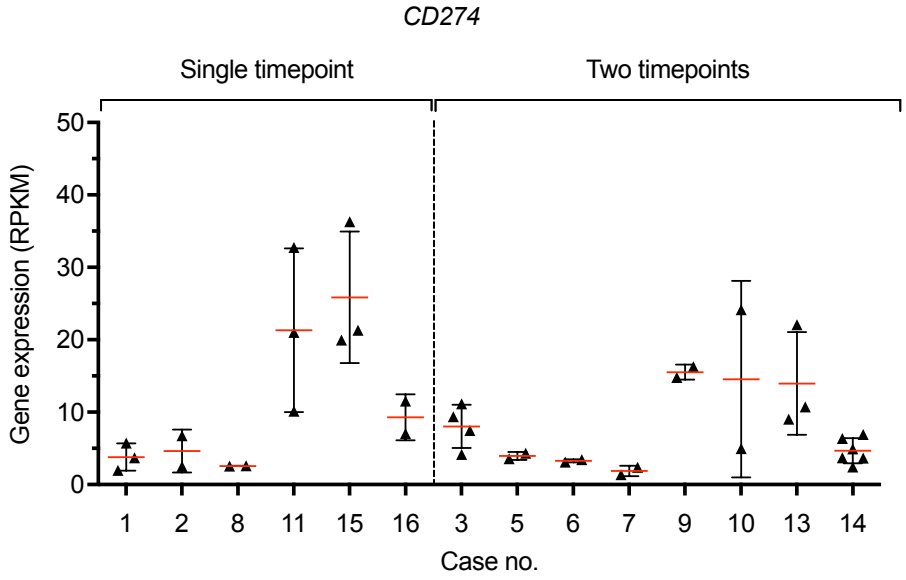
B.1.2 Gene expression of immune genes across the sample cohort



D



E

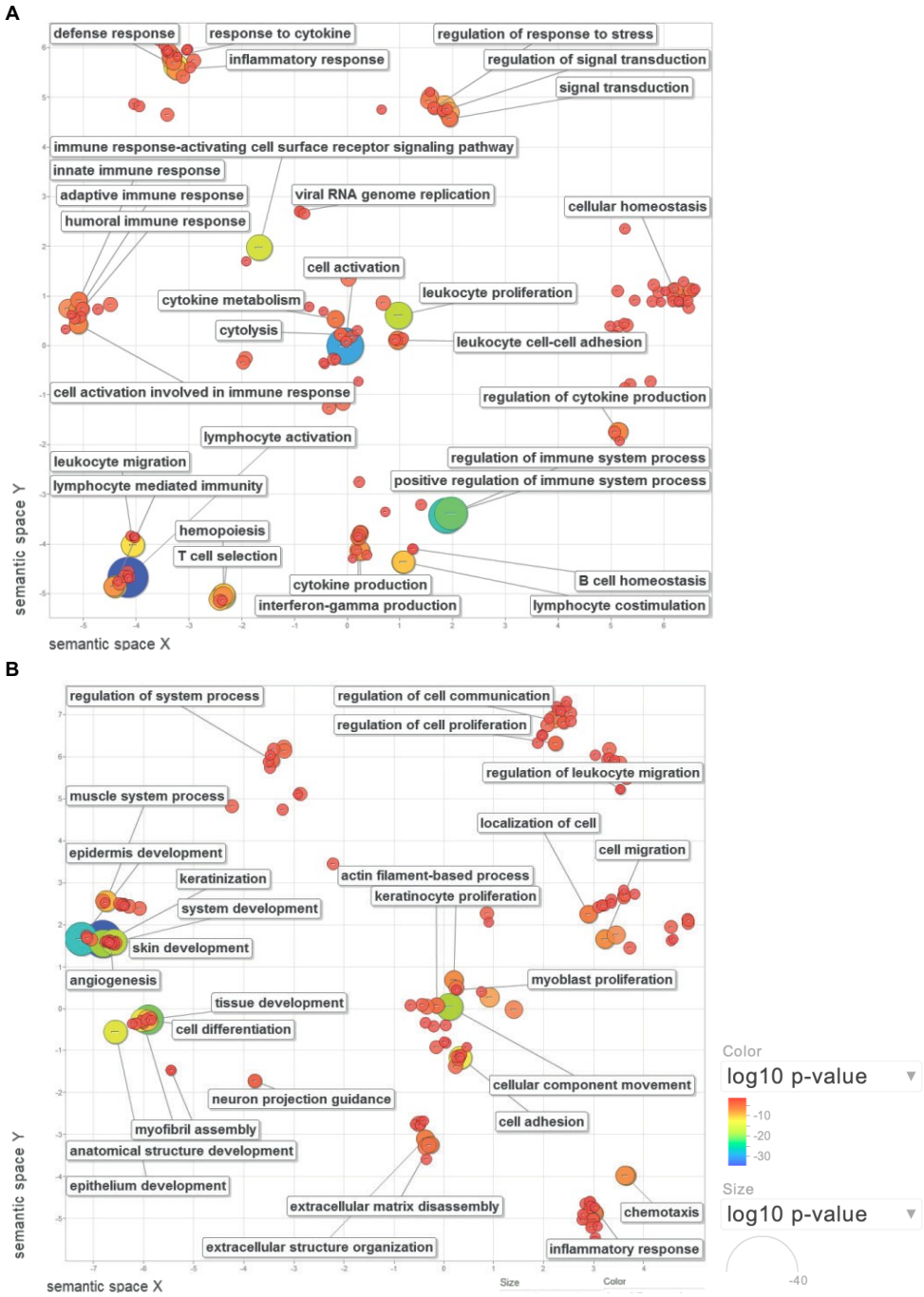


B1.2. Gene expression of immune genes across the sample cohort.

Gene expression (RPKM – Reads per kilobase per million mapped reads) of immunologically relevant genes for the spatial and temporal tumour replicates. Samples from the same patient display a high level of similarity between both spatial and temporal replicates for the genes CD3E, GZMA, IFNG, CTLA4 and CD274 (PDL1), (A-E) respectively.

B.2 Differentially expressed genes between HPV(+) and (-) (no TIL correction).

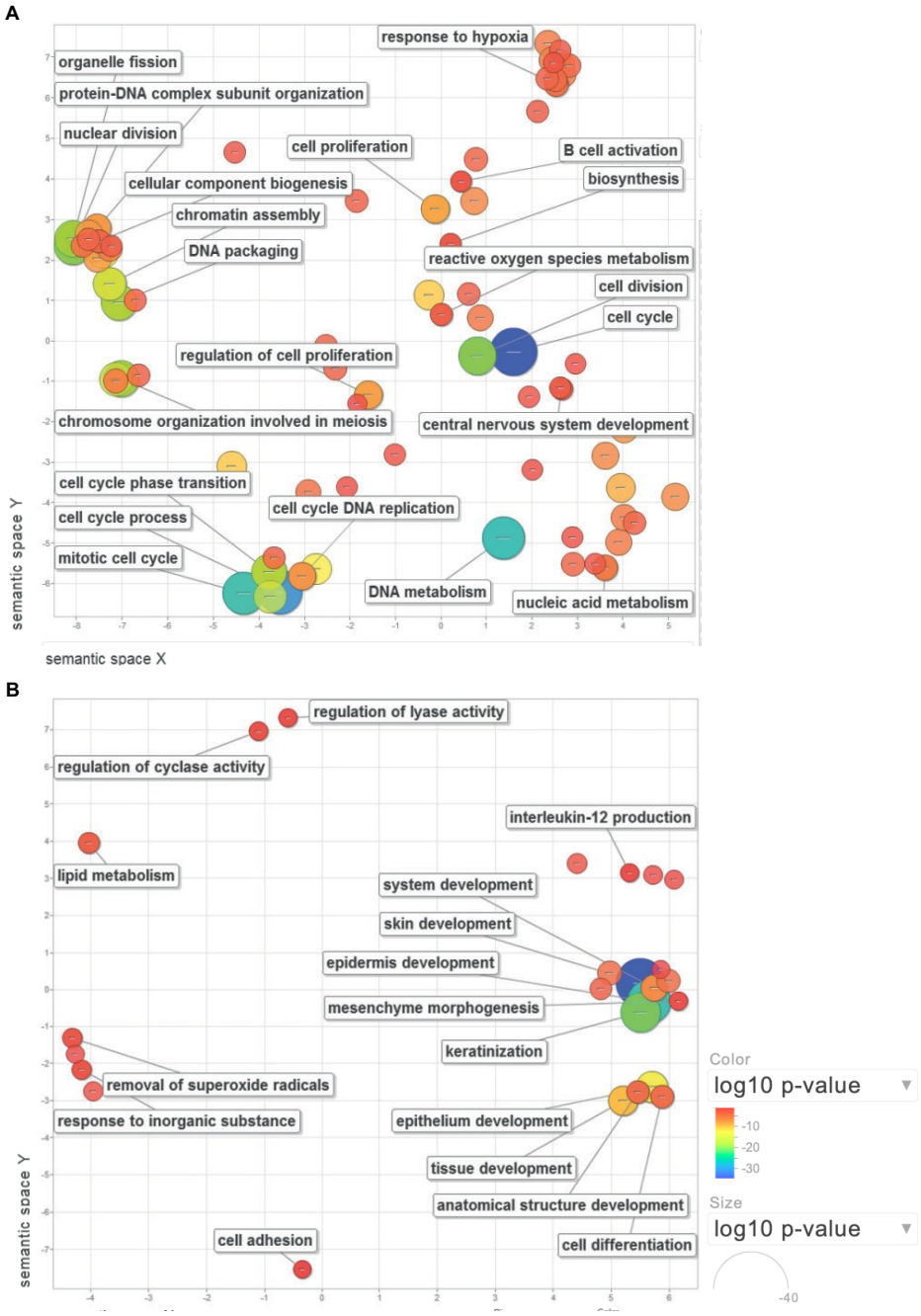
B.2.1 Gene ontology analysis using REVIGO scatterplots



B.2.1. Gene ontology analysis using REVIGO scatterplots.

Gene ontology analysis of the greater and lesser expressed genes between HPV(+) and (-), where REVIGO scatterplots show the GO terms after redundancy reduction in a 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05). (A) Gene ontology terms associated with the greater expression between HPV(+) and (-) samples, a clear immunological signature was observed. (B) Gene ontology terms associated with the lesser expression between HPV(+) and (-) samples, a signature of tissue development and keratinisation is present.

B.2.2 Gene ontology analysis using REVIGO scatterplots



B.2.2. Gene ontology analysis using REVIGO scatterplots on immune cell corrected differential gene analysis.

Gene ontology analysis of the greater and lesser expressed genes between HPV(+) and (-) corrected for immune cell density, where REVIGO scatterplots show the GO terms after redundancy reduction in a 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05). (A) Gene ontology terms associated with the greater expression between HPV(+) and (-) samples, a cell cycle and proliferation signature was observed, with a single immune GO (B cell activation). (B) Gene ontology terms associated with the lesser expression between HPV(+) and (-) samples, a signature of tissue development and keratinisation remained.

B.2.3 A list of differentially expressed genes between HPV(+) and HPV(-) tumours identified by RNA-Seq analysis.

*Genes that overlap with DEGs identified from TIL corrected data				
Gene Symbol	**Gene/Protein name**	**Fold Change (Log)**	***q*-value**	
SYCP2	synaptonemal complex protein 2	4.57	1.86E-26	*
STAG3	stromal antigen 3	5.05	3.29E-24	*
ZFR2	zinc finger RNA binding protein 2	5.84	1.93E-21	*
SMC1B	structural maintenance of chromosomes 1B	5.64	1.19E-19	*
RNF212	ring finger protein 212	5.19	1.94E-19	*
TEX15	testis expressed 15	6.93	8.93E-19	*
KEL	Kell blood group, metallo-endopeptidase	6.59	7.88E-17	*
LOC254559	hypothetical LOC254559	4.13	4.83E-15	*
MAP7D2	MAP7 domain containing 2	5.26	5.87E-15	*
NEFH	neurofilament, heavy polypeptide	5.82	2.69E-13	*
WNK3	WNK lysine deficient protein kinase 3	3.22	2.07E-11	*
SLFN13	schlafen family member 13	2.32	6.96E-11	*
TAF7L	TAF7-like RNA polymerase II, TATA box binding protein (TBP)-associated fact	4.40	8.31E-10	*
SYCE2	synaptonemal complex central element protein 2	2.84	2.32E-09	*
MS4A1	membrane-spanning 4-domains, subfamily A, member 1	3.96	2.46E-09	
SOX30	SRY (sex determining region Y)-box 30	3.63	3.28E-09	*
FCRL2	Fc receptor-like 2	3.44	3.51E-09	
C14orf39	chromosome 14 open reading frame 39	5.17	4.38E-09	*
UPB1	ureidopropionase, beta	3.90	4.38E-09	*
CD22	CD22 molecule	3.75	7.07E-09	
SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	3.80	8.05E-09	*
ZNF541	zinc finger protein 541	4.83	1.08E-08	*
CXCR5	chemokine (C-X-C motif) receptor 5	3.64	1.20E-08	
ZYG11A	zyg-11 homolog A (C. elegans)	3.03	1.33E-08	*
ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	3.48	1.39E-08	*
BANK1	B-cell scaffold protein with ankyrin repeats 1	2.53	1.46E-08	
SCG5	secretogranin V (7B2 protein)	-1.47	4.33E-02	
LOC643650	hypothetical protein LOC643650	-1.08	4.33E-02	
HCN2	hyperpolarization activated cyclic nucleotide-gated potassium channel 2	-1.35	4.40E-02	
TGM1	transglutaminase 1 (K polypeptide epidermal type I, protein-glutamine-gamr	-1.75	4.41E-02	
RAET1G	retinoic acid early transcript 1G	-1.30	4.42E-02	
LOC375295	hypothetical protein LOC375295	-1.29	4.43E-02	
CA2	carbonic anhydrase II	-1.26	4.46E-02	
LOC729178	hypothetical protein LOC729178	-1.80	4.46E-02	*
HS3ST3A1	heparan sulfate (glucosamine) 3-O-sulfotransferase 3A1	-1.37	4.52E-02	
NACAD	NAC alpha domain containing	-1.77	4.53E-02	
PGLYRP3	peptidoglycan recognition protein 3	-1.45	4.61E-02	
STC1	stanniocalcin 1	-1.27	4.66E-02	
S100A1	S100 calcium binding protein A1	-1.47	4.68E-02	
SAA2	serum amyloid A2	-1.94	4.69E-02	
AKR1C2	aldo-keto reductase family 1, member C2 (dihydrodiol dehydrogenase 2; bile	-1.76	4.72E-02	
C5orf38	chromosome 5 open reading frame 38	-1.24	4.74E-02	
RAET1L	retinoic acid early transcript 1L	-1.43	4.78E-02	
TMEM86A	transmembrane protein 86A	-1.77	4.78E-02	
AKAP6	A kinase (PRKA) anchor protein 6	-1.29	4.78E-02	
TTLL11	tubulin tyrosine ligase-like family, member 11	-1.20	4.86E-02	
COL4A1	collagen, type IV, alpha 1	-1.00	4.89E-02	
HEYL	hairy/enhancer-of-split related with YRPW motif-like	-1.04	4.89E-02	
ISM1	isthmin 1 homolog (zebrafish)	-1.19	4.90E-02	
OSBP2	oxysterol binding protein 2	-1.15	4.90E-02	
PTX3	pentraxin-related gene, rapidly induced by IL-1 beta	-1.62	4.98E-02	
HIST1H2BG	histone cluster 1, H2bg	-1.26	5.00E-02	

B.2.3 A list of differentially expressed genes between HPV(+) and HPV(-) tumours identified by RNA-Seq analysis.

B.2.4 Gene ontology analysis of DEGs expressed in HPV(+) vs HPV(-) tumours

DEGs identified with a the FDR adjusted *p* -value <0.05 (i.e., *q* -value <0.05) and a fold change of >2 or <-2.

Genes expressed to a greater extent in HPV(+) compared to HPV(-) tumors

Gene Ontology term	Set Size	Candidates	*q* -value
GO:0045321 leukocyte activation	631	118 (18.8%)	8.29E-44
GO:0046649 lymphocyte activation	539	108 (20.2%)	2.81E-42
GO:0001775 cell activation	847	125 (14.8%)	5.98E-35
GO:0042110 T cell activation	392	84 (21.6%)	2.15E-34
GO:0002682 regulation of immune system process	1153	144 (12.6%)	1.43E-32
GO:0006955 immune response	1367	157 (11.5%)	7.78E-32
GO:0002694 regulation of leukocyte activation	376	77 (20.7%)	3.35E-30
GO:0050865 regulation of cell activation	407	79 (19.6%)	1.74E-29
GO:0002684 positive regulation of immune system process	696	102 (14.8%)	7.42E-28
GO:0009897 external side of plasma membrane	218	53 (24.5%)	1.58E-24
GO:0044459 plasma membrane part	2165	192 (8.9%)	1.63E-24
GO:0098552 side of membrane	299	60 (20.3%)	1.26E-23
GO:0030098 lymphocyte differentiation	267	57 (21.4%)	4.28E-23
GO:0050776 regulation of immune response	779	100 (12.9%)	1.06E-22
GO:0002252 immune effector process	591	78 (13.3%)	9.70E-19
GO:0046651 lymphocyte proliferation	226	47 (21.1%)	1.11E-18
GO:0032943 mononuclear cell proliferation	228	47 (20.9%)	1.43E-18
GO:0070661 leukocyte proliferation	238	48 (20.4%)	1.61E-18
GO:0042113 B cell activation	203	44 (21.8%)	3.87E-18
GO:0031226 intrinsic component of plasma membrane	1338	127 (9.5%)	8.60E-18
GO:0005887 integral component of plasma membrane	1286	123 (9.6%)	1.26E-17
GO:0002757 immune response-activating signal transduction	371	58 (15.7%)	7.86E-17
GO:0002429 immune response-activating cell surface receptor signaling pathway	230	45 (19.7%)	8.04E-17
GO:0009986 cell surface	561	71 (12.7%)	4.89E-16
GO:0002253 activation of immune response	423	60 (14.3%)	7.48E-16

Genes expressed to a lesser extent in HPV(+) compared to HPV(-) tumors

Gene Ontology term	Set Size	Candidates	*q* -value
GO:0043588 skin development	289	66 (22.9%)	1.69E-30
GO:0008544 epidermis development	289	58 (20.1%)	1.03E-23
GO:0009888 tissue development	1441	128 (8.9%)	4.49E-18
GO:0031424 keratinization	46	21 (45.7%)	5.21E-16
GO:0006928 cellular component movement	1571	128 (8.2%)	2.20E-15
GO:0048513 organ development	2552	176 (6.9%)	7.88E-15
GO:0048731 system development	3615	223 (6.2%)	4.42E-14
GO:0044707 single-multicellular organism process	6199	326 (5.3%)	2.59E-12
GO:0048856 anatomical structure development	4248	241 (5.7%)	1.67E-11
GO:0071944 cell periphery	4682	260 (5.6%)	1.76E-11
GO:0007155 cell adhesion	1032	87 (8.4%)	2.37E-11
GO:0060429 epithelium development	757	73 (9.7%)	2.69E-11
GO:0043292 contractile fiber	204	33 (16.2%)	1.28E-10
GO:0030017 sarcomere	172	30 (17.4%)	2.06E-10
GO:0044767 single-organism developmental process	5037	268 (5.3%)	3.14E-10
GO:0007275 multicellular organismal development	4394	243 (5.5%)	3.28E-10
GO:0044449 contractile fiber part	188	30 (16.0%)	6.75E-10
GO:0005886 plasma membrane	4582	247 (5.4%)	9.80E-10
GO:0044459 plasma membrane part	2165	138 (6.4%)	2.36E-09
GO:0005856 cytoskeleton	1905	127 (6.7%)	2.72E-09
GO:0030154 cell differentiation	3153	185 (5.9%)	8.26E-09
GO:0003012 muscle system process	314	39 (12.4%)	8.26E-09
GO:0048869 cellular developmental process	3345	189 (5.7%)	4.33E-08
GO:0042330 taxis	647	56 (8.7%)	9.18E-08
GO:0031674 I band	116	21 (18.1%)	1.10E-07

B.2.4 Gene ontology analysis of DEGs expressed in HPV(+) vs HPV(-) tumours.

B.2.5 Pathway analysis of DEGs expressed to in HPV(+) vs HPV(-) tumours.

DEGs identified with a the FDR adjusted *p*-value <0.05 (i.e., *q*-value <0.05) and a fold change of >2 or <-2.

Genes expressed to a greater extent in HPV(+) compared to HPV(-) tumors

Pathway Name	Set Size	Candidates	Pathway Source	*q*-value
Primary immunodeficiency - Homo sapiens (human)	36	18 (50.0%)	KEGG	1.23E-12
TCR	244	43 (17.7%)	NetPath	1.27E-12
Cell adhesion molecules (CAMs) - Homo sapiens (human)	145	32 (22.5%)	KEGG	3.27E-12
Adaptive Immune System	569	66 (11.6%)	Reactome	4.58E-11
TCR signaling in naïve CD4+ T cells	68	21 (30.9%)	PID	1.16E-10
TCR signaling in naïve CD8+ T cells	55	19 (34.5%)	PID	1.28E-10
Intestinal immune network for IgA production - Homo sapiens (human)	49	17 (36.2%)	KEGG	7.72E-10
BCR	129	27 (20.9%)	NetPath	1.03E-09
Immune System	1007	90 (9.0%)	Reactome	2.31E-09
Immunoregulatory interactions between a Lymphoid and a non-Lymphoid	116	24 (20.9%)	Reactome	1.23E-08
Generation of second messenger molecules	36	14 (38.9%)	Reactome	1.23E-08
IL12-mediated signaling events	64	18 (28.1%)	PID	1.30E-08
TCR signaling	65	18 (27.7%)	Reactome	1.60E-08
Hematopoietic cell lineage - Homo sapiens (human)	88	20 (22.7%)	KEGG	6.88E-08
Cytokine-cytokine receptor interaction - Homo sapiens (human)	265	36 (13.6%)	KEGG	8.73E-08
TCR Signaling Pathway	92	20 (21.7%)	Wikipathways	1.39E-07
Costimulation by the CD28 family	72	17 (23.6%)	Reactome	5.68E-07
PD-1 signaling	28	11 (39.3%)	Reactome	6.74E-07
Ick and fyn tyrosine kinases in initiation of tcr activation	13	8 (61.5%)	BioCarta	7.80E-07
Translocation of ZAP-70 to Immunological synapse	23	10 (43.5%)	Reactome	8.87E-07
B Cell Receptor Signaling Pathway	94	19 (20.2%)	Wikipathways	9.61E-07
Phosphorylation of CD3 and TCR zeta chains	25	10 (40.0%)	Reactome	2.13E-06
Leishmaniasis - Homo sapiens (human)	74	16 (22.2%)	KEGG	2.84E-06
NF-kappa B signaling pathway - Homo sapiens (human)	91	18 (19.8%)	KEGG	2.84E-06
Integrin cell surface interactions	66	15 (22.7%)	Reactome	5.04E-06

Genes expressed to a lesser extent in HPV(+) compared to HPV(-) tumors

Pathway Name	Set Size	Candidates	Pathway Source	*q*-value
Striated Muscle Contraction	38	13 (34.2%)	Wikipathways	1.71E-06
Striated Muscle Contraction	32	11 (34.4%)	Reactome	1.52E-05
Assembly of collagen fibrils and other multimeric structures	44	12 (27.3%)	Reactome	4.35E-05
Extracellular matrix organization	261	30 (11.5%)	Reactome	7.10E-05
Type I hemidesmosome assembly	9	6 (66.7%)	Reactome	7.64E-05
Collagen formation	88	16 (18.2%)	Reactome	7.77E-05
Basigin interactions	26	9 (34.6%)	Reactome	7.77E-05
Muscle contraction	52	12 (23.1%)	Reactome	0.000119
Degradation of the extracellular matrix	84	15 (17.9%)	Reactome	0.000155
Laminin interactions	23	8 (34.8%)	Reactome	0.000231
Antagonism of Activin by Follistatin	4	4 (100.0%)	Reactome	0.000284
Focal adhesion - Homo sapiens (human)	207	24 (11.6%)	KEGG	0.000331
Amoebiasis - Homo sapiens (human)	109	16 (14.7%)	KEGG	0.000688
Cell junction organization	86	14 (16.3%)	Reactome	0.000699
Beta1 integrin cell surface interactions	66	12 (18.2%)	PID	0.0009
Validated transcriptional targets of AP1 family members Fra1 and Fra2	37	9 (24.3%)	PID	0.000908
Activation of Matrix Metalloproteinases	30	8 (26.7%)	Reactome	0.00126
Transport of glucose and other sugars, bile salts and organic acids, metal	97	14 (14.4%)	Reactome	0.00223
Alpha6 beta4 integrin-ligand interactions	11	5 (45.5%)	PID	0.00249
Hair Follicle Development- Induction (Part 1 of 3)	43	9 (20.9%)	Wikipathways	0.00265
Urokinase-type plasminogen activator (uPA) and uPAR-mediated signalin	44	9 (20.5%)	PID	0.00306
Cell-Cell communication	128	16 (12.5%)	Reactome	0.00313
a6b1 and a6b4 Integrin signaling	45	9 (20.0%)	PID	0.00337
Anchoring fibril formation	7	4 (57.1%)	Reactome	0.00413
AP-1 transcription factor network	70	11 (15.7%)	PID	0.00505

Appendices B.2.5 Pathway analysis of DEGs expressed to in HPV(+) vs HPV(-) tumours

B.3 Differentially expressed genes between HPV(+) and (-) corrected for immune cell numbers.

B.3.1 A list of differentially expressed genes between HPV(+) and HPV(-) tumours identified by RNA-Seq analysis followed by correction for TIL number.

*Genes that overlap with DEGs identified from uncorrected data

Gene Symbol	Gene/Protein name	Fold Change (Log)	*q*-value	
SYCP2	synaptonemal complex protein 2	4.57	1.86E-26	*
STAG3	stromal antigen 3	5.05	3.29E-24	*
ZFR2	zinc finger RNA binding protein 2	5.84	1.93E-21	*
SMC1B	structural maintenance of chromosomes 1B	5.64	1.19E-19	*
RNF212	ring finger protein 212	5.19	1.94E-19	*
TEX15	testis expressed 15	6.93	8.93E-19	*
KEL	Kell blood group, metallo-endopeptidase	6.59	7.88E-17	*
LOC254559	hypothetical LOC254559	4.13	4.83E-15	*
MAP7D2	MAP7 domain containing 2	5.26	5.87E-15	*
NEFH	neurofilament, heavy polypeptide	5.82	2.69E-13	*
WNK3	WNK lysine deficient protein kinase 3	3.22	2.07E-11	*
SLFN13	schlafen family member 13	2.32	6.96E-11	*
TAF7L	TAF7-like RNA polymerase II, TATA box binding protein (TBP)-associated	4.40	8.31E-10	*
SYCE2	synaptonemal complex central element protein 2	2.84	2.32E-09	*
SOX30	SRY (sex determining region Y)-box 30	3.63	3.28E-09	*
C14orf39	chromosome 14 open reading frame 39	5.17	4.38E-09	*
UPB1	ureidopropionase, beta	3.90	4.38E-09	*
SPIB	Spi-B transcription factor (Spi-1/PU.1 related)	3.80	8.05E-09	*
ZNF541	zinc finger protein 541	4.83	1.08E-08	*
ZYG11A	zyg-11 homolog A (C. elegans)	3.03	1.33E-08	*
ABCA3	ATP-binding cassette, sub-family A (ABC1), member 3	3.48	1.39E-08	*
KLHL35	kelch-like 35 (Drosophila)	3.49	1.46E-08	*
KCNS1	potassium voltage-gated channel, delayed-rectifier, subfamily S, membe	3.58	1.56E-08	*
HMSD	histocompatibility (minor) serpin domain containing	3.18	3.11E-08	*
IL17REL	interleukin 17 receptor E-like	5.51	3.99E-08	*
MOCS1	molybdenum cofactor synthesis 1	-1.41	3.54E-02	
PRR5	microfibrillar-associated protein 4	-1.44	3.54E-02	
MFAP4	proline rich 5 (renal)	-2.02	3.54E-02	
GSTA4	glutathione S-transferase alpha 4	-1.20	3.59E-02	*
INHBB	inhibin, beta B	-1.32	3.59E-02	*
DKFZp434J0226	hypothetical LOC93429	-5.24	3.65E-02	
GADD45B	growth arrest and DNA-damage-inducible, beta	-1.73	3.70E-02	
IRX4	iroquois homeobox 4	-3.54	3.71E-02	
GEM	GTP binding protein overexpressed in skeletal muscle	-1.81	3.72E-02	
RGS11	regulator of G-protein signaling 11	-2.26	3.77E-02	
SMPD3	sphingomyelin phosphodiesterase 3, neutral membrane (neutral sphingo	-1.53	3.82E-02	*
CACNA1I	calcium channel, voltage-dependent, T type, alpha 1I subunit	-2.26	3.87E-02	
ADRA2A	adrenergic, alpha-2A-, receptor	-3.16	4.20E-02	
CES4A	carboxylesterase 4A	-1.99	4.21E-02	
ADCY4	adenylate cyclase 4	-1.42	4.42E-02	
FOXF1	forkhead box F1	-1.62	4.42E-02	
PALM	paralemmin	-1.93	4.42E-02	
LOC729178	hypothetical protein LOC729178	-1.80	4.46E-02	*
CWH43	cell wall biogenesis 43 C-terminal homolog (S. cerevisiae)	-4.96	4.54E-02	
TNFSF9	tumor necrosis factor (ligand) superfamily, member 9	-2.18	4.56E-02	
LGR6	leucine-rich repeat-containing G protein-coupled receptor 6	-3.13	4.62E-02	
KCTD15	potassium channel tetramerisation domain containing 15	-1.33	4.74E-02	
RPS26	ribosomal protein S26	-1.10	4.77E-02	
CHRD	chordin	-2.11	4.82E-02	
NOP16	NOP16 nucleolar protein homolog (yeast)	-1.12	4.92E-02	

B.3.1 A list of differentially expressed genes between HPV(+) and HPV(-) tumours identified by RNA-Seq analysis followed by correction for TIL number.

B.3.2 Gene ontology analysis of DEGs in HPV(+) vs HPV(-) tumours (TIL corrected data).

DEGs identified with a the FDR adjusted *p*-value <0.05 (i.e., *q*-value <0.05) and a fold change of >2 or <-2.

Genes expressed to a greater extent in HPV(+) compared to HPV(-) tumors				
	Gene Ontology term	Set Size	Candidates	*q*-value
GO:0007049	cell cycle	1576	48 (3.1%)	1.86E-10
GO:0022402	cell cycle process	1157	38 (3.3%)	4.49E-09
GO:0005694	chromosome	758	29 (3.8%)	2.06E-08
GO:0006259	DNA metabolic process	1001	34 (3.4%)	4.88E-08
GO:0000278	mitotic cell cycle	901	32 (3.6%)	4.88E-08
GO:0044427	chromosomal part	658	25 (3.8%)	4.24E-07
GO:0051301	cell division	735	26 (3.5%)	1.35E-06
GO:0000280	nuclear division	498	21 (4.2%)	3.22E-06
GO:0048285	organelle fission	528	21 (4.0%)	4.03E-06
GO:0006323	DNA packaging	208	13 (6.3%)	7.52E-06
GO:0070192	chromosome organization involved in meiosis	38	7 (18.4%)	8.36E-06
GO:0044770	cell cycle phase transition	469	19 (4.1%)	1.64E-05
GO:0044454	nuclear chromosome part	317	15 (4.7%)	1.85E-05
GO:0000228	nuclear chromosome	377	16 (4.2%)	2.18E-05
GO:0007126	meiotic nuclear division	167	11 (6.6%)	3.66E-05
GO:0051276	chromosome organization	867	25 (2.9%)	4.88E-05
GO:0051321	meiotic cell cycle	177	11 (6.2%)	4.88E-05
GO:0007129	synapsis	33	6 (18.2%)	5.31E-05
GO:0031497	chromatin assembly	174	11 (6.4%)	7.26E-05
GO:0000083	regulation of transcription involved in G1/S transition of mit	27	5 (18.5%)	0.000317
GO:0031981	nuclear lumen	2871	50 (1.7%)	0.000365
GO:0044428	nuclear part	3209	55 (1.7%)	0.000366
GO:0007127	meiosis I	76	7 (9.2%)	0.000469
GO:0048610	cellular process involved in reproduction	456	16 (3.5%)	0.000471
GO:0019953	sexual reproduction	679	20 (2.9%)	0.000471

Genes expressed to a lesser extent in HPV(+) compared to HPV(-) tumors				
	Gene Ontology term	Set Size	Candidates	*q*-value
GO:0043588	skin development	289	30 (10.4%)	4.79E-18
GO:0008544	epidermis development	289	26 (9.0%)	3.66E-14
GO:0031424	keratinization	46	12 (26.1%)	1.93E-11
GO:0060429	epithelium development	757	29 (3.8%)	8.03E-07
GO:0009888	tissue development	1441	40 (2.8%)	7.15E-06
GO:0048731	system development	3615	69 (1.9%)	8.88E-05
GO:0048513	organ development	2552	54 (2.1%)	8.88E-05
GO:0035878	nail development	5	3 (60.0%)	0.000706
GO:0007275	multicellular organismal development	4394	75 (1.7%)	0.00106
GO:0005882	intermediate filament	196	10 (5.2%)	0.00307
GO:0016616	oxidoreductase activity, acting on the CH-OH gro	116	8 (6.9%)	0.00313
GO:0005200	structural constituent of cytoskeleton	96	7 (7.3%)	0.00316
GO:0016614	oxidoreductase activity, acting on CH-OH group o	131	8 (6.1%)	0.00345
GO:0045550	geranylgeranyl reductase activity	2	2 (100.0%)	0.00345
GO:0048856	anatomical structure development	4248	72 (1.7%)	0.00356
GO:0017171	serine hydrolase activity	180	9 (5.0%)	0.00356
GO:0045095	keratin filament	99	7 (7.1%)	0.00382
GO:0005615	extracellular space	1234	29 (2.4%)	0.00397
GO:0030154	cell differentiation	3153	58 (1.8%)	0.0041
GO:0048869	cellular developmental process	3345	59 (1.8%)	0.00425
GO:0008236	serine-type peptidase activity	178	9 (5.1%)	0.00587
GO:0006629	lipid metabolic process	1226	28 (2.3%)	0.00718
GO:0005921	gap junction	31	4 (12.9%)	0.00871
GO:0044707	single-multicellular organism process	6199	93 (1.5%)	0.00926
GO:0035880	embryonic nail plate morphogenesis	2	2 (100.0%)	0.0101

B.3.2 Gene ontology analysis of DEGs in HPV(+) vs HPV(-) tumours (TIL corrected data).

B.3.3 Pathway analysis of DEGs in HPV(+) vs HPV(-) tumours (TIL corrected data)

Genes expressed to a greater extent in HPV(+) compared to HPV(-) tumors				
Pathway Name	**Set Size**	**Candidates**	**Pathway Source**	***q*-value**
Cell Cycle	476	26 (5.5%)	Reactome	8.75E-10
RB in Cancer	87	12 (13.8%)	Wikipathways	1.62E-08
Mitotic G1-G1/S phases	84	11 (13.1%)	Reactome	1.20E-07
G1 to S cell cycle control	68	10 (14.7%)	Wikipathways	1.67E-07
Cell Cycle, Mitotic	433	21 (4.8%)	Reactome	2.61E-07
Activation of the pre-replicative complex	30	7 (23.3%)	Reactome	1.12E-06
DNA Replication Pre-Initiation	35	7 (20.0%)	Reactome	2.65E-06
M/G1 Transition	35	7 (20.0%)	Reactome	2.65E-06
G1/S Transition	60	8 (13.3%)	Reactome	7.20E-06
DNA Replication	42	7 (16.7%)	Wikipathways	8.00E-06
Mitotic M-M/G1 phases	293	15 (5.1%)	Reactome	1.25E-05
Cell cycle	124	10 (8.1%)	KEGG	1.83E-05
DNA strand elongation	31	6 (19.4%)	Reactome	1.83E-05
E2F transcription factor network	75	8 (10.7%)	PID	2.67E-05
DNA replication	36	6 (16.7%)	KEGG	4.03E-05
DNA Replication	57	7 (12.3%)	Reactome	4.20E-05
ATR signaling pathway	37	6 (16.2%)	PID	4.20E-05
p53 signaling pathway	13	4 (30.8%)	BioCarta	0.000143
Cell Cycle	103	8 (7.8%)	Wikipathways	0.000217
Cell Cycle Checkpoints	75	7 (9.3%)	Reactome	0.000218
Oncogene Induced Senescence	30	5 (16.7%)	Reactome	0.000218
S Phase	77	7 (9.1%)	Reactome	0.000239
Synthesis of DNA	52	6 (11.5%)	Reactome	0.000239
Activation of ATR in response to replication stress	37	5 (13.5%)	Reactome	0.000533
TP53 Network	19	4 (21.1%)	Wikipathways	0.000533

Genes expressed to a lesser extent in HPV(+) compared to HPV(-) tumors				
Pathway Name	**Set Size**	**Candidates**	**Pathway Source**	***q*-value**
Antagonism of Activin by Follistatin	4	2 (50.0%)	Reactome	0.0809
ABCA transporters in lipid homeostasis	17	3 (17.6%)	Reactome	0.0809
The canonical retinoid cycle in rods (twilight vision)	21	3 (14.3%)	Reactome	0.0809
Arachidonic acid metabolism	50	4 (8.0%)	Reactome	0.0809
Ovarian steroidogenesis - Homo sapiens (human)	51	4 (7.8%)	KEGG	0.0809
Staphylococcus aureus infection - Homo sapiens (human	57	4 (7.3%)	KEGG	0.0809
Steroid hormone biosynthesis - Homo sapiens (human)	57	4 (7.0%)	KEGG	0.0809
Visual phototransduction	96	5 (5.2%)	Reactome	0.0809
Diseases associated with visual transduction	96	5 (5.2%)	Reactome	0.0809
Prostaglandin Synthesis and Regulation	30	3 (10.0%)	Wikipathways	0.0809
Oxidative Stress	30	3 (10.0%)	Wikipathways	0.0809
Arachidonic acid metabolism - Homo sapiens (human)	64	4 (6.2%)	KEGG	0.0928
ABC-family proteins mediated transport	36	3 (8.3%)	Reactome	0.103
Gap junction trafficking	12	2 (16.7%)	Reactome	0.103
Bile salt and organic anion SLC transporters	12	2 (16.7%)	Reactome	0.103
Platinum Pathway, Pharmacokinetics/Pharmacodynamic	13	2 (15.4%)	PharmGKB	0.105
Signaling by Activin	13	2 (15.4%)	Reactome	0.105
Vitamin A and Carotenoid Metabolism	40	3 (7.5%)	Wikipathways	0.105
classical complement pathway	14	2 (14.3%)	BioCarta	0.105
Gap junction trafficking and regulation	14	2 (14.3%)	Reactome	0.105

B.3.3 Pathway analysis of DEGs in HPV(+) vs HPV(-) tumours (TIL corrected data).

B.4 Lymphocyte marker genes for CD19, CD8 and CD4 cells.

B cells a	B cells b	B cells c	NK cells a	NK cells b	CD4+ T-cells a	CD4+ T-cells b	CD8+ T-cells
BACH2	KYNU	UBE2J1	ADAMTS1	NCR1	TRAT1	NELL2	CD8B
ZNF395	FUBP1	VPREB3	AKR1C3	NMUR1	ICOS	MYBL1	TBX21
ARHGAP17	ADAM19	BLNK	ASCL2	OSBPL5	TNFRSF25	SH2D1A	CST7
IFT57	STX7	TCL1A	BNC2	PDGFD	MAL	LCP2	GZMH
TSPAN13	TFEB	SPIB	BZRAP1	PDGFRB	DPP4	LCK	ADRB2
CNTNAP2	TPD52	POU2AF1	CCL3	PDZD4	AQP3	KLRB1	CD8A
OPN3	SYK	PNOC	CD160	PLEKHF1	BCL11B	ITPKB	PRF1
GGA2	SNX2	CD79B	CD247	PRSS23	GPSM3	ITK	IL2RB
FAIM3	SHMT2	CD79A	CHST12	PRSS30P	LEF1	ITGA6	DKK3
KMO	RNASE6	CD19	CMKLR1	PTGDR	LEPROTL1	INPP4A	
SYPL1	RFX5	BLK	COL13A1	PTGDS	SYNE2	IL7R	
ST6GAL1	PPP3CA	TNFRSF17	COLQ	RNF165	KLRK1	DNAJB1	
PTPRK	POU2F2	TMEM156	CX3CR1	S1PR5	AAK1	GZMK	
ABCB4	PMAIP1	PPAPDC1B	FASLG	SGSM1	SEMA4D	GBP2	
NT5E	PLCG2	FCRLA	FEZ1	SH2D1B	TACC3	GBP1	
GPR18	PIK3C2B	CPNE5	FGFBP2	SIGLEC17P	RASGRP1	GATA3	
FCER2	ODC1	FAM129C	GK5	SPON2	SPOCK2	GALT	
CSNK1G3	NUP88	FCRL5	GOLGA8T	SPTSSB	LPIN2	FYN	
CCR6	NCF4	FCRL1	GZMB	TKTL1	IFITM1	FYB	
CD40	MEF2C	KIAA0125	HOPX	TTC38	ZAP70	FLT3LG	
CD37	SMAD3	PAX5	IGFBP7	ZBTB16	UPP1	CTSW	
CXCR5	LRMP	FCRL2	IL12RB2	CD244	TXK	CDC25B	
NAPSB	LGALS9	EAF2	KIR2DL1	IVNS1ABP	TNFAIP3	CD28	
SLC2A5	IGL@	QRSL1	KIR3DL2	YPEL1	TIAM1	CD6	
CD83	IGKC	STAP1	KLRC3	KIR2DL2	TCF7	CD3G	
E2F5	IGHM	SEL1L3	KLRD1	KIR2DS5	STAT4	CD3E	
SSPN	IGHG1	EZR	LAIR2	KIR2DS1	SORL1	CD3D	
ZCCHC7	IGHD	GNG7	LDB2	KIR3DL3	SELPLG	CD2	
ZNF154	HLA-DRB6	MS4A1	LIM2	TGFBR1	SATB1	CCND2	
GLDC	HLA-DRA	METTL21D	LINGO2	CLIC3	RORA	APBA2	
PRKCE	HLA-DPB1	TCF4	LOC401321	KIR2DL3	RGS10	IL6ST	
RAB30	HLA-DPA1	CD200	LOC727787	KIR2DL4	RBMS1	MAN1C1	
HLA-DOA	HLA-DMB	ITPR1	MGC24103	KIR3DL1	RARRES3	FHIT	
CYBASC3	HLA-DMA	IRF4	MYOM2	KLRF1	PRKCQ	INPP4B	
SETBP1	HHEX	IL4R	NCAM1		PIK3R1	PRKCA	
TEAD2	GM2A	IGJ					
BTLA	FCGR2B	IRF8					
TLR10	CYBB	HLA-DQB1					
RALGPS2	CTSH	HLA-DQA1					
PLEKHF2	CD74	HLA-DOB					
SAV1	CCNG2	CR2					
POLD4	BTK	CD72					
BCL11A	BCL7A	CD22					
SWAP70	RHOH	CD1C					
SP140	BIRC3	ADK					
ADAM28	ANXA4	KLHL14					
STAG3	ALOX5	CD24					
AIM2	OSBPL10	HIST1H2BK					
UVRAG	BANK1	WDR11					
FAM46C	RRAS2	SIDT2					
FKBP11	CORO1A	SNX10					
EIF2AK3	MTSS1	BACE2					
KLF6	LY86	NGLY1					

B.4 Marker gene sets whose expression was associated with the different lymphocyte cell subsets (B-cells, NK cells and CD4+ and CD8+ T-cells)

Gene sets obtained from the following resources: CTen, IRIS, HeamAtlas, Palmer et al. Grigoryev et al. and Whitney et al.

B.5 Advanta Immuno-Oncology gene expression assay

B.5.1 IO gene expression panel

Panel A (PN 101-7403)

	1	2	3	4	5	6	7	8	9	10	11	12
A	GAPDH*	TFRC	VEGFA	EOMES	GZMA	GZMB	CCL22	ITGAM	TMEM55B	TNF	CD1C	TNFRSF4
B	ACTB	GUSB	MS4A1	LAG3	PTPRC	IDO1	CCL28	ITGAX	VPS33B	FOXP3	CD70	ARG1
C	B2M	CCL2	CD244	PDCD1LG2	KLRK1	IL7R	CCR5	LGALS9	RORC	IL10	CLEC4C	CSF2
D	HLA-C	HMOX1	CD274	TNFRSF14	PRF1	CXCL9	CD40	MAP4K1	MICA†	CCR7	IL2RA	IL17A
E	HLA-B	NCAM1	CD276	VTCN1	CD3E	CX3CL1	CD40LG	PTGER2	BTLA	CD27	TNFRSF9	IL17F
F	HLA-A	EPCAM	CD69	CD86	CD4	CXCL10	ITGB2	PTGER4	CD28	CXCR3	IL7	IL2
G	CXCL8 (IL8)	SDHA	CD80	CD48	IL12A	ICAM1	IL13	PTGS2	ICOS	IFNG	MICB†	IL4
H	HAVCR2**	VCAM1	CTLA4	TGFB1	CD8A	IL1B	IL6	SP2	TNFSF4	TBX21	PDCD1	TNFSF9

	1	2	3	4	5	6	7	8	9	10	11	12
A	GAPDH*	TFRC	LAPTM5	STAT1	CD22	CD2	IL2RG	PYGL	STAT2	EBI3		
B	ACTB	GUSB	NKG7*	CYBB	IGLJ3*	CD37	SLAMF7	LCK	STAT5A	FASLG		
C	B2M	IGKC*	DGAT2	IRF9	CD160	CD53	SLAMF8	GZMH	STAT6	IFNA2*		
D	IGHA1	HLA-DPB1	CD63	STAT3	NT5E	CXCR4	TNFAIP8	ISG15*	TLR7	CD19		
E	HLA-DRB1	CCL21	GNLY	STAT5B	CCL3	GZMK	FYB	IFIT2	APOBEC3B	TNFRSF18		
F	NRAS	CD52	IL15	ARG2	CCL4	HLA-DMB	CA4	CD1D	FCRLA	TLR8		
G	IGHM	CTSS	CCL5	ERBB2	HLA-DQB1	IGSF6	KREMEN1	GATA3	IGHG1	IL12B		
H	JCHAIN (IGJ)	FCER1G	IFI27	APOBEC3A	CCL18	IL10RA	LRG1	JAK2	TNFSF18			

Reference genes

Open wells for user customization

* Assays that can detect gDNA.

**Assays that may also detect gDNA when gDNA is present in the RT-preamplification reaction.

† The **MICA** assay can also detect the MICB gene.

B.5.1. Immuno-Oncology gene expression panel (Fluidigm)

Details of the investigative genes and control genes in panel A and B of the gene expression assay. Panel A = 91 genes, Panel B =74 genes + 17 and Control = 5 genes.

B.5.2 IO gene expression protocol version and details

FLUIDIGM

PN 101-7087 B1
PROTOCOL

Advanta IO Gene Expression Assay Using the GE 24.192 IFC

Contents

Materials

Required Kit

For maps of the assay plates, see Appendix A. For a list of the available bundle kits, see Appendix B. For a list of the reagent kit components, see Appendix C.

IMPORTANT Store reagents as soon as they are received, according to manufacturer's storage recommendations.

Kit	Source	Part Number	Storage
Advanta IO Gene Expression Assay—Panels A&B, GE 24.192, 2 IFCs	Fluidigm	101-7678	
Kit contains:			
Advanta IO Gene Expression Reagent Kit—GE 24.192, 2 IFCs			–20 °C
Advanta IO Gene Expression Assay Plate–Panel A			Room temperature
Advanta IO Gene Expression Assay Plate–Panel B			Room temperature
Control Line Fluid 192.24/24.192 for 2 IFCs			Room temperature
24.192 Dynamic Array™ IFC for Gene Expression			Room temperature

Suggested Reagent

Kit	Source	Part Number	Storage
Advanta IO Gene Expression Control 1	Fluidigm	101-7676	–20 °C

Required Consumables

Product	Source	Part Number
PCR-compatible clear adhesive film*	Major laboratory supplier (MLS)	–
1.5 mL microtubes	MLS	–
25-mL reagent reservoir†	MLS	–
96-well PCR plates	MLS	–
8-well PCR tube strips	MLS	–
Filtered pipette tips	MLS	–
Clear tape	MLS	–

* Recommended: MicroAmp® Clear Adhesive Film (Thermo Fisher Scientific PN 4306311).

† Recommended: Accuflow 25-mL Reagent Reservoir, Certified Sterile (Accuflow PN D620025)

B.5.2. Immuno-Oncology gene expression panel (Fluidigm) protocol version.

Details of the Kit components and protocol version used for the qPCR assay.

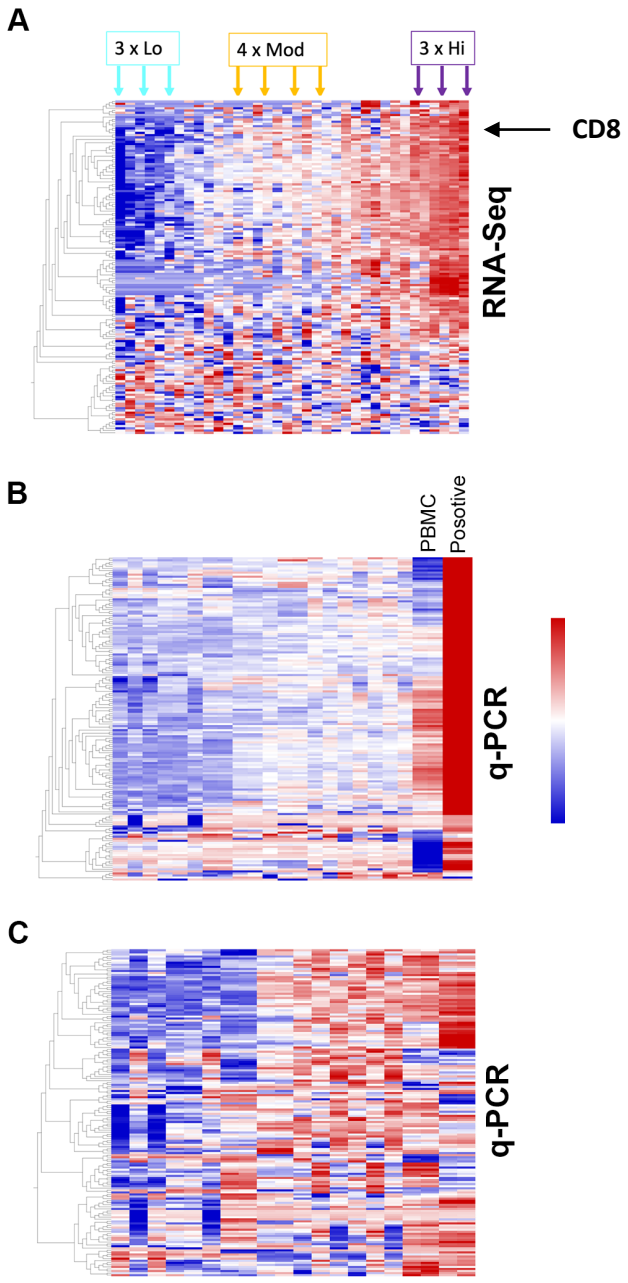
B.5.3 Immunological and tumour microenvironment gene markers

Immune marker genes	Stromal marker genes	Glycolytic marker genes
PTPRC	ACTA2	SLC2A1
CD3E	POSTN	ALDOA
CD4	COL1A1	PKM2
CD8A		
ITGAE		
MS4A1		
SELL		
FOXP3		
IL10		
IL2RA		
CD86		
GZMA		
IFNG		
PRF1		
TNFRSF9		
PDCD1		
CD274		
PDCD1LG2		
HAVCR2		
LAG3		
CTLA4		
TIGIT		
ICOS		
CD27		

B.5.3. Immunological and tumour microenvironment gene markers.

A curated list of immune gene markers, stromal gene markers and glycolytic gene markers visualised in the qPCR data and RNA-Seq data.

B.5.4 Immuno-oncology gene expression panel A and B visualised in the Southampton HNSCC RNA-Seq data.



B.5.4. Immuno-oncology gene expression panel A and B visualised in the Southampton HNSCC RNA-Seq data.

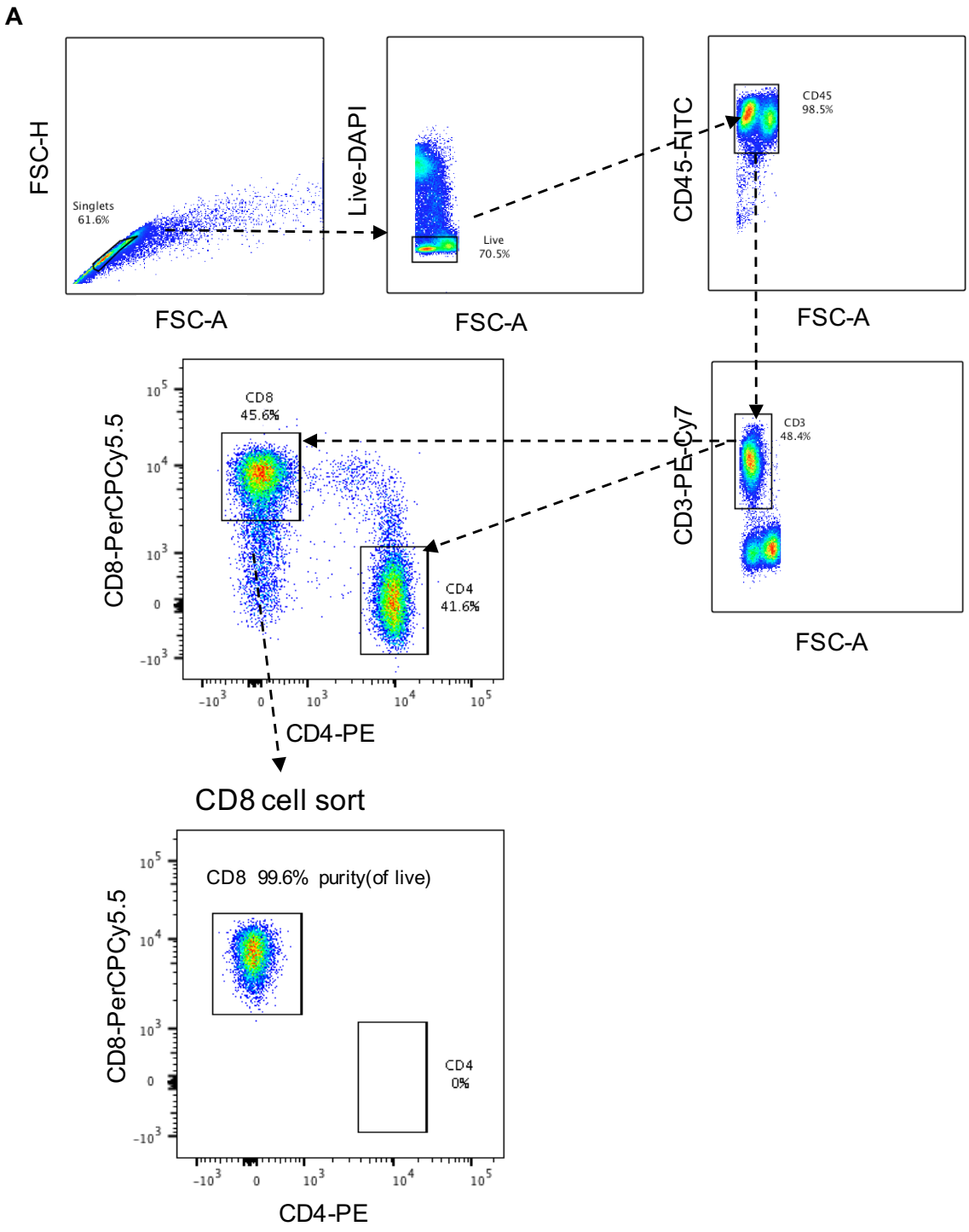
Heatmap visualisation (sample order 1st PCA) of the Immuno-oncology gene expression Panels A and B (panel A = 91 genes, panel B = 74 genes). (A) Heatmap visualisation of the Panels A and B in our HNSCC RNA-Seq data. The top cluster consists of the immune genes. The bottom cluster of genes are associated with the oncology markers. The panels capture the immune density across the samples n=35. (B) qPCR gene expression data for a subset of HNSCC cases (n=10) that display a range of immune densities. The gene expression data is derived from qPCR data (ΔΔct[GOI]) for the Immuno-oncology panels A and B. A positive control and PBMC sample show high expression of the immune gene markers, this results in a loss of resolution in the tumours samples. (C) This shows the overall expression of gene panel A and B (excluding the control samples), the duplicate samples are highly consistent and a clear clustering of two sample groups (Immune “hot” and “cold”) was now present.

Appendix C

C.1 Immune cell sorting optimisation

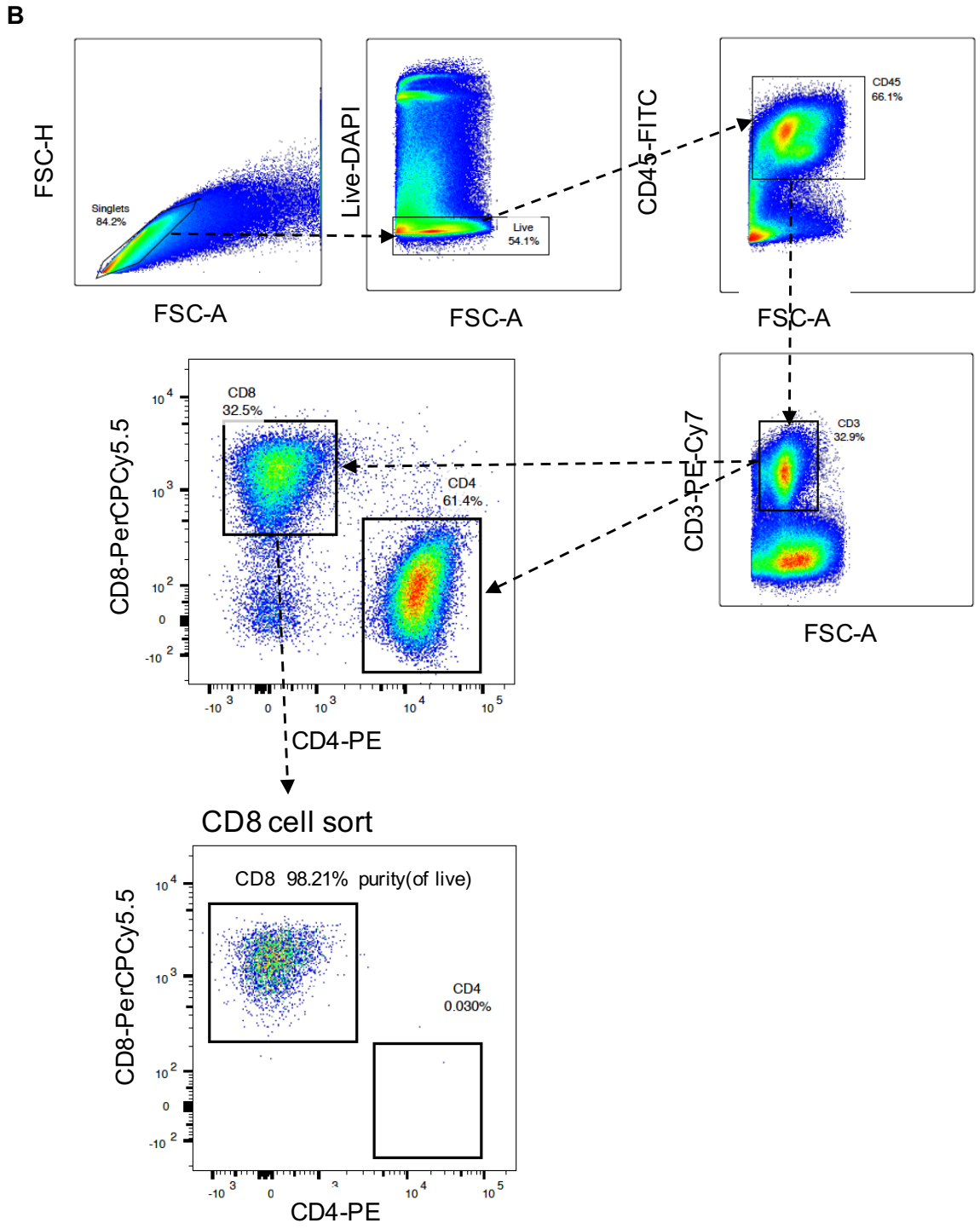
C.1.1 Optimisation of Immune cell sorting strategy and panel 1

PBMC



Optimisation of Immune cell sorting strategy and panel 1.

Tumour



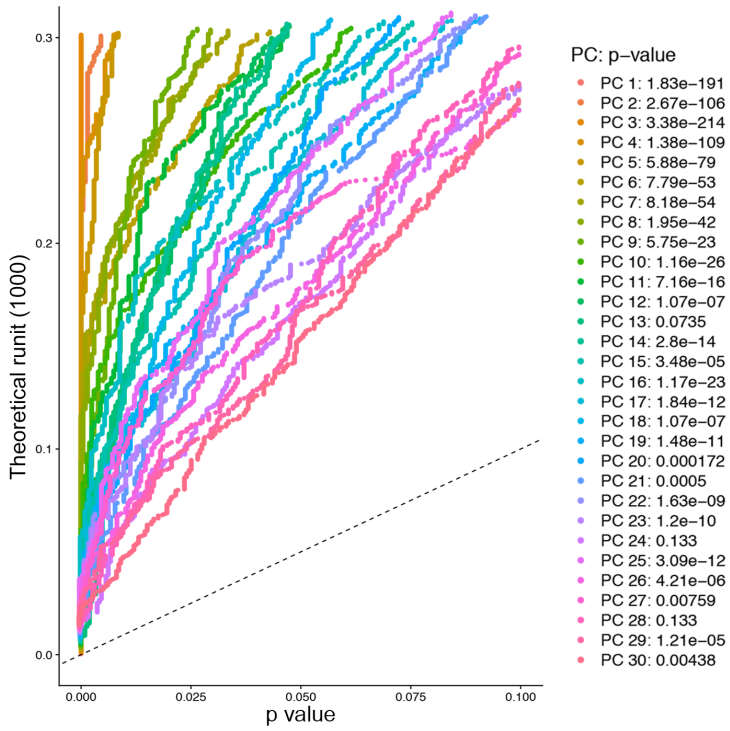
C.1.1. Optimisation of Immune cell sorting strategy and panel 1.

(A) PBMC and (B) HNSCC display an immune cell sorting panel that allows CD4+ and CD8+ cells to be isolated from tumours at very high purity (>98%).

Appendix D

D.1 Single cell B lymphocyte dimensionality reduction

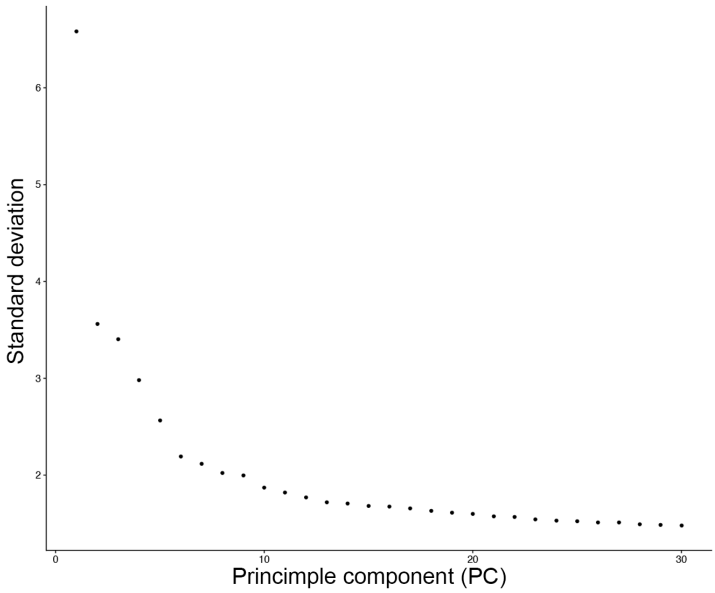
D.1.1 Jackstraw plot



D.1.1. Single cell dimensionality reduction using PCA.

The Jackstraw plot projects the statistical significance of the principle components (PC) against the theoretical runit.

D.1.2 Elbow plot

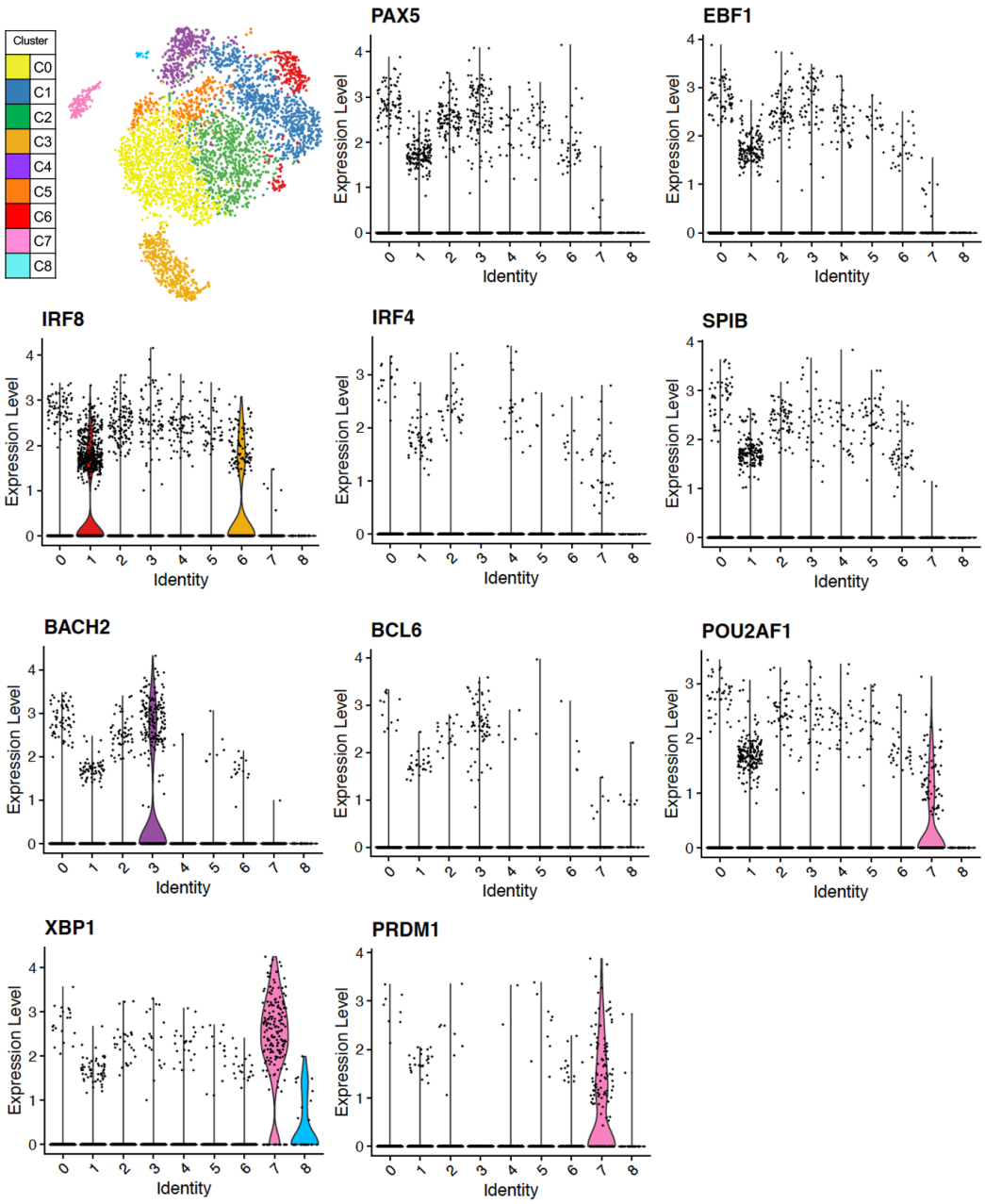


D.1.2. Single cell dimensionality reduction using PCA.

An Elbow plot represents the standard deviation for each PC 1-30. The PCs are selected based on the p<0.05 and where the “Elbow” begins to flatten (PC1-12).

D.2 B cell lineage transcription factors

D.2.1 Violin plots of key B cell transcription factors

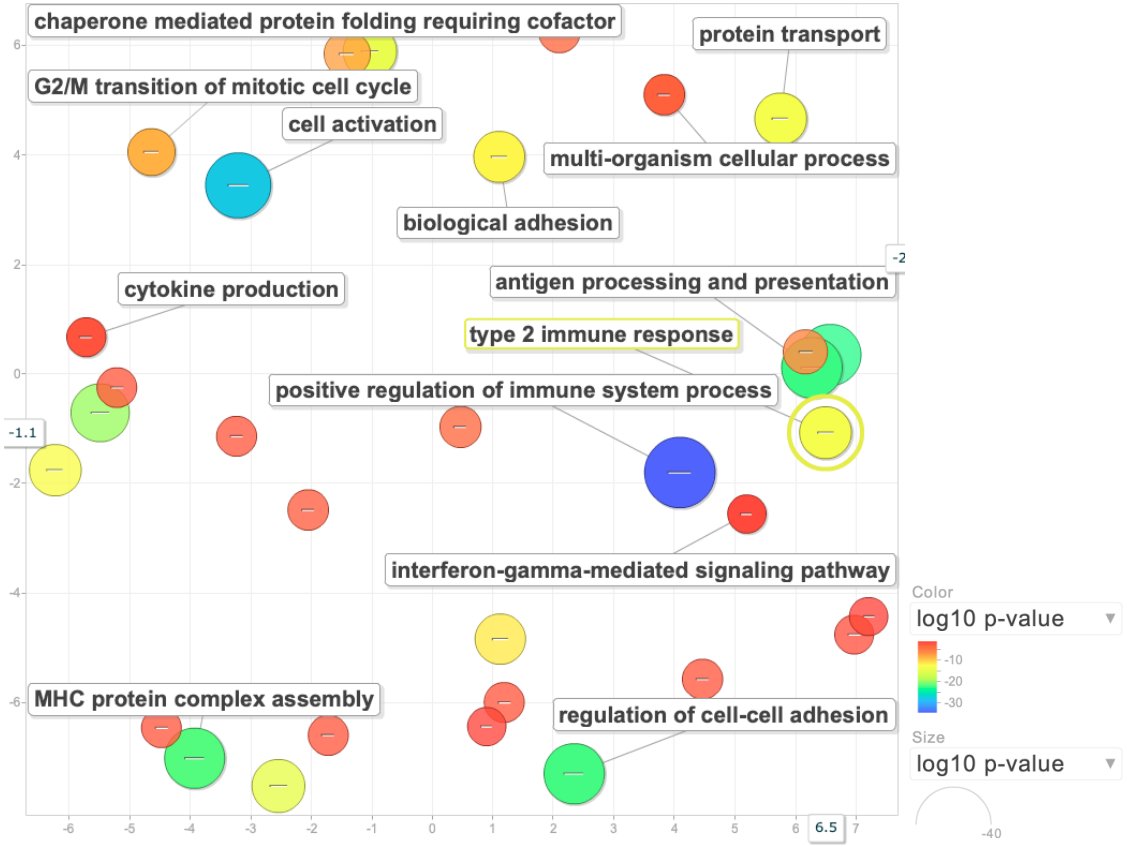


D.2.1. Violin plots of key B cell lineage transcription factors in clusters C0 to C8.

Seurat normalised expression level for each gene as a violin plot. The B cell transcription factors assist in allocating B cell subtypes (Naïve, PAX5, EBF1; Activated, BACH2, IRF4^{lo}, IRF8; Germinal centre, BCL6, BACH2, PAX5; Memory cells, PAX5, POUF2AF1 (OBF1), SPIB; Plasmablast and plasma cell, IRF4, BLIMP1, XBP1 (Basso and Dalla-Favera, 2015, Nutt et al., 2015).

D.3 Gene ontology analysis of B cell clusters

D.3.1 Cluster LN GC (Lymph node germinal centre B cells) REVIGO reduced gene ontology



D3.1. REVIGO gene ontology reduction for LN GC (Lymph node germinal centre B cells).
Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

LN GC (Lymph node germinal centre B cells)-Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
Histocompatibility complex|C1-set domain containing	2.32E-06	8	42	HLA-DMA,HLA-DMB,HLA-DOB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5
RNA binding motif containing	7.99E-06	14	213	SRSF10,HNRNPD,HNRNPH1,U2AF1,TRA2A,SFPQ,SRSF3,SRSF5,RBM23,RBM39,SLTM,RBM6,SPEN,SRSF9
Actin related protein 2/3 complex |WD repeat domain containing	7.45E-05	4	9	ACTR3,ACTR2,ARPC2,ARPC5L
Canonical high mobility group	1.53E-04	4	11	HMGB1,HMGB2,HMGN1,HMGN2
RNA binding motif containing|Serine and arginine rich splicing factors	1.90E-04	4	12	SRSF10,SRSF3,SRSF5,SRSF9
Zinc fingers C2HC-type|PHD finger proteins|Lysine acetyltransferases	1.63E-03	7	90	ASH1L,BPTF,KMT2C,KMT2A,MTF2,KDM5A,PHF3
DEAD-box helicases	2.09E-03	5	42	DDX3X,DDX5,DDX6,DDX24,DDX17
AT-rich interaction domain containing	1.01E-02	3	15	ARID4B,ARID5B,KDM5A
Ubiquitin conjugating enzymes E2	1.80E-02	4	41	UBE2D2,UBE2D3,UBE2L3,UBE2J1
CD molecules|Tumor necrosis factor superfamily	2.64E-02	12	394	EVI2B,ITGAE,PTPRC,MS4A1,CD22,CD86,CD37,CD53,CD74,CD79B,CD81,IL4R

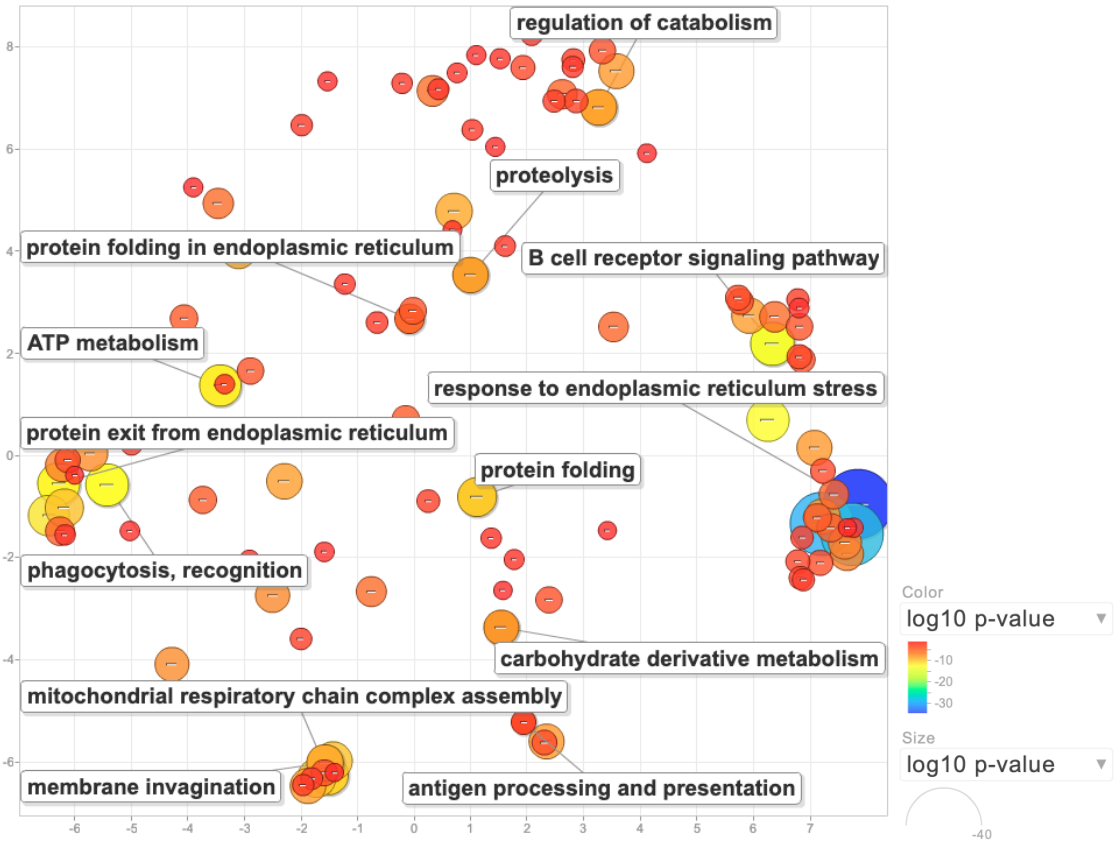
LN GC (Lymph node germinal centre B cells)-Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
immune response	6.14E-08	62	1572	PAG1,SWAP70,HLA-DMA,HLA-DMB,GAPDH,HLA-DOB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,NCF1,ACTB,ETS1,SNAP23,HMGB1,HMGB2,BCL6,VPREB3,MEF2C,STAP1,DDX3X,GCSAM,RFTN1,UBE2D2,UBE2D3,PTPRC,SUMO1,STX7,AIM2,SKAP2,RAC1,HSP90AA1,SEC14L1,KLHL6,EZR,WIPF1,CALM2,CAMK2D,DBNL,CAPZA1,POU2F2,REL,ST6GAL1,RGS1,ACTR3,ACTR2,ARPC2,ARF6,MS4A1,EGR1,CD86,ATG12,CD37,CD74,ELF1,CD79B,CD81,LSM14A,PAX5,PUM1,IL4R
antigen processing and presentation of exogenous peptide antigen via MHC class II	3.06E-04	10	94	HLA-DMA,HLA-DMB,HLA-DOB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,CD74,DYNLL1
cell activation	3.54E-04	37	1001	PAG1,SWAP70,HLA-DMA,HLA-DMB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,ACTB,SNAP23,HMGB1,BCL11A,BCL6,MEF2C,STAP1,BLOC1S6,PTPRC,ZFP36L1,SKAP2,RAC1,HSP90AA1,POU2F2,LMO4,MYL12A,RHOH,MS4A1,EGR1,CD86,CD37,PRKAR1A,CD74,CD79B,CD81,ATM,IL4R,TPD52
cell activation	3.54E-04	37	1001	PAG1,SWAP70,HLA-DMA,HLA-DMB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,ACTB,SNAP23,HMGB1,BCL11A,BCL6,MEF2C,STAP1,BLOC1S6,PTPRC,ZFP36L1,SKAP2,RAC1,HSP90AA1,POU2F2,LMO4,MYL12A,RHOH,MS4A1,EGR1,CD86,CD37,PRKAR1A,CD74,CD79B,CD81,ATM,IL4R,TPD52
B cell activation	1.30E-03	15	257	SWAP70,HLA-DQB1,BCL11A,BCL6,MEF2C,PTPRC,SKAP2,POU2F2,MS4A1,CD86,CD74,CD79B,CD81,ATM,TPD52
cell cycle	1.45E-03	52	1766	MIS18BP1,ERH,ATRX,PPP1R12A,ETS1,ZNF207,RABGAP1,FOXN3,MDM4,DDX3X,WAC,PTPRC,UBE2L3,CCNL1,SCIMP,MBD4,CCNI,HSP90AA1,RAD21,EZR,RAN,RANBP2,SFPQ,SRSF5,OFD1,CALM2,RB1,RBBP4,LAMTOR5,CAMK2D,CNTRL,SYF2,CHORDC1,YWHAE,PRPF40A,ACTR3,ACTR2,WTAP,RPS27L,PPP1CB,ARF6,PPP2R5C,KHDRBS1,PEBP1,CDK13,PRKAR1A,DYNLL1,ATM,PUM1,LSM10,STAG3,PCM1
regulation of cell-cell adhesion	1.52E-03	20	430	MYADM,PAG1,SWAP70,HLA-DMA,HLA-DMB,HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,ETS1,HMGB1,BCL6,PTPRC,RAC1,HSP90AA1,CD86,PRKAR1A,CD74,IL4R
B cell receptor signaling pathway	3.01E-03	7	65	MEF2C,STAP1,GCSAM,RFTN1,PTPRC,KLHL6,CD79B
regulation of B cell proliferation	9.10E-03	6	59	BCL6,MEF2C,PTPRC,CD74,CD81,ATM
positive regulation of B cell proliferation	9.31E-03	5	39	BCL6,MEF2C,PTPRC,CD74,CD81
lymphocyte costimulation	9.41E-03	7	83	HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,RAC1,CD86
interferon-gamma-mediated signaling pathway	1.05E-02	7	85	HLA-DQA1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,SUMO1,CAMK2D
cellular response to endogenous stimulus	1.27E-02	37	1278	NCOA3,SETX,MEF2C,HNRNPD,HNRNPH1,DDX5,IQGAP1,BPTF,PARP1,SULF2,UBE2D3,UBE2L3,UCP2,NUCKS1,TAF7,GNG7,DNAJA1,THRAP3,DNMT1,EZR,RAN,RAP1B,DDX17,SRSF3,SRSF5,OFD1,RB1,LAMTOR5,POLR2K,ATP6V1G1,SHOC2,ZNF106,EGR1,PRKAR1A,H2AFZ,CD81,ACAP2
regulation of type 2 immune response	1.76E-02	4	28	BCL6,CD86,CD74,IL4R
regulation of alpha-beta T cell activation	3.67E-02	6	86	HLA-DQB1,HMGB1,BCL6,PTPRC,CD86,IL4R

D.3.1. Gene ontology analysis of B cell cluster LN GC (Lymph node germinal centre B cells).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in LN GC. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.2 REVIGO gene ontology reduction for PB (Plasmablast) cluster



D.3.2. REVIGO gene ontology reduction for PB (Plasmablast) cluster.

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

PB (Plasmablast cluster) -Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
NADH:ubiquinone oxidoreductase supernumerary subunits	1.25E-12	11	30	NDUFA1,NDUFA2,NDUFA3,NDUFA6,NDUFAB1,NDUFB4,NDUFB7,NDUFB10,NDUFC2,NDUFS6,NDUFB11
Oligosaccharyltransferase complex subunits	5.65E-12	8	12	MLEC,RPN1,RPN2,OST4,DAD1,DDOST,OSTC,KRTCAP2
Peroxiredoxins	8.92E-06	4	6	PRDX5,PRDX4,PRDX2,PRDX1
Heat shock 70kDa proteins	2.72E-05	5	17	HSPA1A,HSPA1B,HSPA5,HSPA6,HYOU1
Transmembrane p24 trafficking proteins	7.23E-05	4	10	TMED2,TMED10,TMED4,TMED9
F-type ATPases|Mitochondrial complex V: ATP synthase subunits	8.20E-04	4	18	ATP5MC1,ATP5MC3,ATP5PF,ATP5MF
F-type ATPases|Mitochondrial complex V: ATP synthase subunits	8.82E-04	4	19	ATP5MC1,ATP5MC3,ATP5PF,ATP5MF
Protein disulfide isomerases	1.17E-03	4	21	PDIA3,PDIA4,PDIA6,P4HB
CD molecules|Tumor necrosis factor superfamily	1.74E-03	14	394	ITGA6,TNFRSF17,BSG,SDC1,ICAM2,SLC44A1,CD27,GYPC,CD38,CD59,CD63,CD79A,FCRL5,SLAMF7
DNAJ (HSP40) heat shock proteins	2.29E-03	5	49	DNAJB11,DNAJB9,DNAJC1,SEC63,DNAJC3
Basic leucine zipper proteins	2.29E-03	5	49	JUN,CREB3L2,ATF5,XBP1,ATF4
ARF GTPase family	3.77E-03	4	31	ARF1,ARF4,ARL1,SAR1B
Rhomboid family	4.08E-03	3	14	DERL3,RHBDD1,DERL1
Immunoglobulin lambda locus at 22q11.2	4.08E-03	6	89	IGLV6-57,IGLV3-1,IGLV2-14,IGLC7,IGLC2,IGLC3
Immunoglobulin heavy locus at 14q32.33	9.44E-03	8	187	IGHV1-69,IGHA1,IGHA2,IGHG1,IGHG2,IGHG3,IGHG4,IGHM
Proteasome	1.00E-02	4	43	PSMB5,PSMB6,PSMD8,PSME2
CD molecules|Tumor necrosis factor receptor superfamily|Death inducing signaling complex	1.25E-02	2	6	CFLAR,CASP10
MRH domain containing	1.25E-02	2	6	ERLEC1,OS9
Death effector domain containing|Ripoptosome|Death inducing signaling complex	1.64E-02	2	7	CFLAR,CASP10
Blood group antigens	4.94E-02	3	37	BSG,AQP3,GYPC

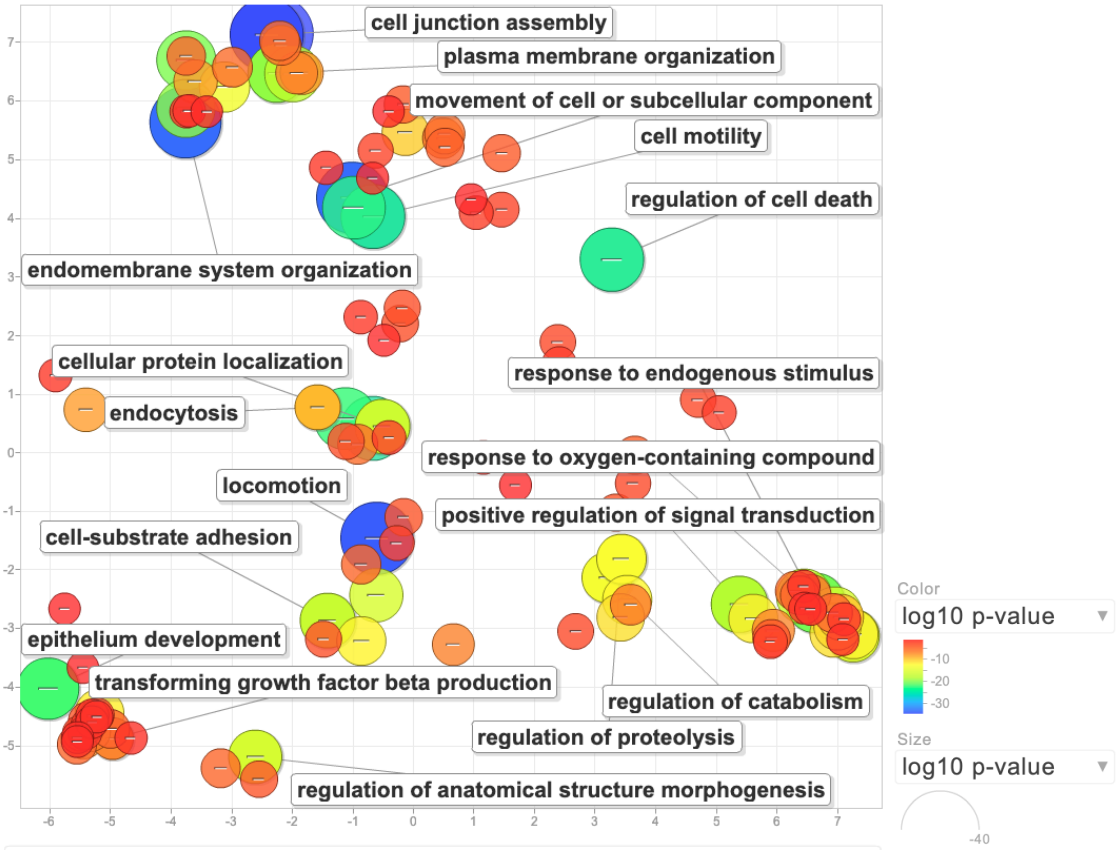
PB (Plasmablast cluster) -Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
B cell activation	7.98E-08	21	257	MZB1,IGLL5,PRDM1,IGLC7,MIF,CASP3,XBP1,LGALS1,IGHA1,IGHA2,CD27,IGHG1,IGHG2,IGHG3,IGHG4,IGHM,CD38,IGKC,CD79A,IGLC2,IGLC3
B cell mediated immunity	6.16E-07	15	146	IRF7,IGLL5,IGLC7,IGKV4-1,IGHA1,IGHA2,CD27,IGHG1,IGHG2,IGHG3,IGHG4,IGHM,IGKC,IGLC2,IGLC3
B cell proliferation	9.42E-03	7	93	MZB1,PRDM1,MIF,CASP3,CD27,CD38,CD79A
B cell receptor signaling pathway	3.59E-11	15	65	IGLL5,IGLC7,IGHA1,IGHA2,IGHG1,IGHG2,IGHG3,IGHG4,IGHM,CD38,IGKC,CD79A,IGLC2,IGLC3,FCRL5
mitochondrial respiratory chain complex assembly	3.14E-09	14	77	PET100,NDUFA1,NDUFA2,NDUFA3,NDUFA6,NDUFAB1,NDUFB4,NDUFB7,NDUFB10,NDUFC2,NDUFS6,NDUFS8,NDUFB11,UQCC2
antigen processing and presentation	5.10E-04	14	232	HLA-C,SEC24A,PSMB5,PSMB6,PSMD8,PSME2,SEC24D,SEC13,CALR,CANX,PDIA3,ARF1,SAR1B,CYBA
protein exit from endoplasmic reticulum	1.05E-11	13	39	HSP90B1,HM13,ERLEC1,DERL3,SEC61B,OS9,SEL1L,UBE2J1,RHBDD1,DERL1,LMAN1,TMED9,HERPUD1
protein folding	4.99E-09	22	236	DNAJB11,HSP90B1,MLEC,FKBP11,TXN,DNAJC1,HSPA1A,HSPA1B,HSPA5,FKBP2,HSPA6,CALR,PRDX4,CANX,PPIB,PDIA3,TXNDC11,SIL1,PDIA4,PDIA6,LMAN1,P4HB
response to type I interferon	6.39E-03	7	86	HLA-C,IRF7,ISG20,ZBP1,IFNAR2,ISG15,IFI6
response to unfolded protein	7.27E-24	34	174	DNAJB11,HSP90B1,ERN1,ACADVL,SRPRA,SSR1,DNAJB9,SERP1,SRPRB,UBXN4,DERL3,MANF,SEC61B,SDF2L1,SEC61A1,HSPA1A,HSPA5,HSPA6,HSPB1,CREB3L2,RHBDD1,HYOU1,DERL1,CALR,XBP1,TMBIM6,PREB,PDIA6,SEC61G,ATF4,EIF2AK4,SEC63,DNAJC3,HERPUD1
immunoglobulin mediated immune response	4.76E-07	15	143	IRF7,IGLL5,IGLC7,IGKV4-1,IGHA1,IGHA2,CD27,IGHG1,IGHG2,IGHG3,IGHG4,IGHM,IGKC,IGLC2,IGLC3
leukocyte mediated immunity	4.22E-04	18	356	IRF7,IGLL5,CADM1,IGLC7,IGKV4-1,IGHA1,IGHA2,CD27,IGHG1,IGHG2,IGHG3,IGHG4,IGHM,IGKC,PRDX1,IGLC2,IGLC3,SLAMF7
autophagy	1.55E-03	23	590	ATP6V0B,PSAP,TP53INP1,ERN1,PDK1,DAP,UBC,HAX1,SCFD1,HSPB1,PIM2,CISD2,TMEM59,CASP3,XBP1,FIS1,LARP1B,TMBIM6,ZNF593,TMEM208,EIF2AK4,MTDH,FBXW7
tumor necrosis factor-mediated signaling pathway	1.91E-04	12	155	TRADD,PSMB5,PSMB6,PSMD8,PSME2,TNFRSF17,UBC,HSPA1A,HSPA1B,TXNDC17,GSTP1,CD27
proteolysis	4.54E-08	62	1697	TRADD,HSP90B1,SPCS2,PSMB5,PSMB6,DAP,SEC11C,PSMD8,PSME2,DNAJB9,HM13,MDM2,ERLEC1,CFLAR,SPCS3,UBC,UBXN4,IGLC7,ATRAID,DERL3,SEC61B,SDF2L1,OS9,AURKAIP1,DNAJC1,PRDX5,USP48,IGKV4-1,HSPA5,RBX1,SEL1L,UBE2J1,RHBDD1,DERL1,NRDC,SPCS1,TMEM59,CASP3,XBP1,CASP10,FIS1,CAV1,PDIA3,LAP3,RPS27L,ISG15,TIMP1,CD27,IGHG1,TMEM208,IGHG2,IGHG3,IGHG4,IGKC,CD59,IGLC2,IGLC3,IFI6,DNAJC3,HERPUD1,TRIB1,FBXW7
regulation of catabolic process	2.65E-08	39	777	ATP6V0B,PSAP,TP53INP1,ERN1,PSMB5,PSMB6,DAP,PSMD8,PSME2,MDM2,ERLEC1,UBC,BTG2,ATRAID,DERL3,MIF,OS9,HAX1,SCFD1,HSPB1,UBE2J1,DDIT4,RHBDD1,PIM2,NRDC,CISD2,ANXA2,TMEM59,CASP3,XBP1,RDX,UQCC2,CAV1,TIMP1,EIF2AK4,HERPUD1,TRIB1,MTDH,FBXW7

D.3.2. Gene ontology analysis of B cell cluster PB (Plasmablast).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in PB. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.3 REVIGO gene ontology reduction for PC (Plasma cell)



D.3.3. REVIGO gene ontology reduction for PC (Plasma cell).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

PC (Plasma cell cluster)-Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
CD molecules|Tumor necrosis factor superfamily	1.73E-06	39	394	ERBB2,CD46,F3,TNFRSF12A,ALCAM,SDC1,FUT3,MUC1,CEACAM6,TNFRSF10B,TNFRSF21,NRP1,BMPR1B,SLC44A1,CD9,CD47,TREM1,CD59,CD63,CD151,TNFRSF1A,CDH1,PLAUR,ICAM1,CDCP1,IL1R1,IL13RA1,INSR,CD55,ITGA1,ITGA2,ITGA3,ITGAV,ITGB1,ITGB4,KDR,LAMP2,BCAM,EPCAM
S100 calcium binding proteins|EF-hand domain containing	4.63E-05	8	21	S100A14,S100A2,S100A6,S100A9,S100A10,S100A11,S100A13,S100A16
EF-hand domain containing	8.38E-05	24	219	EHD2,S100A2,S100A6,S100A9,S100A11,CIB1,RASEF,SRI,TBC1D9,DST,CALM2,CALU,CAPN1,CAPN2,CETN2,FKBP9,GCA,PLS3,S100A16,PRKCSH,MYL12B,CAPN8,PDCD6,RCN1
Plakins	2.09E-04	5	8	EVPL,DST,PLEC,PPL,DSP
LIM domain containing	3.35E-04	11	59	FHL2,NEBL,LIMCH1,LMO3,SCEL,CRIP2,CSRP1,TES,PXN,LMO7,LPP
NADH:ubiquinone oxidoreductase supernumerary subunits	3.35E-04	8	30	NDUFB2,NDUFB4,NDUFB7,NDUFB10,NDUFC1,NDUFS5,NDUFS6,NDUFA11
WD repeat domain containing|Dyneins, cytoplasmic	2.37E-03	5	13	DYNLT1,DYNLRB1,DYNC1H1,DYNC1I2,DYNC1LI2
Zinc fingers FYVE-type|Pleckstrin homology domain containing|Rho guanine nucleotide exchange factors	4.72E-03	19	206	NET1,ARHGAP27,KIF1B,SPTBN1,BCR,MPRIP,FGD4,PSD3,GRB7,TRIOBP,PHLDA2,PLEKHA5,PLEKHH2,OSBPL3,RAPH1,PLEKHM1,FARP1,SBF2,PLEKHA1
Integrin beta subunits	5.24E-03	4	9	ITGB1,ITGB4,ITGB5,ITGB6
Erb-b2 receptor tyrosine kinases	5.24E-03	3	4	ERBB2,ERBB3,EGFR
Short chain dehydrogenase/reductase superfamily	8.78E-03	10	76	NMRAL1,UXS1,HTATIP2,DHRS7,GALE,BLVRB,HPGD,SDR16C5,RDH10,DHRS9
Charged multivesicular body proteins|ESCRT-III	9.82E-03	4	11	CHMP4C,CHMP5,CHMP2A,CHMP4B
STEAP family	9.82E-03	3	5	STEAP1,STEAP2,STEAP4
Rho family GTPases	1.17E-02	5	20	RHOB,RHOC,RND3,RHOV,RAC1
Protein phosphatase 1 regulatory subunits	1.30E-02	16	181	PPP1R9A,SH2D4A,OCLN,AHCYL1,BCL2L1,STAU1,PARD3,TNS1,PHACTR3,ZFYVE16,SYTL2,PTK2,ITPR3,PPP1R3B,FARP1,SPRED1
Mucins	1.30E-02	5	21	MUC15,MUC1,MUC5B,MUC21,MUC16
Annexins	1.53E-02	4	13	ANXA1,ANXA2,ANXA4,ANXA11
Claudins	1.81E-02	5	23	CLDN12,CLDN1,CLDN4,CLDN3,CLDN7
14-3-3 phospho-serine/phospho-threonine binding proteins	2.12E-02	3	7	SFN,YWHAG,YWHAH
Charged multivesicular body proteins|ESCRT-III	2.12E-02	3	7	CHMP4C,CHMP2A,CHMP4B
Basic leucine zipper proteins	2.14E-02	7	49	BATF,FOSL2,NFE2L1,NFE2L2,CEBPB,CEBPD,MAFK
Cathepsins	2.14E-02	4	15	CTSA,CTSB,CTSD,CTSL
Blood group antigens	2.27E-02	6	37	ACHE,FUT3,CD151,SMIM1,CD55,BCAM
Protein phosphatase 1 regulatory subunits|Prefoldin subunits	2.64E-02	3	8	PFDN6,PFDN1,PFDN2
Transmembrane channel likes	2.64E-02	3	8	TMC4,TMC6,TMC5
Tumor necrosis factor receptor superfamily	3.59E-02	5	29	TNFRSF12A,TNFRSF10B,TNFRSF21,TNFRSF1A,LTBR
CD molecules|Complement system|Integrin alpha subunits	3.59E-02	4	18	ITGA1,ITGA2,ITGA3,ITGAV
Paraoxonases	3.82E-02	2	3	PON2,PON3
Apolipoproteins|Sushi domain containing	3.82E-02	7	57	C1RL,CD46,CFB,SRPX2,CFH,CD55,SEZ6L2
Armadillo repeat containing	3.82E-02	6	43	PKP4,PKP3,PKP2,CTNNB1,CTNND1,JUP
Endogenous ligands	4.20E-02	17	237	ADM,CCL20,CX3CL1,FN1,ANXA1,APP,CFLAR,C3,CALM2,CXCL2,ICAM1,PSAP,RAC1,EDN1,LGALS1,LGALS3,LGALS3BP
Transmembrane p24 trafficking proteins	4.20E-02	3	10	TMED10,TMED4,TMED3

D.3.3. Gene ontology analysis of B cell cluster PC (Plasma cell).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in PC (Plasma cell). The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

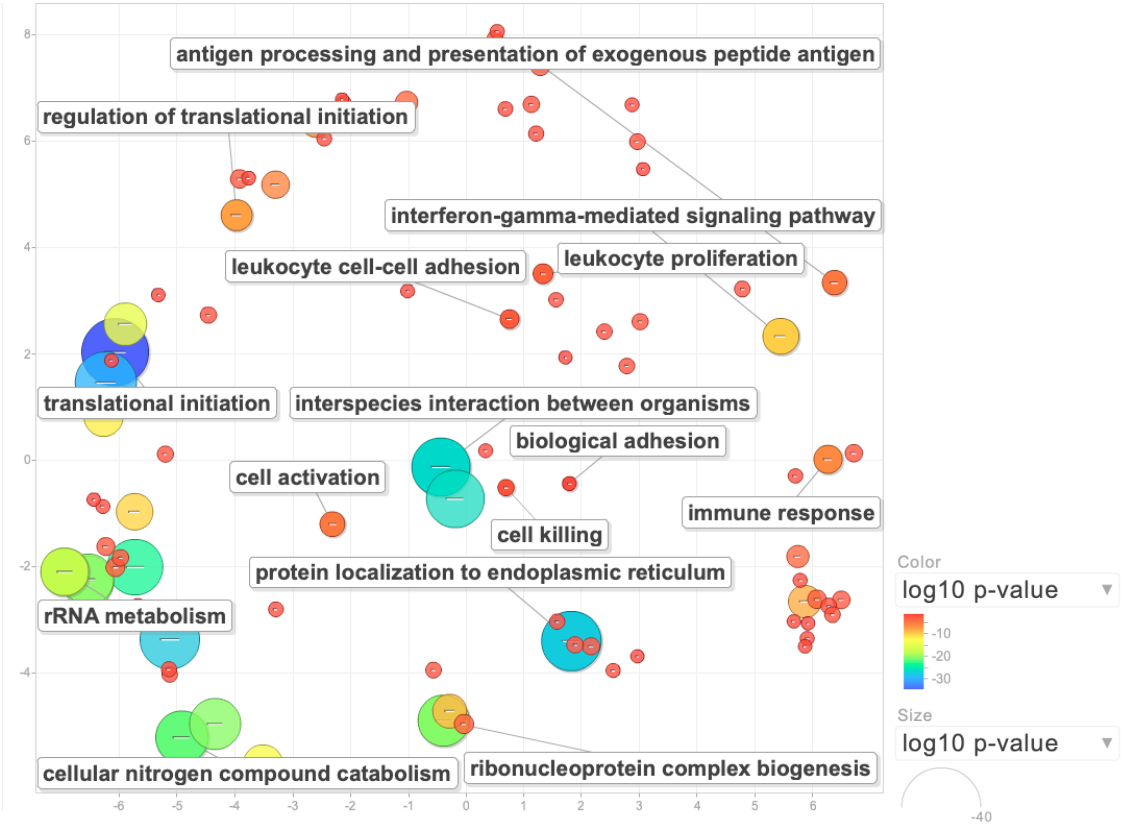
PC (Plasma cell cluster) -Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
movement of cell or subcellular component	7.29E-16	186	1882	EPS8,ERBB2,ERBB3,NET1,S100A14,ETV1,ACTN4,ACVR1B,RRAS,RREB1,MDK,F3,S100A2,TNFRSF12A,DOCK5,S100A9,MET,TUBB4B,MTUS1,MIF,ALCAM,CCL20,CX3CL1,SDC1,SDC4,KLC3,FLNA,SEMA4B,CIB1,CDC42BPA,FN1,ANXA1,KIF21A,TUBB,MYO1B,SIX1,APP,ARF4,RHOB,RHOC,ARHGAP5,SPINT2,PHPT1,ENAH,NUMB,ATP1B1,FUT8,ATP2A2,ATP2B4,DNAH11,MYH9,NCKAP1,MYO6,AVL9,ADAM9,KIF1B,SPTBN1,SRI,CEACAM6,GOLPH3,BLOC1S1,BCR,ST14,STAT3,IQGAP1,NRP1,STC1,PARD6B,BMPR1B,SRPX2,DST,NF1,NFE2L2,NFIB,TSPO,GPC1,DSTN,CALR,ABHD2,GPX1,GRB7,CXCL2,NR2F1,GSN,THBS1,CD9,SLC16A3,TIMP1,NKX2-1,FURIN,CD47,TREM1,PAK1,CD63,CD151,PARD3,TNNT1,CDH1,TNS1,AGRN,TPM1,TPM4,CD2AP,PHLDA2,PLXNB2,TXN,SUN2,PKP2,PLAT,HES1,PLAU,PLAUR,SMURF2,HSPB1,VEGFA,CERS2,VIM,FNDC3B,DDIT4,ICAM1,ID1,CXCL14,CLDN7,CELSR1,IGFBP3,PRKCA,PKN2,CTNNA1,CTNNB1,IL1R1,MAP2K2,SLK,INSR,ITGA1,ITGA2,FNBP1L,ITGA3,ITGAV,ITGB1,ITGB4,PTGS2,PTK2,PTP4A1,SEMA4C,JUP,PTPRF,PTPRK,PXN,DYNLRB1,KDR,RAB1A,DNAH5,WIPF2,RAB13,DYNC1H1,DYNC1I2,DYNC1LI2,RAC1,KRAS,DPYSL3,RAB25,DSC2,PDCD6,DSG2,DSP,RTN4,RDX,LAMC1,EDN1,LGALS3,TSPAN1,NTN4,LRPPRC,EFNA1,EFNA5,LMNA,EGFR,MYH14,LRP5,EMP2,CTTN,SLIRP,TACSTD2,EPCAM,TRIB1,SMAD3
localization of cell	1.35E-13	147	1428	EPS8,ERBB2,ERBB3,NET1,S100A14,ACTN4,ACVR1B,RRAS,RREB1,MDK,F3,S100A2,TNFRSF12A,DOCK5,S100A9,MET,MTUS1,MIF,CCL20,CX3CL1,SDC1,SDC4,FLNA,SEMA4B,CIB1,CDC42BPA,FN1,ANXA1,SIX1,ARF4,RHOB,RHOC,ARHGAP5,SPINT2,PHPT1,NUMB,ATP1B1,FUT8,ATP2B4,DNAH11,MYH9,NCKAP1,AVL9,ADAM9,CEACAM6,GOLPH3,BCR,ST14,STAT3,IQGAP1,NRP1,STC1,PARD6B,SRPX2,DST,NF1,NFE2L2,TSPO,CALR,ABHD2,GPX1,GRB7,CXCL2,NR2F1,GSN,THBS1,SLC16A3,TIMP1,NKX2-1,FURIN,CD47,TREM1,PAK1,CD63,CD151,PARD3,CDH1,TNS1,TPM1,CD2AP,PHLDA2,PLXNB2,SUN2,PKP2,PLAT,HES1,PLAU,SMURF2,HSPB1,VEGFA,CERS2,VIM,FNDC3B,DDIT4,ICAM1,ID1,CXCL14,CLDN7,CELSR1,IGFBP3,PRKCA,PKN2,CTNNA1,CTNNB1,IL1R1,MAP2K2,SLK,INSR,ITGA1,ITGA2,ITGA3,ITGAV,ITGB1,ITGB4,PTGS2,PTK2,PTP4A1,SEMA4C,JUP,PTPRF,PTPRK,KDR,RAB1A,DNAH5,RAB13,RAC1,KRAS,DPYSL3,RAB25,PDCD6,RTN4,RDX,LAMC1,EDN1,LGALS3,TSPAN1,EFNA1,LMNA,EGFR,LRP5,EMP2,CTTN,SLIRP,TACSTD2,EPCAM,TRIB1,SMAD3
epithelium development	3.63E-13	136	1296	FOXQ1,EVPL,ACVR1B,RREB1,ADM,ESRP2,MET,AHR,TAGLN2,SDC1,SDC4,FLNA,FLNB,ANXA1,FOSL2,ANXA4,AGR2,AGPAT2,SIX1,RHOB,RHOC,SPINT2,ENAH,NUMB,ATP2A2,SOX4,NCKAP1,SPINT1,MYO6,ADAM9,CCND1,SCEL,ST14,IQGAP1,NRP1,STC1,NF1,NFATC4,NFIB,GNAQ,GNAS,NOTCH2,SFN,PLOD3,GPX1,NUMA1,TEAD1,GRSF1,GSN,TGM2,NKX2-1,ERRFI1,PAK1,CD63,TNFRSF1A,PARD3,CDH1,CDH3,CDKN1A,CEBPB,PHB2,PLXNB2,PGK1,GRHL1,CLTC,UGCG,TJP2,HES1,PLAUR,MLPH,PLS3,SMURF2,CNN3,VEGFA,VIM,FNDC3B,COL17A1,TNC,ICAM1,ID1,XDH,CLDN3,CRABP2,SLC40A1,PPL,CELSR1,CST6,CTNNB1,CTNND1,CTSB,CTSD,MAP2K2,PRSS8,PSAP,VEZF1,PSMA7,INSR,PSMB5,PSMC1,PSMC3,PSMD2,C1GALT1,ITGA2,PTGS2,SEMA4C,PTPRS,EHF,PXN,KDR,RAB1A,RAB13,RAC1,KRAS,RAB25,RARA,KRT15,DSP,RDH10,SMURF1,LAMB3,RDX,LAMC2,EDN1,LGALS3,RHEB,NTN4,LLGL2,EGFR,DHRS9,LRP5,ELF3,EMP1,TACSTD2,EPCAM,SMAD3,SMAD5
regulation of cell death	5.42E-12	157	1650	ERBB2,ERBB3,NET1,ACTN4,MDK,F3,PICALM,TNFRSF12A,ADM,S100A9,MET,NDRG1,MIF,AHR,FHL2,CX3CL1,FLNA,CIB1,SH3RF1,FN1,PHLDA1,AAMDC,ANXA1,ANXA4,AGR2,HTATIP2,YBX3,SIX1,PDXK,ARF4,RHOB,MUC1,PTGES3,SOD2,SOX4,PLK2,PERP,GOLPH3,BCL2L1,BCL3,TNFRSF10B,TNFRSF21,CHMP2A,GFER,TMBIM1,STAT3,NRP1,CFLAR,BMP1,BMP3,TSC22D1,BMPR1B,BNIP3,BNIP3L,BOK,NF1,IER3,NFATC4,FGD4,NFE2L2,SQSTM1,TAX1BP1,TOX3,TSPO,SERINC3,GNAQ,NME4,PTTG1IP,NOTCH2,SFN,TRIM2,CALR,GPX1,CASP4,TMBIM6,GSN,TGFA,GSTP1,TGM2,THBS1,TIMP1,P4HB,PA2G4,PDCD5,CD59,TNFRSF1A,CDH1,AGRN,CDKN1A,HSP90B1,CEBPB,TRIO,HEBP2,PHB2,TXN,PIM1,PLAUR,PRDX5,HSPA1B,HSPB1,VEGFA,DDIT4,GDF15,ICAM1,ID1,POR,XDH,CLDN7,NIBAN2,KANK2,YWHAG,SLC40A1,YWHAH,PIM3,IGFBP3,VSTM2L,PRKAA1,CST3,PRKCA,CTNNA1,CTNNB1,CTSB,SLK,PSAP,TMEM219,DAD1,PSMC1,ITGA1,ITGAV,ITGB1,PTGS2,PTK2,PTPRF,KDR,RAC1,KRAS,RARA,PDCD6,KRT18,RTN4,DUSP6,FIS1,LDHA,CHMP4B,EDN1,LGALS1,LGALS3,RPS27L,EFNA1,LMNA,EGFR,ANGPTL4,LRP5,LTBR,CTTN,EPCAM,SMAD3
cytoskeleton organization	1.80E-11	121	1164	EPS8,ACTN4,EHD2,S100A9,S100A10,MKLN1,TAOK1,BAIAP2,ALDOA,FKBP4,SDC4,CHMP4C,FLNA,FLNB,CIB1,CDC42BPA,NEBL,ANXA1,CHMP5,PPP1R9A,ARPC1A,TUBB,MYO1B,RND3,HOOK3,PHPT1,DCTN6,STAU2,ENAH,LIMCH1,PLK2,MYH9,NCKAP1,GAPDH,SPTBN1,BCR,CHMP2A,CFLAR,DST,NF1,BAIAP2L1,FGD4,RALBP1,SVIL,CAMSAP2,GOLGA2,DSTN,ARHGEF28,CALR,CAPN2,NUMA1,CASP4,TRIOBP,DYNLT1,GSN,PAK1,PAM,PARD3,AGRN,PACSIN2,TPM1,TPR,HSP90B1,CDC42EP4,CETN2,CD2AP,PFDN1,PFDN2,CHD3,TRIP10,SUN2,CLTC,FBXW5,PKP2,PLEKHH2,CNN3,VIM,ICAM1,ID1,RICTOR,KANK2,YWHAH,CELSR1,CSNK1A1,CSNK1D,CSRP1,PRKAR1A,CTNNA1,CTNNB1,SLK,SPIRE1,ITGB1,ITGB5,PTK2,JUP,STMN3,PXN,DNAH5,WIPF2,RAB13,DYNC1H1,DYNC1LI2,RAC1,KRAS,KRT8,DPYSL3,KRT18,KRT19,DSP,RDX,CHMP4B,EDN1,LLGL2,EFNA5,LMNA,MYH14,EMP2,CTTN,TACSTD2,MARCKS,SMAD3
response to endogenous stimulus	1.91E-11	161	1740	NR2F6,FSTL3,STUB1,MDK,F3,ADM,ESRP2,ADSS2,MGST1,MGST3,ATP6V1D,PGRMC2,AHR,BAIAP2,FHL2,SDC1,FKBP4,CIB1,FN1,ANXA1,ANXA2,FOSL2,SLC34A2,APP,TRIM33,NPTN,ASPH,DDRGK1,FUT8,ATP2A2,PTGES3,ATP2B4,AHCYL1,ATP6V1E1,GALNT3,SPINT1,ADAM9,LMO3,CCND1,BCL2L1,CFB,STAT3,IQGAP1,STC1,BMP1,BMP3,BMPR1B,GNG12,BAIAP2L1,BUD31,TSPO,C3,GNAI1,GNAS,TMED10,GNG10,NOTCH2,GPC1,PLOD3,CALM2,CALR,ABHD2,CAPN2,CASP4,GRN,CAV2,TMBIM6,NR2F1,GSN,GSTP1,THBS1,CD9,TIMP1,TIMP2,NKX2-1,PDCD5,FURIN,ERRFI1,PAK1,PAM,PARD3,CDH1,CDKN1A,TPR,HSP90B1,CEBPB,PHB2,CBX3,HNMT,HNRNPU,SERPINA1,VKORC1,HPGD,PKM,NUCKS1,HES1,PLAU,PLOD2,SMURF2,DDIT4,STRN3,TNC,GDF15,ICAM1,POLR2I,POLR2L,ID1,POR,CLDN4,KANK2,YWHAG,YWHAH,CSNK1D,PRKAA1,PRKAR1A,PRKCA,LEPROT,CTNNA1,CTNNB1,IL1R1,CTSB,CTSL,ZFYVE16,INSR,ITGA2,ITGA3,STEAP2,ITGB1,PMEPA1,ITGB5,PTGS2,PTK2,ITPR3,DHX15,DEFB1,JUP,PTPRF,PTPRK,PXN,KDR,TWSG1,RAB13,MED13,KRAS,RARA,DSG2,KRT19,DUSP6,SMURF1,RDX,LDHA,EDN1,NAMPT,EFNA5,EGFR,SPRED1,LTBP1,LTBP3,MARCKS,SMAD3,SMAD5
regulation of anatomical structure morphogenesis	2.23E-11	118	1128	EPS8,ERBB2,NET1,ACHE,TBC1D2,ACTN4,RRAS,RREB1,F3,EHD2,TNFRSF12A,ADM,MET,S100A10,S100A13,MKLN1,BAIAP2,ALDOA,CX3CL1,FLNA,SEMA4B,CIB1,FN1,ANXA1,PPP1R9A,HTATIP2,SIX1,RHOB,STAU2,NUMB,MYH9,ADAM9,TBC1D2B,BCR,GDI1,TMBIM1,NRP1,CFLAR,SRPX2,BNIP3,NF1,NFATC4,FGD4,NFE2L2,NFIB,C3,GOLGA2,CALR,TRIOBP,GSN,THBS1,GTF2I,NKX2-1,TNFRSF1A,CDH1,AGRN,TPM1,CDC42EP4,PHB2,MGLL,CHI3L1,PLXNB2,PDCL3,CLTC,HES1,PLAU,SMURF2,HSPB1,VEGFA,VIM,GDF15,ICAM1,ID1,POR,XDH,CPE,CRABP2,YWHAH,CELSR1,CSNK1A1,CSNK1D,CST3,PRKCA,CTNNB1,MAP2K2,MYL12B,CUX1,PSMA7,PSMB5,PSMC1,PSMD2,PTGS2,PTK2,SEMA4C,PTPRF,PXN,KDR,WIPF2,RAC1,PDCD6,RTN4,DUSP6,SMURF1,RDX,FIS1,EDN1,LGALS3,NTN4,EFNA1,EFNA5,MYH14,ANGPTL4,EMP2,CTTN,TACSTD2,MARCKS,SMAD3
cell adhesion	1.90E-10	143	1530	ERBB2,ERBB3,FSTL3,NET1,ACHE,ACTN4,CD46,RREB1,TNFRSF12A,S100A9,S100A10,MKLN1,ALCAM,CX3CL1,SDC4,FLNA,CIB1,FN1,BATF,ANXA1,PPFIBP1,PKP4,AGR2,APP,CLSTN1,CASK,RHOB,RND3,ARHGAP5,NPTN,SPINT2,ATP1B1,MUC1,ATP2A2,SOX4,PAG1,MYH9,ADAM9,PERP,GOLPH3,BCL3,TNFRSF21,IQGAP1,TINAGL1,SRPX2,DST,NF1,NINJ1,GNAS,CALR,TRIOBP,CLDN12,CLDN1,GSN,GSTP1,TGM2,THBS1,CD9,PKP3,CD47,PAK1,CD59,CD63,CD151,PARD3,CDH1,CDH3,TPM1,MAGI1,MUC21,CEBPB,HLA-A,CD2AP,PLXNB2,PKP2,HES1,PLAU,HSPB1,VEGFA,FNDC3B,COL17A1,TNC,ICAM1,ID1,RICTOR,CLDN4,CLDN3,CLDN7,PPP1CB,CELSR1,CSRP1,PRKAR1A,PRKCA,PKN2,CTNNA1,CTNNB1,CTNND1,SLK,CD55,ITGA1,ITGA2,ITGA3,ITGAV,ITGB1,TMEM8A,ITGB4,ITGB5,ITGB6,PTK2,JUP,PTPRF,PTPRK,PTPRS,PXN,KDR,RAB1A,RAC1,VMP1,RARA,DSC2,DSG2,DSP,MUC16,LAMB3,RDX,LAMC1,LAMC2,LGALS1,LGALS3,LGALS3BP,EFNA1,EFNA5,EGFR,LMO7,LPP,BCAM,MPZL2,EMP2,CTTN,TACSTD2,EPCAM,MARCKS,SMAD3
tissue migration	7.02E-06	34	244	RREB1,S100A2,MET,CIB1,RHOB,MYH9,ADAM9,NRP1,STC1,SRPX2,NF1,NFE2L2,GPX1,THBS1,CD63,HSPB1,VEGFA,ID1,PRKCA,PKN2,ITGA2,ITGA3,ITGAV,ITGB1,PTGS2,PTK2,KDR,RAB13,RAB25,PDCD6,EDN1,EFNA1,EMP2,TACSTD2
cell-cell junction assembly	5.40E-05	18	95	ACTN4,PKP4,RHOC,OCLN,PARD6B,CLDN1,CD9,PKP3,PARD3,CDH1,PKP2,CLDN3,PRKCA,PKN2,CTNNA1,ITGB1,JUP,RAB13
defense response	4.23E-03	116	1651	S100A14,C1RL,CD46,TMEM173,MDK,F3,ADM,S100A9,TUBB4B,HIST2H2BE,IFITM3,MIF,CCL20,CX3CL1,SDC1,FN1,BATF,ANXA1,TUBB,APP,SLPI,ATP1B1,SNRPD1,GAPDH,IL37,BCL3,TNFRSF10B,BCR,TNFRSF21,GFER,CFB,STAT3,BMPR1B,BNIP3,BNIP3L,SMPDL3B,IER3,NFATC4,NFE2L1,NFE2L2,SERINC3,C3,STK39,TAPBP,NOTCH2,NTHL1,GPX1,GPX4,CASP4,NFKBIZ,MUC5B,CXCL2,GSTP1,TGM2,THBS1,CD47,TREM1,PAK1,CD59,OSMR,TNFRSF1A,CFH,HSP90B1,CEBPB,HLA-A,PHB2,HMGB3,MGLL,CHI3L1,SERPINA1,AP1S1,PRDX5,HSPA1B,HIST1H2BK,DDIT4,ICAM1,PTGES,RICTOR,CLDN7,IGFBP4,CST3,PRKCA,IL1R1,IL1RAP,CTSB,CTSL,ROMO1,VEZF1,PSMA7,PSMB5,CD55,PSMC1,PSMC3,PSMD2,ITGA2,ITGAV,ITGB1,ITGB6,PTGS2,PTK2,DEFB1,RAB1A,RAC1,KRAS,SMURF1,LCN2,EDN1,LGALS3,LGALS3BP,MLF2,EGFR,EIF4G1,ELF3,LTBR,SMAD3,STOM
positive regulation of programmed cell death	9.01E-03	53	655	NET1,F3,TNFRSF12A,ADM,S100A9,PHLDA1,ANXA1,HTATIP2,RHOB,SOX4,BCL2L1,TNFRSF10B,TSC22D1,BMPR1B,BNIP3,BNIP3L,BOK,NF1,NFATC4,FGD4,SQSTM1,TSPO,SERINC3,NOTCH2,SFN,GSN,TGM2,THBS1,PDCD5,TNFRSF1A,AGRN,CDKN1A,TRIO,PLAUR,XDH,YWHAG,YWHAH,IGFBP3,CTNNA1,CTNNB1,SLK,ITGA1,ITGB1,PTGS2,PTPRF,RAC1,PDCD6,DUSP6,FIS1,LDHA,RPS27L,LTBR,SMAD3
regulation of cell division	2.26E-02	27	291	MDK,MET,CHMP4C,MACC1,CIB1,CHMP5,PKP4,ANAPC11,CHMP2A,SVIL,SFN,CALM2,CALR,CAV2,TGFA,ZFYVE19,TPR,CETN2,PHB2,VEGFA,YWHAH,PRKAR1A,PKN2,INSR,CHMP4B,EDN1,LRP5

D.3.3. Gene ontology analysis of B cell cluster PC (Plasma cell).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in PC (Plasma cell). The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.4 REVIGO gene ontology reduction for B1 CD8^{Hi} (Blue)



D.3.4. REVIGO gene ontology reduction for B1 CD8^{Hi}(Blue).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

B1 CD8Hi(Blue) -Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
Histocompatibility complex|C1-set domain containing	1.91E-12	10	42	HLA-A,HLA-C,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,HLA-E
Eukaryotic translation initiation factor 3	1.03E-05	4	13	EIF3E,EIF3D,EIF3F,EIF3G
RNA binding motif containing|Serine and arginine rich splicing factors	4.61E-04	3	12	SRSF2,SRSF5,SRSF7
Basic leucine zipper proteins|Jun transcription factor family	9.61E-04	2	3	JUN,JUNB
Regulators of G-protein signaling	2.08E-03	3	21	RGS19,RGS1,RGS2
Actins	3.80E-03	2	6	ACTB,ACTG1
MAP kinase phosphatases	1.24E-02	2	11	DUSP1,DUSP2
CD molecules|Tumor necrosis factor superfamily	1.60E-02	8	394	SELL,CXCR4,CD37,CD48,TNFRSF13B,CD69,CD79A,LY9
Cathepsins	1.84E-02	2	15	CTSH,CTSS
Minor histocompatibility antigens|FERM domain containing	1.84E-02	3	51	RPS4Y1,ARHGDIB,CTSH
Zinc fingers CCHC-type	4.70E-02	2	25	SRSF7,CNBP
Tubulins	4.75E-02	2	26	TUBA1B,TUBA1A

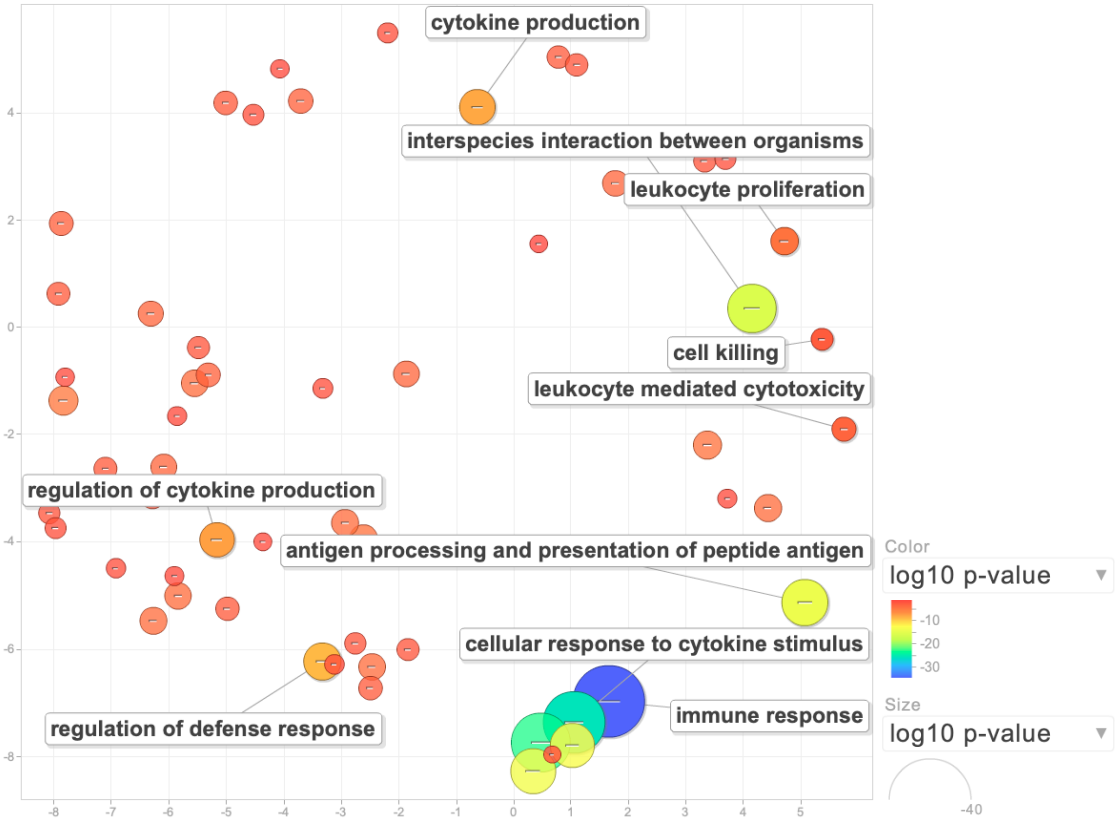
B1 CD8Hi(Blue)-Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
translational initiation	1.40E-36	36	194	RPL22,NPM1,RPL30,RPL29,RPL37,RPL36A,RPLP0,RPLP1,RPS2,RPS3,RPS3A,RPS4Y1,RPS7,EIF3E,RPS12,RPSA,RPS23,RPL14,PPP1R15A,RPL10A,EIF4A1,EIF4A2,EIF4B,EIF3D,EIF3F,EIF3G,EIF1,RPL3,RPL4,RPL7,RPL7A,RPL8,RPL9,RPL10,RPL18,RPL19
protein localization to endoplasmic reticulum	1.56E-29	27	125	RPL22,RPL30,RPL29,RPL37,RPL36A,RPLP0,RPLP1,RPS2,RPS3,RPS3A,RPS4Y1,RPS7,RPS12,RPSA,RPS23,RPL14,PPP1R15A,RPL10A,RPL3,RPL4,RPL7,RPL7A,RPL8,RPL9,RPL10,RPL18,RPL19
multi-organism metabolic process	9.26E-28	30	207	RPL22,RPL30,RPL29,RPL37,RPL36A,RPLP0,RPLP1,RPS2,RPS3,RPS3A,RPS4Y1,RPS7,RPS12,RPSA,RPS23,RPL14,RPL10A,JUN,EIF3D,EIF3F,EIF3G,RPL3,RPL4,RPL7,RPL7A,RPL8,RPL9,RPL10,RPL18,RPL19
interferon-gamma-mediated signaling pathway	5.20E-12	13	85	HLA-A,HLA-C,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,HLA-E,IRF8,SOCS1,HSP90AB1
response to cytokine	1.46E-09	28	825	MYC,RPLP0,HLA-A,HLA-C,GAPDH,HLA-DPA1,RPS2,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,HLA-E,IRF8,KLF2,TUBA1B,JUN,JUNB,CORO1A,UBC,CXCR4,NFKBIA,TNFRSF13B,SOCS1,LTB,RPL3,HSP90AB1
response to interferon-gamma	1.57E-09	14	163	HLA-A,HLA-C,GAPDH,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,HLA-E,IRF8,SOCS1,HSP90AB1
antigen processing and presentation of exogenous peptide antigen	2.67E-06	11	166	HLA-A,HLA-C,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,HLA-E,CTSS
cell activation	6.40E-06	25	1001	RPL22,HLA-A,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,ACTB,HLA-E,ACTG1,GPR183,SAMSN1,TBC1D10C,JUN,CORO1A,ZFP36L2,CLIC1,CXCR4,CD37,CD48,TNFRSF13B,CD79A,LY9
cell killing	9.41E-03	6	134	HLA-A,B2M,HLA-E,CORO1A,CTSH,HSP90AB1
biological adhesion	1.92E-02	23	1542	SELL,RPL22,HLA-A,HLA-DPA1,HLA-DPB1,HLA-DQA1,HLA-DQB1,HLA-DRA,HLA-DRB1,B2M,ACTB,HLA-E,ACTG1,RPSA,CYTIP,ARHGDIB,CORO1A,ZFP36L2,CLIC1,CXCR4,CD48,LY9,HSP90AB1

D.3.4 Gene ontology analysis of B cell cluster B1 CD8^{Hi} (Blue).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in B1. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.5 REVIGO gene ontology reduction for B2 CD8^{Hi} (Red)



D.3.5. REVIGO gene ontology reduction for B2 CD8^{Hi} (Red).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

B2 CD8Hi (Red)-Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
Histocompatibility complex|C1-set domain containing	4.56E-15	13	42	HLA-A,HLA-B,HLA-C,HLA-DMB,HLA-DPA1,HLA-DPB1,HLA-DQA2,HLA-DQB1,HLA-DRA,B2M,HLA-DRB5,HLA-E,HLA-F
Histocompatibility complex|C1-set domain containing	2.16E-13	12	44	HLA-A,HLA-B,HLA-C,HLA-DMB,HLA-DPA1,HLA-DPB1,HLA-DQA2,HLA-DQB1,HLA-DRA,HLA-DRB5,HLA-E,HLA-F
Actin related protein 2/3 complex |WD repeat domain containing	1.65E-09	6	9	ARPC5,ARPC4,ARPC3,ARPC1B,ACTR3,ARPC2
Proteasome	4.42E-09	9	43	PSMA1,PSMA4,PSMA5,PSMA7,PSMB8,PSMB9,PSMD13,PSME1,PSME2
CD molecules|Tumor necrosis factor superfamily	3.02E-06	17	394	TNFSF10,CD164,TNFRSF14,ITGB2,FCGR2B,CXCR4,BST2,CCR7,IFITM1,CD40,CD48,CD53,CD69,CD74,CD79A,FCRL5,IL2RG
Minor histocompatibility antigens|FERM domain containing	1.28E-04	6	51	RPS4Y1,TRIM22,TYMP,SP110,ARHGDIB,CTSH
MAP kinase phosphatases	1.87E-03	3	11	DUSP1,DUSP2,DUSP4
ENAH/VASPs	3.43E-03	2	3	VASP,EVL
Cathepsins	3.90E-03	3	15	CTSC,CTSH,CTSS
Pyrin domain containing|Pyrin and HIN domain family	8.20E-03	2	5	MNDA,IFI16
Regulators of G-protein signaling	8.20E-03	3	21	RGS19,RGS1,RGS2
Protein disulfide isomerases	8.20E-03	3	21	ERP29,PDIA3,PDIA6
Actins	1.04E-02	2	6	ACTB,ACTG1
SH2 domain containing	1.88E-02	5	101	STAT1,PTPN6,BTK,SOCS3,LYN
Caspase recruitment domain containing	1.89E-02	3	30	IFIH1,CASP1,CASP4
Caspases	4.20E-02	2	13	CASP1,CASP4
DEAD-box helicases	4.33E-02	3	42	DDX60,EIF4A1,EIF4A2
Myosin light chains|EF-hand domain containing	4.33E-02	2	14	MYL12A,MYL12B

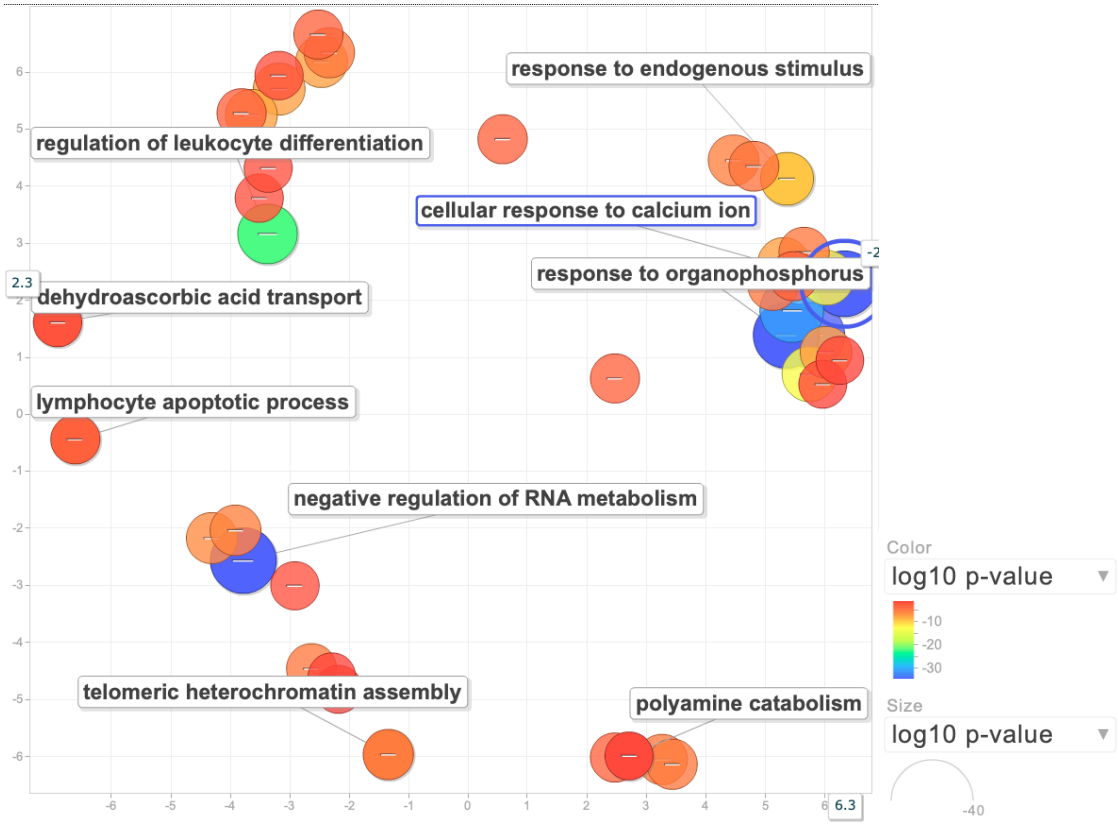
B2 CD8Hi (Red)-Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
cellular response to cytokine stimulus	5.91E-37	68	713	USP18,RPLP0,HLA-A,HLA-B,HLA-C,GAPDH,HLA-DPA1,HLA-DPB1,RPS2,HLA-DQA2,HLA-DQB1,HLA-DRA,PSMA1,PSMA4,PSMA5,B2M,HLA-DRB5,PSMA7,TNFRSF14,HLA-E,HLA-F,PSMB8,PSMB9,GBP1,GBP2,IRF2,IRF7,ISG20,PSMD13,PSME1,PSME2,ADAR,TRIM22,STAT1,KLF2,PTPN1,IRF9,PTPN6,UBC,GBP4,CXCR4,BST2,CPNE1,CCR7,SOCS3,IRF8,IFITM1,OAS1,OAS3,IFITM2,IFI16,IFI35,IFIT2,IFIT3,RBCK1,LGALS9,CORO1A,MT2A,ISG15,CD40,CD74,XAF1,IFI6,IL2RG,HCLS1,MX1,MX2,CYBA
type I interferon signaling pathway	2.05E-33	30	81	USP18,HLA-A,HLA-B,HLA-C,HLA-E,HLA-F,PSMB8,GBP2,IRF2,IRF7,ISG20,ADAR,STAT1,PTPN1,IRF9,PTPN6,BST2,IRF8,IFITM1,OAS1,OAS3,IFITM2,IFI35,IFIT2,IFIT3,ISG15,XAF1,IFI6,MX1,MX2
cellular response to type I interferon	2.40E-33	30	82	USP18,HLA-A,HLA-B,HLA-C,HLA-E,HLA-F,PSMB8,GBP2,IRF2,IRF7,ISG20,ADAR,STAT1,PTPN1,IRF9,PTPN6,BST2,IRF8,IFITM1,OAS1,OAS3,IFITM2,IFI35,IFIT2,IFIT3,ISG15,XAF1,IFI6,MX1,MX2
interferon-gamma-mediated signaling pathway	1.68E-23	24	85	HLA-A,HLA-B,HLA-C,HLA-DPA1,HLA-DPB1,HLA-DQA2,HLA-DQB1,HLA-DRA,B2M,HLA-DRB5,HLA-E,HLA-F,GBP1,GBP2,IRF2,IRF7,TRIM22,STAT1,IRF9,SOCS3,IRF8,OAS1,OAS3,MT2A
cellular response to interferon-gamma	7.49E-22	27	140	HLA-A,HLA-B,HLA-C,GAPDH,HLA-DPA1,HLA-DPB1,HLA-DQA2,HLA-DQB1,HLA-DRA,B2M,HLA-DRB5,HLA-E,HLA-F,GBP1,GBP2,IRF2,IRF7,TRIM22,STAT1,IRF9,GBP4,SOCS3,IRF8,OAS1,OAS3,LGALS9,MT2A
antigen processing and presentation	2.15E-20	31	232	HLA-A,HLA-B,HLA-C,HLA-DMB,HLA-DPA1,HLA-DPB1,HLA-DQA2,HLA-DQB1,HLA-DRA,PSMA1,PSMA4,PSMA5,B2M,HLA-DRB5,PSMA7,HLA-E,HLA-F,PSMB8,PSMB9,PSMD13,PSME1,PSME2,FCGR2B,CCR7,KIF2A,TAP1,PDIA3,CD74,CTSH,CTSS,CYBA
regulation of defense response	2.99E-15	46	820	USP18,HLA-A,HLA-B,PSMA1,PSMA4,PSMA5,B2M,PSMA7,HLA-E,PSMB8,PSMB9,PHB2,APOBEC3G,IRF7,PSMD13,PSME1,PSME2,ADAR,ITGB2,STAT1,IFIH1,PTPN1,PTPN6,UBC,FCGR2B,BTK,PARP9,CCR7,TAP1,PLSCR1,TANK,DDX60,SOCS3,CASP1,CASP4,IFI16,ZFP36,LGALS9,NMI,GPSM3,CD74,LYN,EIF2AK2,CTSS,HERC5,CYBA
cytokine production	2.13E-13	40	700	UBE2L6,HLA-A,HLA-B,HLA-DPA1,HLA-DPB1,B2M,HLA-DRB5,TNFRSF14,HLA-E,GBP1,IRF7,KLF2,IFIH1,IRF9,UBC,FCGR2B,BST2,BTK,TSPO,CCR7,TNFAIP8,LITAF,DDX60,IRF8,CASP1,CASP4,IFI16,BTN3A2,ZFP36,LGALS9,NMI,ISG15,SRGN,CD40,GPSM3,CD74,LYN,EIF2AK2,HERC5,CYBA
regulation of cytokine production	1.18E-12	37	631	UBE2L6,HLA-A,HLA-B,HLA-DPA1,HLA-DPB1,B2M,HLA-DRB5,TNFRSF14,HLA-E,GBP1,IRF7,KLF2,IFIH1,UBC,FCGR2B,BST2,BTK,TSPO,CCR7,LITAF,DDX60,IRF8,CASP1,CASP4,IFI16,ZFP36,LGALS9,NMI,ISG15,SRGN,CD40,GPSM3,CD74,LYN,EIF2AK2,HERC5,CYBA
NIK/NF-kappaB signaling	2.15E-07	13	119	PSMA1,PSMA4,PSMA5,PSMA7,PSMB8,PSMB9,PSMD13,PSME1,PSME2,UBC,GLRX,CPNE1,EIF2AK2
leukocyte mediated cytotoxicity	2.72E-06	11	103	HLA-A,HLA-B,CTSC,B2M,HLA-E,PTPN6,TAP1,LGALS9,CORO1A,PRDX1,CTSH
B cell activation	1.09E-04	14	257	HLA-DQB1,SAMSN1,PTPN6,FCGR2B,BST2,BTK,MNDA,GPR183,CD40,CD74,CD79A,LYL1,LYN,IL2RG
cellular response to interferon-beta	1.43E-04	5	23	GBP2,STAT1,GBP4,IFI16,IFIT3
B cell receptor signaling pathway	2.03E-03	6	65	PTPN6,BTK,MNDA,CD79A,FCRL5,LYN
B cell proliferation	2.13E-03	7	93	FCGR2B,MNDA,GPR183,CD40,CD74,CD79A,LYN
cellular response to interferon-alpha	4.01E-03	3	12	OAS1,IFIT2,IFIT3
cellular response to interleukin-4	4.56E-03	4	29	RPLP0,RPS2,CORO1A,IL2RG
B cell mediated immunity	5.68E-03	8	146	HLA-DQB1,HLA-DRB5,IRF7,PTPN6,FCGR2B,BTK,CD40,CD74
B cell differentiation	1.12E-02	7	129	HLA-DQB1,PTPN6,BTK,GPR183,CD79A,LYL1,IL2RG
B cell apoptotic process	2.93E-02	3	27	BTK,CD74,LYN

D.3.5. Gene ontology analysis of B cell cluster B2 CD8^{Hi} (Red).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in B2. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.6 REVIGO gene ontology reduction for B3 CD8^{Lo}(Green).



D.3.6. REVIGO gene ontology reduction for B2 CD8^{Lo} (Green).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

B3 CD8Lo (Green)-Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
Basic leucine zipper proteins	2.63E-05	4	49	JUND,CREM,FOS,FOSB
Basic leucine zipper proteins|Fos transcription factor family	1.98E-04	2	4	FOS,FOSB
CD molecules|Tumor necrosis factor superfamily	2.41E-04	6	394	CD44,CD55,CD69,CCR7,CD83,LY9
Blood group antigens	1.07E-02	2	37	CD44,CD55
Nuclear hormone receptors	1.49E-02	2	49	NR4A2,NR4A1
Basic leucine zipper proteins|Jun transcription factor family	2.23E-02	1	3	JUND
Phosphatase and actin regulators|Protein phosphatase 1 regulatory subunits	2.47E-02	1	4	PHACTR1
Yippee like family	2.47E-02	1	5	YPEL5
ESCRT-III associated factors	2.47E-02	1	5	CHMP1B
BTG/Tob family	2.67E-02	1	6	BTG1
Proteoglycans|Structural maintenance of chromosomes proteins|Cohesin complex	2.83E-02	1	7	CD44
WD repeat domain containing|BEACH domain containing	3.33E-02	1	9	NBEAL1
C-C motif chemokine receptors	3.41E-02	1	10	CCR7
Charged multivesicular body proteins|ESCRT-III	3.48E-02	1	11	CHMP1B

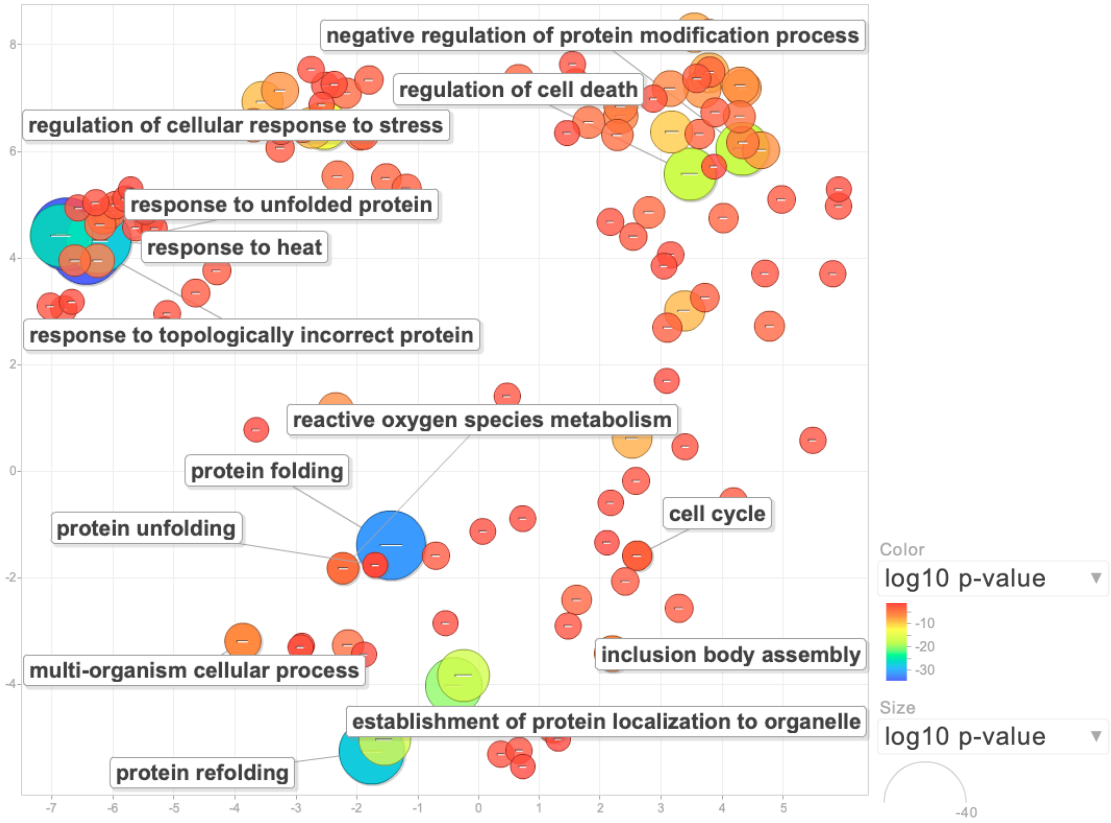
B3 CD8Lo (Green).-Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
CD4-positive, alpha-beta T cell activation	2.40E-04	5	76	CD44,CD55,CCR7,CD83,LY9
response to endogenous stimulus	3.75E-03	12	1740	CD44,NR4A2,SIK1,H3F3B,JUND,CCR7,NR4A1,CREM,FOS,FOSB,SDCBP,IER2
cellular response to calcium ion	8.80E-03	3	53	JUND,FOS,FOSB
regulation of leukocyte activation	1.02E-02	6	510	CD44,CD55,NFKBID,CCR7,CD83,IGHD
regulation of homotypic cell-cell adhesion	1.12E-02	5	348	CD44,CD55,NFKBID,CCR7,CD83
regulation of cell differentiation	1.14E-02	10	1699	CD44,STK4,NFKBID,SIK1,JUND,CCR7,CD83,FOS,SDCBP,BTG1
regulation of cell-cell adhesion	1.90E-02	5	430	CD44,CD55,NFKBID,CCR7,CD83
cell motility	3.15E-02	8	1428	CD44,NR4A2,CCR7,NR4A1,SDCBP,BTG1,IER2,PHACTR1
regulation of apoptotic process	4.12E-02	8	1519	CD44,NR4A2,STK4,NFKBID,SIK1,CCR7,NR4A1,BTG1
defense response	4.91E-02	8	1651	CD44,CD55,NFKBID,CCR7,CD83,LY9,IGHD,FOS

D.3.6. Gene ontology analysis of B cell cluster B3 CD8^{Lo} (Green).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in B3. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.7 REVIGO gene ontology reduction for B4 CD8^{Lo} (Purple).



D.3.7. REVIGO gene ontology reduction for B4 CD8^{Lo} (Purple).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

B4 CD8Lo -Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
Heat shock 70kDa proteins	2.44E-09	5	17	HSPH1,HSPA1A,HSPA1B,HSPA6,HSPA8
DNAJ (HSP40) heat shock proteins	2.51E-05	4	49	DNAJB1,DNAJB6,DNAJB4,DNAJA1
Heat shock 90kDa proteins	2.97E-04	2	5	HSP90AA1,HSP90AB1
Chaperonins	2.31E-03	2	15	HSPD1,HSPE1
Basic leucine zipper proteins|Jun transcription factor family	3.04E-02	1	3	JUN
Ragulator complex	4.04E-02	1	5	LAMTOR4
Peroxiredoxins	4.04E-02	1	6	PRDX1
Armadillo repeat containing|NineTeen complex	4.04E-02	1	7	HSPA8
Gelsolin/villins	4.04E-02	1	8	CAPG
Zinc fingers AN1-type	4.04E-02	1	8	ZFAND2A
Small heat shock proteins	4.26E-02	1	11	HSPB1
BCL2 homology region 3 (BH3) only	4.26E-02	1	11	PMAIP1
MAP kinase phosphatases	4.26E-02	1	11	DUSP2

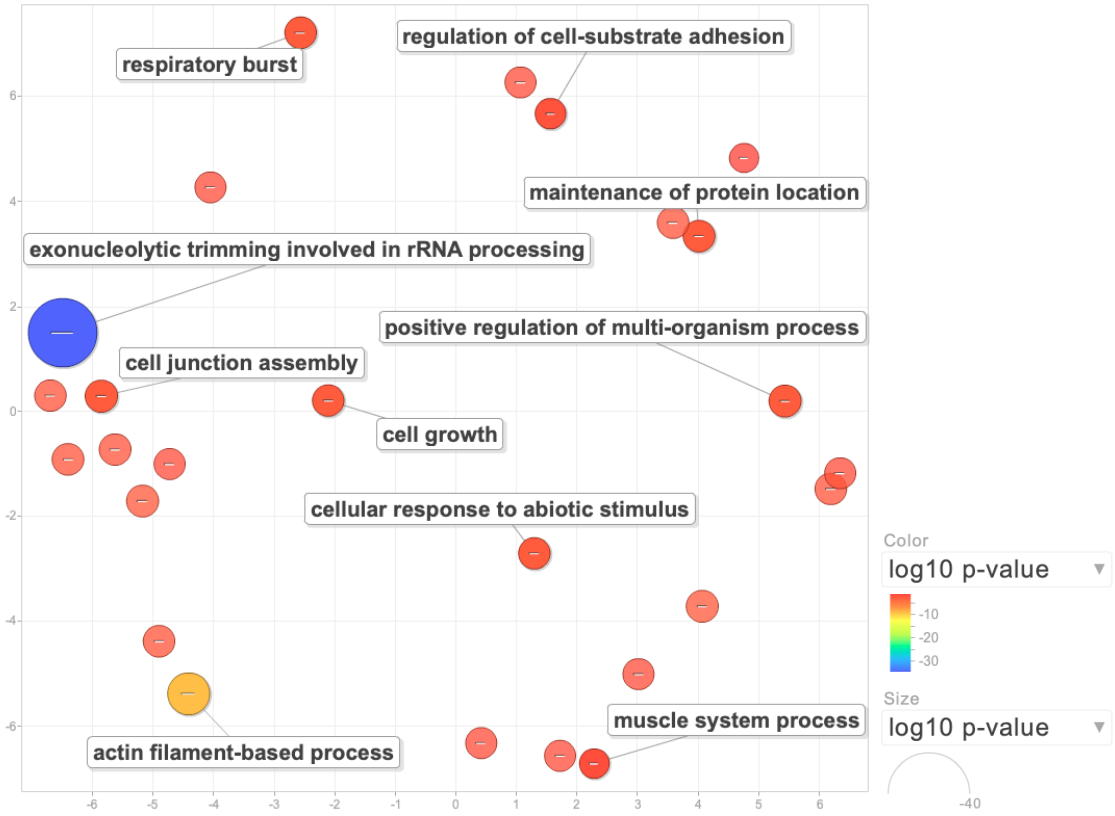
B4 CD8Lo -Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
regulation of cellular response to stress	2.53E-10	17	752	GADD45B,DNAJB1,HSPH1,PRDX1,DNAJB6,PPP1R15A,CHORDC1,DNAJA1,HSPA1A,HSPA1B,PTGES3,HSPA8,HSPB1,PMAIP1,SOD1,HSP90AA1,HSP90AB1
regulation of tumor necrosis factor-mediated signaling pathway	2.29E-04	4	51	UBB,UBC,HSPA1A,HSPA1B
response to heat	5.39E-15	14	171	DNAJB1,HSPH1,DNAJB6,DNAJB4,CHORDC1,DNAJA1,HSPA1A,HSPA1B,PTGES3,HSPA6,HSPA8,SOD1,HSP90AA1,HSP90AB1
response to topologically incorrect protein	5.39E-15	14	184	HSPD1,HSPE1,DNAJB1,HSPH1,TSPYL2,DNAJB4,PPP1R15A,DNAJA1,HSPA1A,HSPA6,HSPA8,HSPB1,HSP90AA1,HSP90AB1
regulation of cell death	1.98E-10	23	1650	HSPD1,RHOB,HSPE1,GADD45B,JUN,DDIT4,UBB,UBC,CIB1,HSPH1,DNAJB6,ID3,CACYBP,PPP1R15A,DNAJA1,HSPA1A,HSPA1B,PTGES3,HSPA8,HSPB1,PMAIP1,SOD1,HSP90AB1
regulation of cytokine production	8.96E-03	7	631	HSPD1,UBB,UBC,HSPA1A,HSPA1B,HSPB1,SOD1
regulation of cell cycle	1.55E-04	12	1008	RHOB,GADD45B,JUN,UBB,UBC,LAMTOR4,TSPYL2,ID3,PPP1R15A,CHORDC1,HSPA8,SERTAD1
regulation of cytokine production	8.96E-03	7	631	HSPD1,UBB,UBC,HSPA1A,HSPA1B,HSPB1,SOD1
protein folding	5.39E-15	15	236	HSPD1,HSPE1,DNAJB1,HSPH1,DNAJB6,DNAJB4,CHORDC1,DNAJA1,HSPA1A,HSPA1B,PTGES3,HSPA6,HSPA8,HSP90AA1,HSP90AB1
autophagy	6.71E-03	7	590	VMP1,UBB,UBC,LAMTOR4,HSPA8,HSPB1,HSP90AA1
biological adhesion	1.10E-02	11	1542	HSPD1,RHOB,VMP1,CIB1,CD99,HSPH1,DNAJB6,HSPB1,SOD1,HSP90AA1,HSP90AB1
cell cycle	4.35E-04	15	1766	RHOB,GADD45B,JUN,UBB,UBC,CIB1,LAMTOR4,TSPYL2,ID3,PPP1R15A,CHORDC1,RGS2,HSPA8,HSP90AA1,SERTAD1
protein refolding	1.43E-13	8	23	HSPD1,DNAJB1,HSPA1A,HSPA1B,PTGES3,HSPA6,HSPA8,HSP90AA1
protein localization to organelle	6.86E-05	12	919	JUN,RPL36A,LAMTOR4,HSPH1,PRDX1,DNAJB6,PPP1R15A,DNAJA1,HSPA8,PMAIP1,HSP90AA1,HSP90AB1

D.3.7. Gene ontology analysis of B cell cluster B4 CD8^{Lo} (Purple).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in B4. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

D.3.8 REVIGO gene ontology reduction for B5 CD8^{Lo} (Orange).



D.3.8. REVIGO gene ontology reduction for B5 CD8^{Lo} (Orange).

Gene ontology results (FDR<0.05) for cluster specific genes, where REVIGO scatterplots show the GO terms after redundancy reduction in 2D space. The bubble colour indicates the log10 *p*-value (*red*, higher; *blue*, lower); size indicates the frequency of the GO term (all GO terms q <0.05).

B5 CD8Lo -Gene Family

Name	FDR	Hit Count	Total no.	Hit in Query List
S100 calcium binding proteins|EF-hand domain containing	2.09E-05	4	21	S100A4,S100A6,S100A10,S100A11
Mitochondrial complex IV: cytochrome c oxidase subunits	5.82E-04	3	19	MT-CO3,COX6C,COX7A2
EF-hand domain containing	8.37E-04	6	219	S100A4,S100A6,S100A11,MYL6,CALM1,CALM2
Proteoglycans|Structural maintenance of chromosomes proteins|Cohesin complex	2.18E-03	2	7	SRGN,CD44
Actin related protein 2/3 complex |WD repeat domain containing	2.97E-03	2	9	ARPC5,ARPC2
F-type ATPases|Mitochondrial complex V: ATP synthase subunits	9.60E-03	2	18	ATP5MC3,ATP5MF
F-type ATPases|Mitochondrial complex V: ATP synthase subunits	9.60E-03	2	19	ATP5MC3,ATP5MF
Rho family GTPases	9.60E-03	2	20	RHOA,CDC42
Selenoproteins	1.33E-02	2	25	GPX1,GPX4
NADH:ubiquinone oxidoreductase supernumerary subunits	1.72E-02	2	30	NDUFB1,NDUFS5
Endogenous ligands	2.48E-02	4	237	PSAP,CALM1,CALM2,LGALS1
CD molecules|Tumor necrosis factor superfamily	2.68E-02	5	394	CD52,CD99,CD44,CD63,ITGB1
Proteasome	2.68E-02	2	43	PSMB9,PSME2

B5 CD8Lo -Gene ontology biological processes

Name	FDR	Hit Count	Total no.	Hit in Query List
actin filament-based process	5.11E-07	18	685	TPM3,RHOA,VIM,S100A10,PTPN1,MYL6,TLN1,CAPG,TMSB4X,TMSB10,PLEK,PFN1,CDC42,ITGB1,ARPC5,MARCKS,SH3BGRL3,ARPC2
actin filament organization	1.59E-06	13	345	RHOA,S100A10,CAPG,TMSB4X,TMSB10,PLEK,PFN1,CDC42,ITGB1,ARPC5,MARCKS,SH3BGRL3,ARPC2
actin filament organization	1.59E-06	13	345	RHOA,S100A10,CAPG,TMSB4X,TMSB10,PLEK,PFN1,CDC42,ITGB1,ARPC5,MARCKS,SH3BGRL3,ARPC2
Fc receptor signaling pathway	5.31E-04	8	206	CALM1,CALM2,ACTB,PSMB9,PSME2,CDC42,ARPC5,ARPC2
cell projection organization	6.20E-04	19	1423	RHOA,S100A6,VIM,DPYSL2,CAPG,MIF,MBP,ACTB,CD44,PLEK,EMP3,PFN1,CPNE5,CDC42,ITGB1,ARPC5,MARCKS,LGALS1,ARPC2
immune response	6.20E-04	20	1572	PTPN1,CALM1,GAPDH,CALM2,MIF,MBP,GPX1,ACTB,CD44,PSMB9,LTB,PSME2,CDC42,ITGB1,ARPC5,CRIP1,LGALS1,HCST,ARPC2,CYBA
positive regulation of multi-organism process	2.87E-03	6	168	MIF,POLR2L,PFN1,PPIA,LGALS1,CYBA
vesicle-mediated transport	3.91E-03	17	1514	S100A10,PTPN1,DPYSL2,PSAP,CALM1,CALM2,SRGN,TLN1,ACTB,CD63,TMSB4X,PLEK,CDC42,ITGB1,ARPC5,ARPC2,CYBA
exocytosis	8.93E-03	8	438	PSAP,CALM1,CALM2,SRGN,TLN1,CD63,TMSB4X,PLEK
integrin-mediated signaling pathway	1.09E-02	4	95	RHOA,CD63,PLEK,ITGB1
cellular response to oxygen-containing compound	1.18E-02	12	1001	PTPN1,PSAP,YBX1,LITAF,MIF,HNRNPM,CBX3,ITGB1,ATP2B1,MARCKS,LGALS1,CYBA
immune response-activating cell surface receptor signaling pathway	2.55E-02	6	332	ACTB,PSMB9,PSME2,CDC42,ARPC5,ARPC2
negative regulation of apoptotic signaling pathway	2.58E-02	5	228	PTPN1,MIF,GPX1,CD44,MCL1
growth	3.50E-02	11	1091	RHOA,DPYSL2,PSAP,GPX1,ARID5B,CD44,EMP3,CPNE5,CDC42,ITGB1,CYBA

D.3.8. Gene ontology analysis of B cell cluster B5 CD8^{Lo} (Orange).

Gene ontology analysis of the differentially expressed genes (FC >0.25 and FDR<0.05) in B5. The significant (FDR<0.05) gene families and biological processes (overlaps reduced with REVIGO) are shown. The significance, number of identified genes and total gene size for each term is shown.

Bibliography

ABBAS, A. R., BALDWIN, D., MA, Y., OUYANG, W., GURNEY, A., MARTIN, F., FONG, S., VAN LOOKEREN CAMPAGNE, M., GODOWSKI, P., WILLIAMS, P. M., CHAN, A. C. & CLARK, H. F. 2005. Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun,* 6**,** 319-31.

AFFARA, N. I., RUFFELL, B., MEDLER, T. R., GUNDERSON, A. J., JOHANSSON, M., BORNSTEIN, S., BERGSLAND, E., STEINHOFF, M., LI, Y., GONG, Q., MA, Y., WIESEN, J. F., WONG, M. H., KULESZ-MARTIN, M., IRVING, B. & COUSSENS, L. M. 2014. B cells regulate macrophage phenotype and response to chemotherapy in squamous carcinomas. *Cancer Cell,* 25**,** 809-21.

AGUIRRE-GHISO, J. A. 2007. Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer,* 7**,** 834-46.

AHMAD, S. S., QIAN, W., ELLIS, S., MASON, E., KHATTAK, M. A., GUPTA, A., SHAW, H., QUINTON, A., KOVARIKOVA, J., THILLAI, K., RAO, A., BOARD, R., NOBES, J., DALGLEISH, A., GRUMETT, S., MARAVEYAS, A., DANSON, S., TALBOT, T., HARRIES, M., MARPLES, M., PLUMMER, R., KUMAR, S., NATHAN, P., MIDDLETON, M. R., LARKIN, J., LORIGAN, P., WHEATER, M., OTTENSMEIER, C. H. & CORRIE, P. G. 2015. Ipilimumab in the real world: the UK expanded access programme experience in previously treated advanced melanoma patients. *Melanoma Res,* 25**,** 432-42.

AKBAR, A. N., HENSON, S. M. & LANNA, A. 2016. Senescence of T Lymphocytes: Implications for Enhancing Human Immunity. *Trends Immunol,* 37**,** 866-876.

AKIRA, S. & TAKEDA, K. 2004. Toll-like receptor signalling. *Nat Rev Immunol,* 4**,** 499-511.

AL-SHIBLI, K. I., DONNEM, T., AL-SAAD, S., PERSSON, M., BREMNES, R. M. & BUSUND, L. T. 2008. Prognostic effect of epithelial and stromal lymphocyte infiltration in non-small cell lung cancer. *Clin Cancer Res,* 14**,** 5220-7.

ALHABBAB, R., BLAIR, P., SMYTH, L. A., RATNASOTHY, K., PENG, Q., MOREAU, A., LECHLER, R., ELGUETA, R. & LOMBARDI, G. 2018. Galectin-1 is required for the regulatory function of B cells. *Sci Rep,* 8**,** 2725.

ALI, K., SOOND, D. R., PINEIRO, R., HAGEMANN, T., PEARCE, W., LIM, E. L., BOUABE, H., SCUDAMORE, C. L., HANCOX, T., MAECKER, H., FRIEDMAN, L., TURNER, M., OKKENHAUG, K. & VANHAESEBROECK, B. 2014. Inactivation of PI(3)K p110delta breaks regulatory T-cell-mediated immune tolerance to cancer. *Nature,* 510**,** 407-411.

ALIZADEH, A. A., ARANDA, V., BARDELLI, A., BLANPAIN, C., BOCK, C., BOROWSKI, C., CALDAS, C., CALIFANO, A., DOHERTY, M., ELSNER, M., ESTELLER, M., FITZGERALD, R., KORBEL, J. O., LICHTER, P., MASON, C. E., NAVIN, N., PE'ER, D., POLYAK, K., ROBERTS, C. W., SIU, L., SNYDER, A., STOWER, H., SWANTON, C., VERHAAK, R. G., ZENKLUSEN, J. C., ZUBER, J. & ZUCMAN-ROSSI, J. 2015. Toward understanding and exploiting tumor heterogeneity. *Nat Med,* 21**,** 846-53.

ALLIE, S. R., BRADLEY, J. E., MUDUNURU, U., SCHULTZ, M. D., GRAF, B. A., LUND, F. E. & RANDALL, T. D. 2019. The establishment of resident memory B cells in the lung requires local antigen encounter. *Nat Immunol,* 20**,** 97-108.

ANDERS, S., PYL, P. T. & HUBER, W. 2014. HTSeq-a Python framework to work with high-throughput sequencing data. *Bioinformatics*.

ANDERSEN, A. S., KOLDJAER SOLLING, A. S., OVESEN, T. & RUSAN, M. 2014. The interplay between HPV and host immunity in head and neck squamous cell carcinoma. *Int J Cancer,* 134**,** 2755-63.

ANDERSON, P., CALIGIURI, M., O'BRIEN, C., MANLEY, T., RITZ, J. & SCHLOSSMAN, S. F. 1990. Fc gamma receptor type III (CD16) is included in the zeta NK receptor complex expressed by human natural killer cells. *Proc Natl Acad Sci U S A,* 87**,** 2274-8.

ANDREANI, V., RAMAMOORTHY, S., PANDEY, A., LUPAR, E., NUTT, S. L., LAMMERMANN, T. & GROSSCHEDL, R. 2018. Cochaperone Mzb1 is a key effector of Blimp1 in plasma cell differentiation and beta1-integrin function. *Proc Natl Acad Sci U S A,* 115**,** E9630-E9639.

ANG, K. K., HARRIS, J., WHEELER, R., WEBER, R., ROSENTHAL, D. I., NGUYEN-TAN, P. F., WESTRA, W. H., CHUNG, C. H., JORDAN, R. C., LU, C., KIM, H., AXELROD, R., SILVERMAN, C. C., REDMOND, K. P. & GILLISON, M. L. 2010. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med,* 363**,** 24-35.

ANGINOT, A., ESPELI, M., CHASSON, L., MANCINI, S. J. & SCHIFF, C. 2013. Galectin 1 modulates plasma cell homeostasis and regulates the humoral immune response. *J Immunol,* 190**,** 5526-33.

ARGIRIS, A., KARAMOUZIS, M. V., RABEN, D. & FERRIS, R. L. 2008. Head and neck cancer. *Lancet,* 371**,** 1695-709.

ASHRAFI, G. H., HAGHSHENAS, M. R., MARCHETTI, B., O'BRIEN, P. M. & CAMPO, M. S. 2005. E5 protein of human papillomavirus type 16 selectively downregulates surface HLA class I. *Int J Cancer,* 113**,** 276-83.

ATTNER, P., DU, J., NASMAN, A., HAMMARSTEDT, L., RAMQVIST, T., LINDHOLM, J., MARKLUND, L., DALIANIS, T. & MUNCK-WIKLAND, E. 2010. The role of human papillomavirus in the increased incidence of base of tongue cancer. *Int J Cancer,* 126**,** 2879-84.

BABA, Y. & KUROSAKI, T. 2016. Role of Calcium Signaling in B Cell Activation and Biology. *Curr Top Microbiol Immunol,* 393**,** 143-174.

BADOUAL, C., HANS, S., RODRIGUEZ, J., PEYRARD, S., KLEIN, C., AGUEZNAY NEL, H., MOSSERI, V., LACCOURREYE, O., BRUNEVAL, P., FRIDMAN, W. H., BRASNU, D. F. & TARTOUR, E. 2006. Prognostic value of tumor-infiltrating CD4+ T-cell subpopulations in head and neck cancers. *Clin Cancer Res,* 12**,** 465-72.

BARBALAT, R., LAU, L., LOCKSLEY, R. M. & BARTON, G. M. 2009. Toll-like receptor 2 on inflammatory monocytes induces type I interferon in response to viral but not bacterial ligands. *Nat Immunol,* 10**,** 1200-7.

BARBER, D. L., WHERRY, E. J., MASOPUST, D., ZHU, B., ALLISON, J. P., SHARPE, A. H., FREEMAN, G. J. & AHMED, R. 2006. Restoring function in exhausted CD8 T cells during chronic viral infection. *Nature,* 439**,** 682-7.

BARCLAY, A. N., WRIGHT, G. J., BROOKE, G. & BROWN, M. H. 2002. CD200 and membrane protein interactions in the control of myeloid cells. *Trends Immunol,* 23**,** 285-90.

BASSO, K. & DALLA-FAVERA, R. 2010. BCL6: master regulator of the germinal center reaction and key oncogene in B cell lymphomagenesis. *Adv Immunol,* 105**,** 193-210.

BASSO, K. & DALLA-FAVERA, R. 2015. Germinal centres and B cell lymphomagenesis. *Nat Rev Immunol,* 15**,** 172-84.

BEERS, S. A., GLENNIE, M. J. & WHITE, A. L. 2016. Influence of immunoglobulin isotype on therapeutic antibody function. *Blood,* 127**,** 1097-101.

BELANGER, S. & CROTTY, S. 2016. Dances with cytokines, featuring TFH cells, IL-21, IL-4 and B cells. *Nat Immunol,* 17**,** 1135-6.

BENDALL, S. C., SIMONDS, E. F., QIU, P., AMIR EL, A. D., KRUTZIK, P. O., FINCK, R., BRUGGNER, R. V., MELAMED, R., TREJO, A., ORNATSKY, O. I., BALDERAS, R. S., PLEVRITIS, S. K., SACHS, K., PE'ER, D., TANNER, S. D. & NOLAN, G. P. 2011. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science,* 332**,** 687-96.

BESSER, M. J., SHAPIRA-FROMMER, R., ITZHAKI, O., TREVES, A. J., ZIPPEL, D. B., LEVY, D., KUBI, A., SHOSHANI, N., ZIKICH, D., OHAYON, Y., OHAYON, D., SHALMON, B., MARKEL, G., YERUSHALMI, R., APTER, S., BEN-NUN, A., BEN-AMI, E., SHIMONI, A., NAGLER, A. & SCHACHTER, J. 2013. Adoptive transfer of tumor-infiltrating lymphocytes in patients with metastatic melanoma: intent-to-treat analysis and efficacy after failure to prior immunotherapies. *Clin Cancer Res,* 19**,** 4792-800.

BINDEA, G., MLECNIK, B., FRIDMAN, W. H. & GALON, J. 2011. The prognostic impact of anti-cancer immune response: a novel classification of cancer patients. *Semin Immunopathol,* 33**,** 335-40.

BINDEA, G., MLECNIK, B., TOSOLINI, M., KIRILOVSKY, A., WALDNER, M., OBENAUF, A. C., ANGELL, H., FREDRIKSEN, T., LAFONTAINE, L., BERGER, A., BRUNEVAL, P., FRIDMAN, W. H., BECKER, C., PAGES, F., SPEICHER, M. R., TRAJANOSKI, Z. & GALON, J. 2013. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity,* 39**,** 782-95.

BIRKELAND, S. A., STORM, H. H., LAMM, L. U., BARLOW, L., BLOHME, I., FORSBERG, B., EKLUND, B., FJELDBORG, O., FRIEDBERG, M., FRODIN, L. & ET AL. 1995. Cancer risk after renal transplantation in the Nordic countries, 1964-1986. *Int J Cancer,* 60**,** 183-9.

BLACKBURN, S. D., SHIN, H., HAINING, W. N., ZOU, T., WORKMAN, C. J., POLLEY, A., BETTS, M. R., FREEMAN, G. J., VIGNALI, D. A. & WHERRY, E. J. 2009. Coregulation of CD8+ T cell exhaustion by multiple inhibitory receptors during chronic viral infection. *Nat Immunol,* 10**,** 29-37.

BOUGHERARA, H., MANSUET-LUPO, A., ALIFANO, M., NGO, C., DAMOTTE, D., LE FRERE-BELDA, M. A., DONNADIEU, E. & PERANZONI, E. 2015. Real-Time Imaging of Resident T Cells in Human Lung and Ovarian Carcinomas Reveals How Different Tumor Microenvironments Control T Lymphocyte Migration. *Front Immunol,* 6**,** 500.

BOUSSO, P. & ROBEY, E. 2003. Dynamics of CD8+ T cell priming by dendritic cells in intact lymph nodes. *Nat Immunol,* 4**,** 579-85.

BRANDES, M., WILLIMANN, K., BIOLEY, G., LEVY, N., EBERL, M., LUO, M., TAMPE, R., LEVY, F., ROMERO, P. & MOSER, B. 2009. Cross-presenting human gammadelta T cells induce robust CD8+ alphabeta T cell responses. *Proc Natl Acad Sci U S A,* 106**,** 2307-12.

BROWN, S. D., WARREN, R. L., GIBB, E. A., MARTIN, S. D., SPINELLI, J. J., NELSON, B. H. & HOLT, R. A. 2014. Neo-antigens predicted by tumor genome meta-analysis correlate with increased patient survival. *Genome Res,* 24**,** 743-50.

BROZ, M. L., BINNEWIES, M., BOLDAJIPOUR, B., NELSON, A. E., POLLACK, J. L., ERLE, D. J., BARCZAK, A., ROSENBLUM, M. D., DAUD, A., BARBER, D. L., AMIGORENA, S., VAN'T VEER, L. J., SPERLING, A. I., WOLF, D. M. & KRUMMEL, M. F. 2014. Dissecting the tumor myeloid compartment reveals rare activating antigen-presenting cells critical for T cell immunity. *Cancer Cell,* 26**,** 638-52.

BRUBAKER, S. W., BONHAM, K. S., ZANONI, I. & KAGAN, J. C. 2015. Innate immune pattern recognition: a cell biological perspective. *Annu Rev Immunol,* 33**,** 257-90.

BRUHNS, P. 2012. Properties of mouse and human IgG receptors and their contribution to disease models. *Blood,* 119**,** 5640-9.

BRUHNS, P. & JONSSON, F. 2015. Mouse and human FcR effector functions. *Immunol Rev,* 268**,** 25-51.

BRUNO, T. C., EBNER, P. J., MOORE, B. L., SQUALLS, O. G., WAUGH, K. A., ERUSLANOV, E. B., SINGHAL, S., MITCHELL, J. D., FRANKLIN, W. A., MERRICK, D. T., MCCARTER, M. D., PALMER, B. E., KERN, J. A. & SLANSKY, J. E. 2017. Antigen-Presenting Intratumoral B Cells Affect CD4(+) TIL Phenotypes in Non-Small Cell Lung Cancer Patients. *Cancer Immunol Res,* 5**,** 898-907.

BUNTING, M. D., COMERFORD, I. & MCCOLL, S. R. 2011. Finding their niche: chemokines directing cell migration in the thymus. *Immunol Cell Biol,* 89**,** 185-96.

BURBAGE, M. & KEPPLER, S. J. 2018. Shaping the humoral immune response: Actin regulators modulate antigen presentation and influence B-T interactions. *Mol Immunol,* 101**,** 370-376.

BURNET, F. M. 1970. The concept of immunological surveillance. *Prog Exp Tumor Res,* 13**,** 1-27.

BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. & WITTWER, C. T. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem,* 55**,** 611-22.

BUTTS, C., SOCINSKI, M. A., MITCHELL, P. L., THATCHER, N., HAVEL, L., KRZAKOWSKI, M., NAWROCKI, S., CIULEANU, T. E., BOSQUEE, L., TRIGO, J. M., SPIRA, A., TREMBLAY, L., NYMAN, J., RAMLAU, R., WICKART-JOHANSSON, G., ELLIS, P., GLADKOV, O., PEREIRA, J. R., EBERHARDT, W. E., HELWIG, C., SCHRODER, A., SHEPHERD, F. A. & TEAM, S. T. 2014. Tecemotide (L-BLP25) versus placebo after chemoradiotherapy for stage III non-small-cell lung cancer (START): a randomised, double-blind, phase 3 trial. *Lancet Oncol,* 15**,** 59-68.

CAMPBELL, J. D., ALEXANDROV, A., KIM, J., WALA, J., BERGER, A. H., PEDAMALLU, C. S., SHUKLA, S. A., GUO, G., BROOKS, A. N., MURRAY, B. A., IMIELINSKI, M., HU, X., LING, S., AKBANI, R., ROSENBERG, M., CIBULSKIS, C., RAMACHANDRAN, A., COLLISSON, E. A., KWIATKOWSKI, D. J., LAWRENCE, M. S., WEINSTEIN, J. N., VERHAAK, R. G., WU, C. J., HAMMERMAN, P. S., CHERNIACK, A. D., GETZ, G., CANCER GENOME ATLAS RESEARCH, N., ARTYOMOV, M. N., SCHREIBER, R., GOVINDAN, R. & MEYERSON, M. 2016. Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nat Genet,* 48**,** 607-16.

CANCER GENOME ATLAS, N. 2015. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature,* 517**,** 576-82.

CANCER GENOME ATLAS RESEARCH, N. 2014. Comprehensive molecular profiling of lung adenocarcinoma. *Nature,* 511**,** 543-50.

CAO, A., ALLUQMANI, N., BUHARI, F. H. M., WASIM, L., SMITH, L. K., QUAILE, A. T., SHANNON, M., HAKIM, Z., FURMLI, H., OWEN, D. M., SAVCHENKO, A. & TREANOR, B. 2018. Galectin-9 binds IgM-BCR to regulate B cell signaling. *Nat Commun,* 9**,** 3288.

CAPONE, M., ROMAGNOLI, P., BEERMANN, F., MACDONALD, H. R. & VAN MEERWIJK, J. P. 2001. Dissociation of thymic positive and negative selection in transgenic mice expressing major histocompatibility complex class I molecules exclusively on thymic cortical epithelial cells. *Blood,* 97**,** 1336-42.

CASTIGLI, E., WILSON, S. A., SCOTT, S., DEDEOGLU, F., XU, S., LAM, K. P., BRAM, R. J., JABARA, H. & GEHA, R. S. 2005. TACI and BAFF-R mediate isotype switching in B cells. *J Exp Med,* 201**,** 35-9.

CASTINO, G. F., CORTESE, N., CAPRETTI, G., SERIO, S., DI CARO, G., MINERI, R., MAGRINI, E., GRIZZI, F., CAPPELLO, P., NOVELLI, F., SPAGGIARI, P., RONCALLI, M., RIDOLFI, C., GAVAZZI, F., ZERBI, A., ALLAVENA, P. & MARCHESI, F. 2016. Spatial distribution of B cells predicts prognosis in human pancreatic adenocarcinoma. *Oncoimmunology,* 5**,** e1085147.

CAZA, T. & LANDAS, S. 2015. Functional and Phenotypic Plasticity of CD4(+) T Cell Subsets. *Biomed Res Int,* 2015**,** 521957.

CEDENO-LAURENT, F., OPPERMAN, M. J., BARTHEL, S. R., HAYS, D., SCHATTON, T., ZHAN, Q., HE, X., MATTA, K. L., SUPKO, J. G., FRANK, M. H., MURPHY, G. F. & DIMITROFF, C. J. 2012. Metabolic inhibition of galectin-1-binding carbohydrates accentuates antitumor immunity. *J Invest Dermatol,* 132**,** 410-20.

CHATURVEDI, A. K., ENGELS, E. A., ANDERSON, W. F. & GILLISON, M. L. 2008. Incidence trends for human papillomavirus-related and -unrelated oral squamous cell carcinomas in the United States. *J Clin Oncol,* 26**,** 612-9.

CHATURVEDI, A. K., ENGELS, E. A., PFEIFFER, R. M., HERNANDEZ, B. Y., XIAO, W., KIM, E., JIANG, B., GOODMAN, M. T., SIBUG-SABER, M., COZEN, W., LIU, L., LYNCH, C. F., WENTZENSEN, N., JORDAN, R. C., ALTEKRUSE, S., ANDERSON, W. F., ROSENBERG, P. S. & GILLISON, M. L. 2011. Human papillomavirus and rising oropharyngeal cancer incidence in the United States. *J Clin Oncol,* 29**,** 4294-301.

CHAUDHURI, J. & ALT, F. W. 2004. Class-switch recombination: interplay of transcription, DNA deamination and DNA repair. *Nat Rev Immunol,* 4**,** 541-52.

CHEN, J., BARDES, E. E., ARONOW, B. J. & JEGGA, A. G. 2009. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res,* 37**,** W305-11.

CHEN, L. & FLIES, D. B. 2013. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nat Rev Immunol,* 13**,** 227-42.

CHEN, R., MANOCHAKIAN, R., JAMES, L., AZZOUQA, A. G., SHI, H., ZHANG, Y., ZHAO, Y., ZHOU, K. & LOU, Y. 2020. Emerging therapeutic agents for advanced non-small cell lung cancer. *J Hematol Oncol,* 13**,** 58.

CHEN, Z., FILLMORE, C. M., HAMMERMAN, P. S., KIM, C. F. & WONG, K. K. 2014. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer,* 14**,** 535-46.

CHENG, M., CHEN, Y., XIAO, W., SUN, R. & TIAN, Z. 2013. NK cell-based immunotherapy for malignant diseases. *Cell Mol Immunol,* 10**,** 230-52.

CHOU, R. C., KIM, N. D., SADIK, C. D., SEUNG, E., LAN, Y., BYRNE, M. H., HARIBABU, B., IWAKURA, Y. & LUSTER, A. D. 2010. Lipid-cytokine-chemokine cascade drives neutrophil recruitment in a murine model of inflammatory arthritis. *Immunity,* 33**,** 266-78.

CHOW, L. T., BROKER, T. R. & STEINBERG, B. M. 2010. The natural history of human papillomavirus infections of the mucosal epithelia. *APMIS,* 118**,** 422-49.

CHUNG, C. D., PATEL, V. P., MORAN, M., LEWIS, L. A. & MICELI, M. C. 2000. Galectin-1 induces partial TCR zeta-chain phosphorylation and antagonizes processive TCR signal transduction. *J Immunol,* 165**,** 3722-9.

CHUNG, C. H., PARKER, J. S., KARACA, G., WU, J., FUNKHOUSER, W. K., MOORE, D., BUTTERFOSS, D., XIANG, D., ZANATION, A., YIN, X., SHOCKLEY, W. W., WEISSLER, M. C., DRESSLER, L. G., SHORES, C. G., YARBROUGH, W. G. & PEROU, C. M. 2004. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell,* 5**,** 489-500.

CHUNG, W., EUM, H. H., LEE, H. O., LEE, K. M., LEE, H. B., KIM, K. T., RYU, H. S., KIM, S., LEE, J. E., PARK, Y. H., KAN, Z., HAN, W. & PARK, W. Y. 2017. Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. *Nat Commun,* 8**,** 15081.

CILLO, A. R., KURTEN, C. H. L., TABIB, T., QI, Z., ONKAR, S., WANG, T., LIU, A., DUVVURI, U., KIM, S., SOOSE, R. J., OESTERREICH, S., CHEN, W., LAFYATIS, R., BRUNO, T. C., FERRIS, R. L. & VIGNALI, D. A. A. 2020. Immune Landscape of Viral- and Carcinogen-Driven Head and Neck Cancer. *Immunity,* 52**,** 183-199 e9.

CLARKE, J., PANWAR, B., MADRIGAL, A., SINGH, D., GUJAR, R., WOOD, O., CHEE, S. J., ESCHWEILER, S., KING, E. V., AWAD, A. S., HANLEY, C. J., MCCANN, K. J., BHATTACHARYYA, S., WOO, E., ALZETANI, A., SEUMOIS, G., THOMAS, G. J., GANESAN, A. P., FRIEDMANN, P. S., SANCHEZ-ELSNER, T., AY, F., OTTENSMEIER, C. H. & VIJAYANAND, P. 2019. Single-cell transcriptomic analysis of tissue-resident memory T cells in human lung cancer. *J Exp Med*.

CLIFFORD, G. M., POLESEL, J., RICKENBACH, M., DAL MASO, L., KEISER, O., KOFLER, A., RAPITI, E., LEVI, F., JUNDT, G., FISCH, T., BORDONI, A., DE WECK, D., FRANCESCHI, S. & SWISS, H. I. V. C. 2005. Cancer risk in the Swiss HIV Cohort Study: associations with immunodeficiency, smoking, and highly active antiretroviral therapy. *J Natl Cancer Inst,* 97**,** 425-32.

COHEN, N. R., BRENNAN, P. J., SHAY, T., WATTS, G. F., BRIGL, M., KANG, J., BRENNER, M. B. & IMMGEN PROJECT, C. 2013. Shared and distinct transcriptional programs underlie the hybrid nature of iNKT cells. *Nat Immunol,* 14**,** 90-9.

COLMONT, C. S., BENKETAH, A., REED, S. H., HAWK, N. V., TELFORD, W. G., OHYAMA, M., UDEY, M. C., YEE, C. L., VOGEL, J. C. & PATEL, G. K. 2013. CD200-expressing human basal cell carcinoma cells initiate tumor growth. *Proc Natl Acad Sci U S A,* 110**,** 1434-9.

COQUET, J. M., CHAKRAVARTI, S., KYPARISSOUDIS, K., MCNAB, F. W., PITT, L. A., MCKENZIE, B. S., BERZINS, S. P., SMYTH, M. J. & GODFREY, D. I. 2008. Diverse cytokine production by NKT cell subsets and identification of an IL-17-producing CD4-NK1.1- NKT cell population. *Proc Natl Acad Sci U S A,* 105**,** 11287-92.

CROTTY, S. 2011. Follicular helper CD4 T cells (TFH). *Annu Rev Immunol,* 29**,** 621-63.

DANSO-ABEAM, D., HUMBLET-BARON, S., DOOLEY, J. & LISTON, A. 2011. Models of aire-dependent gene regulation for thymic negative selection. *Front Immunol,* 2**,** 14.

DAVIDSON, M. R., GAZDAR, A. F. & CLARKE, B. E. 2013. The pivotal role of pathology in the management of lung cancer. *J Thorac Dis,* 5 Suppl 5**,** S463-78.

DE VEER, M. J., HOLKO, M., FREVEL, M., WALKER, E., DER, S., PARANJAPE, J. M., SILVERMAN, R. H. & WILLIAMS, B. R. 2001. Functional classification of interferon-stimulated genes identified using microarrays. *J Leukoc Biol,* 69**,** 912-20.

DIEU-NOSJEAN, M. C., GIRALDO, N. A., KAPLON, H., GERMAIN, C., FRIDMAN, W. H. & SAUTES-FRIDMAN, C. 2016. Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. *Immunol Rev,* 271**,** 260-75.

DIMOVA, T., BROUWER, M., GOSSELIN, F., TASSIGNON, J., LEO, O., DONNER, C., MARCHANT, A. & VERMIJLEN, D. 2015. Effector Vgamma9Vdelta2 T cells dominate the human fetal gammadelta T-cell repertoire. *Proc Natl Acad Sci U S A,* 112**,** E556-65.

DJENIDI, F., ADAM, J., GOUBAR, A., DURGEAU, A., MEURICE, G., DE MONTPREVILLE, V., VALIDIRE, P., BESSE, B. & MAMI-CHOUAIB, F. 2015. CD8+CD103+ tumor-infiltrating lymphocytes are tumor-specific tissue-resident memory T cells and a prognostic factor for survival in lung cancer patients. *J Immunol,* 194**,** 3475-86.

DOBBIN, K. K., CESANO, A., ALVAREZ, J., HAWTIN, R., JANETZKI, S., KIRSCH, I., MASUCCI, G. V., ROBBINS, P. B., SELVAN, S. R., STREICHER, H. Z., ZHANG, J., BUTTERFIELD, L. H. & THURIN, M. 2016. Validation of biomarkers to predict response to immunotherapy in cancer: Volume II - clinical validation and regulatory considerations. *J Immunother Cancer,* 4**,** 77.

DOMAGALA-KULAWIK, J. 2015. The role of the immune system in non-small cell lung carcinoma and potential for therapeutic intervention. *Transl Lung Cancer Res,* 4**,** 177-90.

DOMEIER, P. P., CHODISETTI, S. B., SONI, C., SCHELL, S. L., ELIAS, M. J., WONG, E. B., COOPER, T. K., KITAMURA, D. & RAHMAN, Z. S. 2016. IFN-gamma receptor and STAT1 signaling in B cells are central to spontaneous germinal center formation and autoimmunity. *J Exp Med,* 213**,** 715-32.

DONNEM, T., HALD, S. M., PAULSEN, E. E., RICHARDSEN, E., AL-SAAD, S., KILVAER, T. K., BRUSTUGUN, O. T., HELLAND, A., LUND-IVERSEN, M., POEHL, M., OLSEN, K. E., DITZEL, H. J., HANSEN, O., AL-SHIBLI, K., KISELEV, Y., SANDANGER, T. M., ANDERSEN, S., PEZZELLA, F., BREMNES, R. M. & BUSUND, L. T. 2015. Stromal CD8+ T-cell Density-A Promising Supplement to TNM Staging in Non-Small Cell Lung Cancer. *Clin Cancer Res,* 21**,** 2635-43.

DRENNAN, M. B., ELEWAUT, D. & HOGQUIST, K. A. 2009. Thymic emigration: sphingosine-1-phosphate receptor-1-dependent models and beyond. *Eur J Immunol,* 39**,** 925-30.

DUHEN, T., DUHEN, R., MONTLER, R., MOSES, J., MOUDGIL, T., DE MIRANDA, N. F., GOODALL, C. P., BLAIR, T. C., FOX, B. A., MCDERMOTT, J. E., CHANG, S. C., GRUNKEMEIER, G., LEIDNER, R., BELL, R. B. & WEINBERG, A. D. 2018. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun,* 9**,** 2724.

DUNN, G. P., BRUCE, A. T., IKEDA, H., OLD, L. J. & SCHREIBER, R. D. 2002. Cancer immunoediting: from immunosurveillance to tumor escape. *Nat Immunol,* 3**,** 991-8.

DUNNE, P. D., MCART, D. G., BRADLEY, C. A., O'REILLY, P. G., BARRETT, H. L., CUMMINS, R., O'GRADY, T., ARTHUR, K., LOUGHREY, M., ALLEN, W. L., MCDADE, S., WAUGH, D. J., HAMILTON, P. W., LONGLEY, D. B., KAY, E. W., JOHNSTON, P. G., LAWLER, M., SALTO-TELLEZ, M. & VAN SCHAEYBROECK, S. 2016. Challenging the cancer molecular stratification dogma: Intratumoral heterogeneity undermines consensus molecular subtypes and potential diagnostic value in colorectal cancer. *Clin Cancer Res*.

DUPAGE, M. & BLUESTONE, J. A. 2016. Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease. *Nat Rev Immunol,* 16**,** 149-63.

DYSON, N. 1998. The regulation of E2F by pRB-family proteins. *Genes Dev,* 12**,** 2245-62.

EBISUNO, Y., TANAKA, T., KANEMITSU, N., KANDA, H., YAMAGUCHI, K., KAISHO, T., AKIRA, S. & MIYASAKA, M. 2003. Cutting edge: the B cell chemokine CXC chemokine ligand 13/B lymphocyte chemoattractant is expressed in the high endothelial venules of lymph nodes and Peyer's patches and affects B cell trafficking across high endothelial venules. *J Immunol,* 171**,** 1642-6.

EHRHARDT, G. R., HSU, J. T., GARTLAND, L., LEU, C. M., ZHANG, S., DAVIS, R. S. & COOPER, M. D. 2005. Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells. *J Exp Med,* 202**,** 783-91.

ELGUETA, R., BENSON, M. J., DE VRIES, V. C., WASIUK, A., GUO, Y. & NOELLE, R. J. 2009. Molecular mechanism and function of CD40/CD40L engagement in the immune system. *Immunol Rev,* 229**,** 152-72.

ENGEL, I., SEUMOIS, G., CHAVEZ, L., SAMANIEGO-CASTRUITA, D., WHITE, B., CHAWLA, A., MOCK, D., VIJAYANAND, P. & KRONENBERG, M. 2016. Innate-like functions of natural killer T cell subsets result from highly divergent gene programs. *Nat Immunol,* 17**,** 728-39.

ERDAG, G., SCHAEFER, J. T., SMOLKIN, M. E., DEACON, D. H., SHEA, S. M., DENGEL, L. T., PATTERSON, J. W. & SLINGLUFF, C. L., JR. 2012. Immunotype and immunohistologic characteristics of tumor-infiltrating immune cells are associated with clinical outcome in metastatic melanoma. *Cancer Res,* 72**,** 1070-80.

EXLEY, M. A., TAHIR, S. M., CHENG, O., SHAULOV, A., JOYCE, R., AVIGAN, D., SACKSTEIN, R. & BALK, S. P. 2001. A major fraction of human bone marrow lymphocytes are Th2-like CD1d-reactive T cells that can suppress mixed lymphocyte responses. *J Immunol,* 167**,** 5531-4.

FAN, X., QUEZADA, S. A., SEPULVEDA, M. A., SHARMA, P. & ALLISON, J. P. 2014. Engagement of the ICOS pathway markedly enhances efficacy of CTLA-4 blockade in cancer immunotherapy. *J Exp Med,* 211**,** 715-25.

FARBER, D. L., YUDANIN, N. A. & RESTIFO, N. P. 2014. Human memory T cells: generation, compartmentalization and homeostasis. *Nat Rev Immunol,* 14**,** 24-35.

FEHRENBACHER, L., SPIRA, A., BALLINGER, M., KOWANETZ, M., VANSTEENKISTE, J., MAZIERES, J., PARK, K., SMITH, D., ARTAL-CORTES, A., LEWANSKI, C., BRAITEH, F., WATERKAMP, D., HE, P., ZOU, W., CHEN, D. S., YI, J., SANDLER, A., RITTMEYER, A. & GROUP, P. S. 2016. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet,* 387**,** 1837-46.

FENG, X., WANG, H., TAKATA, H., DAY, T. J., WILLEN, J. & HU, H. 2011. Transcription factor Foxp1 exerts essential cell-intrinsic regulation of the quiescence of naive T cells. *Nat Immunol,* 12**,** 544-50.

FERRER, G., BOSCH, R., HODGSON, K., TEJERO, R., ROUE, G., COLOMER, D., MONTSERRAT, E. & MORENO, C. 2014. B cell activation through CD40 and IL4R ligation modulates the response of chronic lymphocytic leukaemia cells to BAFF and APRIL. *Br J Haematol,* 164**,** 570-8.

FINAK, G., MCDAVID, A., YAJIMA, M., DENG, J., GERSUK, V., SHALEK, A. K., SLICHTER, C. K., MILLER, H. W., MCELRATH, M. J., PRLIC, M., LINSLEY, P. S. & GOTTARDO, R. 2015. MAST: a flexible

statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol,* 16**,** 278.

FINBERG, R. W., WANG, J. P. & KURT-JONES, E. A. 2007. Toll like receptors and viruses. *Rev Med Virol,* 17**,** 35-43.

FLORES-BORJA, F., BOSMA, A., NG, D., REDDY, V., EHRENSTEIN, M. R., ISENBERG, D. A. & MAURI, C. 2013. CD19+CD24hiCD38hi B cells maintain regulatory T cells while limiting TH1 and TH17 differentiation. *Sci Transl Med,* 5**,** 173ra23.

FOURCADE J, S. Z., BENALLAOUA M, GUILLAUME P, LUESCHER IF, SANDER C, KIRKWOOD JM, KUCHROO V, ZAROUR HM. 2010. Upregulation of Tim-3 and PD-1 expression is associated with tumor antigen-specific CD8+ T cell dysfunction in melanoma patients. *J Exp Med.,* 207**,** 2175-86.

FRAZER, I. H. 2009. Interaction of human papillomaviruses with the host immune system: a well evolved relationship. *Virology,* 384**,** 410-4.

FRIDMAN, W. H., PAGES, F., SAUTES-FRIDMAN, C. & GALON, J. 2012. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer,* 12**,** 298-306.

FU, Y. X., HUANG, G., WANG, Y. & CHAPLIN, D. D. 1998. B lymphocytes induce the formation of follicular dendritic cell clusters in a lymphotoxin alpha-dependent fashion. *J Exp Med,* 187**,** 1009-18.

GAGLIANI, N., AMEZCUA VESELY, M. C., ISEPPON, A., BROCKMANN, L., XU, H., PALM, N. W., DE ZOETE, M. R., LICONA-LIMON, P., PAIVA, R. S., CHING, T., WEAVER, C., ZI, X., PAN, X., FAN, R., GARMIRE, L. X., COTTON, M. J., DRIER, Y., BERNSTEIN, B., GEGINAT, J., STOCKINGER, B., ESPLUGUES, E., HUBER, S. & FLAVELL, R. A. 2015. Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. *Nature,* 523**,** 221-5.

GALON, J., COSTES, A., SANCHEZ-CABO, F., KIRILOVSKY, A., MLECNIK, B., LAGORCE-PAGES, C., TOSOLINI, M., CAMUS, M., BERGER, A., WIND, P., ZINZINDOHOUE, F., BRUNEVAL, P., CUGNENC, P. H., TRAJANOSKI, Z., FRIDMAN, W. H. & PAGES, F. 2006. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science,* 313**,** 1960-4.

GANESAN, A. P., CLARKE, J., WOOD, O., GARRIDO-MARTIN, E. M., CHEE, S. J., MELLOWS, T., SAMANIEGO-CASTRUITA, D., SINGH, D., SEUMOIS, G., ALZETANI, A., WOO, E., FRIEDMANN, P. S., KING, E. V., THOMAS, G. J., SANCHEZ-ELSNER, T., VIJAYANAND, P. & OTTENSMEIER, C. H. 2017. Tissue-resident memory features are linked to the magnitude of cytotoxic T cell responses in human lung cancer. *Nat Immunol,* 18**,** 940-950.

GARG, A. D., COULIE, P. G., VAN DEN EYNDE, B. J. & AGOSTINIS, P. 2017. Integrating Next-Generation Dendritic Cell Vaccines into the Current Cancer Immunotherapy Landscape. *Trends Immunol,* 38**,** 577-593.

GARNETT, T. O. & DUERKSEN-HUGHES, P. J. 2006. Modulation of apoptosis by human papillomavirus (HPV) oncoproteins. *Arch Virol,* 151**,** 2321-35.

GARON, E. B., RIZVI, N. A., HUI, R., LEIGHL, N., BALMANOUKIAN, A. S., EDER, J. P., PATNAIK, A., AGGARWAL, C., GUBENS, M., HORN, L., CARCERENY, E., AHN, M. J., FELIP, E., LEE, J. S., HELLMANN, M. D., HAMID, O., GOLDMAN, J. W., SORIA, J. C., DOLLED-FILHART, M., RUTLEDGE, R. Z., ZHANG, J., LUNCEFORD, J. K., RANGWALA, R., LUBINIECKI, G. M., ROACH, C., EMANCIPATOR, K., GANDHI, L. & INVESTIGATORS, K.-. 2015. Pembrolizumab for the treatment of non-small-cell lung cancer. *N Engl J Med,* 372**,** 2018-28.

GASSER, S., ORSULIC, S., BROWN, E. J. & RAULET, D. H. 2005. The DNA damage pathway regulates innate immune system ligands of the NKG2D receptor. *Nature,* 436**,** 1186-90.

GATTO, D., PAUS, D., BASTEN, A., MACKAY, C. R. & BRINK, R. 2009. Guidance of B cells by the orphan G protein-coupled receptor EBI2 shapes humoral immune responses. *Immunity,* 31**,** 259-69.

GEISSMANN, F., MANZ, M. G., JUNG, S., SIEWEKE, M. H., MERAD, M. & LEY, K. 2010. Development of monocytes, macrophages, and dendritic cells. *Science,* 327**,** 656-61.

GENTLES, A. J., NEWMAN, A. M., LIU, C. L., BRATMAN, S. V., FENG, W., KIM, D., NAIR, V. S., XU, Y., KHUONG, A., HOANG, C. D., DIEHN, M., WEST, R. B., PLEVRITIS, S. K. & ALIZADEH, A. A. 2015. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med,* 21**,** 938-45.

GERLACH, C., VAN HEIJST, J. W., SWART, E., SIE, D., ARMSTRONG, N., KERKHOVEN, R. M., ZEHN, D., BEVAN, M. J., SCHEPERS, K. & SCHUMACHER, T. N. 2010. One naive T cell, multiple fates in CD8+ T cell differentiation. *J Exp Med,* 207**,** 1235-46.

GERMAIN, C., GNJATIC, S., TAMZALIT, F., KNOCKAERT, S., REMARK, R., GOC, J., LEPELLEY, A., BECHT, E., KATSAHIAN, S., BIZOUARD, G., VALIDIRE, P., DAMOTTE, D., ALIFANO, M., MAGDELEINAT, P., CREMER, I., TEILLAUD, J. L., FRIDMAN, W. H., SAUTES-FRIDMAN, C. & DIEU-NOSJEAN, M. C. 2014. 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**,** 832-44.

GHANIZADA, M., JAKOBSEN, K. K., GRONHOJ, C. & VON BUCHWALD, C. 2019. The effects of checkpoint inhibition on head and neck squamous cell carcinoma: A systematic review. *Oral Oncol,* 90**,** 67-73.

GIOVANNONE, N., LIANG, J., ANTONOPOULOS, A., GEDDES SWEENEY, J., KING, S. L., POCHEBIT, S. M., BHATTACHARYYA, N., LEE, G. S., DELL, A., WIDLUND, H. R., HASLAM, S. M. & DIMITROFF, C. J. 2018. Galectin-9 suppresses B cell receptor signaling and is regulated by I-branching of N-glycans. *Nat Commun,* 9**,** 3287.

GLIMELIUS, B. & LAHN, M. 2011. Window-of-opportunity trials to evaluate clinical activity of new molecular entities in oncology. *Ann Oncol,* 22**,** 1717-25.

GOBER, H. J., KISTOWSKA, M., ANGMAN, L., JENO, P., MORI, L. & DE LIBERO, G. 2003. Human T cell receptor gammadelta cells recognize endogenous mevalonate metabolites in tumor cells. *J Exp Med,* 197**,** 163-8.

GRIGORYEV, Y. A., KURIAN, S. M., AVNUR, Z., BORIE, D., DENG, J., CAMPBELL, D., SUNG, J., NIKOLCHEVA, T., QUINN, A., SCHULMAN, H., PENG, S. L., SCHAFFER, R., FISHER, J., MONDALA, T., HEAD, S., FLECHNER, S. M., KANTOR, A. B., MARSH, C. & SALOMON, D. R. 2010. Deconvoluting post-transplant immunity: cell subset-specific mapping reveals pathways for activation and expansion of memory T, monocytes and B cells. *PLoS One,* 5**,** e13358.

GROS, A., ROBBINS, P. F., YAO, X., LI, Y. F., TURCOTTE, S., TRAN, E., WUNDERLICH, J. R., MIXON, A., FARID, S., DUDLEY, M. E., HANADA, K., ALMEIDA, J. R., DARKO, S., DOUEK, D. C., YANG, J. C. & ROSENBERG, S. A. 2014. PD-1 identifies the patient-specific CD8(+) tumor-reactive repertoire infiltrating human tumors. *J Clin Invest,* 124**,** 2246-59.

GROS A, R. P., YAO X, LI YF, TURCOTTE S, TRAN E, WUNDERLICH JR, MIXON A, FARID S, DUDLEY ME, HANADA K, ALMEIDA JR, DARKO S, DOUEK DC, YANG JC, ROSENBERG SA. 2014. PD-1

identifies the patient-specific CD8+ tumor-reactive repertoire infiltrating human tumors. ***J Clin Invest.*** *,* 124**,** 2246-59.

GUERREIRO-CACAIS, A. O., LEVITSKAYA, J. & LEVITSKY, V. 2010. B cell receptor triggering sensitizes human B cells to TRAIL-induced apoptosis. *J Leukoc Biol,* 88**,** 937-45.

GUILLAUMET-ADKINS, A., RODRIGUEZ-ESTEBAN, G., MEREU, E., MENDEZ-LAGO, M., JAITIN, D. A., VILLANUEVA, A., VIDAL, A., MARTINEZ-MARTI, A., FELIP, E., VIVANCOS, A., KEREN-SHAUL, H., HEATH, S., GUT, M., AMIT, I., GUT, I. & HEYN, H. 2017. Single-cell transcriptome conservation in cryopreserved cells and tissues. *Genome Biol,* 18**,** 45.

GUTZEIT, C., CHEN, K. & CERUTTI, A. 2018. The enigmatic function of IgD: some answers at last. *Eur J Immunol,* 48**,** 1101-1113.

GYANCHANDANI, R., LIN, Y., LIN, H. M., COOPER, K. L., NORMOLLE, D. P., BRUFSKY, A. M., FASTUCA, M., CROSSON, W., OESTERREICH, S., DAVIDSON, N. E., BHARGAVA, R., DABBS, D. J. & LEE, A. V. 2016. Intra-tumor heterogeneity affects gene expression profile test prognostic risk stratification in early breast cancer. *Clin Cancer Res*.

HAGN, M. & JAHRSDORFER, B. 2012. Why do human B cells secrete granzyme B? Insights into a novel B-cell differentiation pathway. *Oncoimmunology,* 1**,** 1368-1375.

HAGN, M., SONTHEIMER, K., DAHLKE, K., BRUEGGEMANN, S., KALTENMEIER, C., BEYER, T., HOFMANN, S., LUNOV, O., BARTH, T. F., FABRICIUS, D., TRON, K., NIENHAUS, G. U., SIMMET, T., SCHREZENMEIER, H. & JAHRSDORFER, B. 2012. Human B cells differentiate into granzyme B-secreting cytotoxic B lymphocytes upon incomplete T-cell help. *Immunol Cell Biol,* 90**,** 457-67.

HALLER, O., KOCHS, G. & WEBER, F. 2006. The interferon response circuit: induction and suppression by pathogenic viruses. *Virology,* 344**,** 119-30.

HAMILTON, S. E. & JAMESON, S. C. 2012. CD8 T cell quiescence revisited. *Trends Immunol,* 33**,** 224-30.

HANAHAN, D. & WEINBERG, R. A. 2000. The hallmarks of cancer. *Cell,* 100**,** 57-70.

HANAHAN, D. & WEINBERG, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

HANLEY, C. J., MELLONE, M., FORD, K., THIRDBOROUGH, S. M., MELLOWS, T., FRAMPTON, S. J., SMITH, D. M., HARDEN, E., SZYNDRALEWIEZ, C., BULLOCK, M., NOBLE, F., MOUTASIM, K. A., KING, E. V., VIJAYANAND, P., MIRNEZAMI, A. H., UNDERWOOD, T. J., OTTENSMEIER, C. H. & THOMAS, G. J. 2018. Targeting the Myofibroblastic Cancer-Associated Fibroblast Phenotype Through Inhibition of NOX4. *J Natl Cancer Inst,* 110.

HARTMANN, K., SCHLOMBS, K., LAIBLE, M., GURTLER, C., SCHMIDT, M., SAHIN, U. & LEHR, H. A. 2018. Robustness of biomarker determination in breast cancer by RT-qPCR: impact of tumor cell content, DCIS and non-neoplastic breast tissue. *Diagn Pathol,* 13**,** 83.

HARTY, J. T., TVINNEREIM, A. R. & WHITE, D. W. 2000. CD8+ T cell effector mechanisms in resistance to infection. *Annu Rev Immunol,* 18**,** 275-308.

HAYFLICK, L. & MOORHEAD, P. S. 1961. The serial cultivation of human diploid cell strains. *Exp Cell Res,* 25**,** 585-621.

HEDRICK, S. M. 2008. Thymus lineage commitment: a single switch. *Immunity,* 28**,** 297-9.

HEDRICK, S. M., COHEN, D. I., NIELSEN, E. A. & DAVIS, M. M. 1984. Isolation of cDNA clones encoding T cell-specific membrane-associated proteins. *Nature,* 308**,** 149-53.

HELMINK, B. A., REDDY, S. M., GAO, J., ZHANG, S., BASAR, R., THAKUR, R., YIZHAK, K., SADE-FELDMAN, M., BLANDO, J., HAN, G., GOPALAKRISHNAN, V., XI, Y., ZHAO, H., AMARIA, R. N., TAWBI, H. A., COGDILL, A. P., LIU, W., LEBLEU, V. S., KUGERATSKI, F. G., PATEL, S., DAVIES, M. A., HWU, P., LEE, J. E., GERSHENWALD, J. E., LUCCI, A., ARORA, R., WOODMAN, S., KEUNG, E. Z., GAUDREAU, P. O., REUBEN, A., SPENCER, C. N., BURTON, E. M., HAYDU, L. E., LAZAR, A. J., ZAPASSODI, R., HUDGENS, C. W., LEDESMA, D. A., ONG, S., BAILEY, M., WARREN, S., RAO, D., KRIJGSMAN, O., ROZEMAN, E. A., PEEPER, D., BLANK, C. U., SCHUMACHER, T. N., BUTTERFIELD, L. H., ZELAZOWSKA, M. A., MCBRIDE, K. M., KALLURI, R., ALLISON, J., PETITPREZ, F., FRIDMAN, W. H., SAUTES-FRIDMAN, C., HACOHEN, N., REZVANI, K., SHARMA, P., TETZLAFF, M. T., WANG, L. & WARGO, J. A. 2020. B cells and tertiary lymphoid structures promote immunotherapy response. *Nature,* 577**,** 549-555.

HEMON, P., RENAUDINEAU, Y., DEBANT, M., LE GOUX, N., MUKHERJEE, S., BROOKS, W. & MIGNEN, O. 2017. Calcium Signaling: From Normal B Cell Development to Tolerance Breakdown and Autoimmunity. *Clin Rev Allergy Immunol,* 53**,** 141-165.

HENSLEY, C. T., FAUBERT, B., YUAN, Q., LEV-COHAIN, N., JIN, E., KIM, J., JIANG, L., KO, B., SKELTON, R., LOUDAT, L., WODZAK, M., KLIMKO, C., MCMILLAN, E., BUTT, Y., NI, M., OLIVER, D., TORREALBA, J., MALLOY, C. R., KERNSTINE, K., LENKINSKI, R. E. & DEBERARDINIS, R. J. 2016. Metabolic Heterogeneity in Human Lung Tumors. *Cell,* 164**,** 681-94.

HERBST, R. S., MORGENSZTERN, D. & BOSHOFF, C. 2018. The biology and management of non-small cell lung cancer. *Nature,* 553**,** 446-454.

HETZ, C. 2012. The unfolded protein response: controlling cell fate decisions under ER stress and beyond. *Nat Rev Mol Cell Biol,* 13**,** 89-102.

HEUSINKVELD, M., GOEDEMANS, R., BRIET, R. J., GELDERBLOM, H., NORTIER, J. W., GORTER, A., SMIT, V. T., LANGEVELD, A. P., JANSEN, J. C. & VAN DER BURG, S. H. 2012. Systemic and local human papillomavirus 16-specific T-cell immunity in patients with head and neck cancer. *Int J Cancer,* 131**,** E74-85.

HIRAOKA, K., MIYAMOTO, M., CHO, Y., SUZUOKI, M., OSHIKIRI, T., NAKAKUBO, Y., ITOH, T., OHBUCHI, T., KONDO, S. & KATOH, H. 2006. Concurrent infiltration by CD8+ T cells and CD4+ T cells is a favourable prognostic factor in non-small-cell lung carcinoma. *Br J Cancer,* 94**,** 275-80.

HODI, F. S., O'DAY, S. J., MCDERMOTT, D. F., WEBER, R. W., SOSMAN, J. A., HAANEN, J. B., GONZALEZ, R., ROBERT, C., SCHADENDORF, D., HASSEL, J. C., AKERLEY, W., VAN DEN EERTWEGH, A. J., LUTZKY, J., LORIGAN, P., VAUBEL, J. M., LINETTE, G. P., HOGG, D., OTTENSMEIER, C. H., LEBBE, C., PESCHEL, C., QUIRT, I., CLARK, J. I., WOLCHOK, J. D., WEBER, J. S., TIAN, J., YELLIN, M. J., NICHOL, G. M., HOOS, A. & URBA, W. J. 2010. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med,* 363**,** 711-23.

HOFFMANN, J. C., KRUGER, H., ZIELEN, S., BAYER, B. & ZEIDLER, H. 1998. Human b cell differentiation: dependence on interactions with monocytes and T lymphocytes via CD40, CD80 (B7.1), and the CD2-Ligands CD48 and CD58 (LFA-3). *Cell Biol Int,* 22**,** 21-9.

HOGAN, S. P., ROSENBERG, H. F., MOQBEL, R., PHIPPS, S., FOSTER, P. S., LACY, P., KAY, A. B. & ROTHENBERG, M. E. 2008. Eosinophils: biological properties and role in health and disease. *Clin Exp Allergy,* 38**,** 709-50.

HOLLERN, D. P., XU, N., THENNAVAN, A., GLODOWSKI, C., GARCIA-RECIO, S., MOTT, K. R., HE, X., GARAY, J. P., CAREY-EWEND, K., MARRON, D., FORD, J., LIU, S., VICK, S. C., MARTIN, M., PARKER, J. S., VINCENT, B. G., SERODY, J. S. & PEROU, C. M. 2019. B Cells and T Follicular Helper Cells Mediate Response to Checkpoint Inhibitors in High Mutation Burden Mouse Models of Breast Cancer. *Cell,* 179**,** 1191-1206 e21.

HOLMANNOVA, D., KOLACKOVA, M., KONDELKOVA, K., KUNES, P., KREJSEK, J. & ANDRYS, C. 2012. CD200/CD200R paired potent inhibitory molecules regulating immune and inflammatory responses; Part I: CD200/CD200R structure, activation, and function. *Acta Medica (Hradec Kralove),* 55**,** 12-7.

HONDA, K., YANAI, H., NEGISHI, H., ASAGIRI, M., SATO, M., MIZUTANI, T., SHIMADA, N., OHBA, Y., TAKAOKA, A., YOSHIDA, N. & TANIGUCHI, T. 2005. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature,* 434**,** 772-7.

HONG, J. J., AMANCHA, P. K., ROGERS, K., ANSARI, A. A. & VILLINGER, F. 2013. Re-evaluation of PD-1 expression by T cells as a marker for immune exhaustion during SIV infection. *PLoS One,* 8**,** e60186.

HORNE, Z. D., JACK, R., GRAY, Z. T., SIEGFRIED, J. M., WILSON, D. O., YOUSEM, S. A., NASON, K. S., LANDRENEAU, R. J., LUKETICH, J. D. & SCHUCHERT, M. J. 2011. Increased levels of tumor-infiltrating lymphocytes are associated with improved recurrence-free survival in stage 1A non-small-cell lung cancer. *J Surg Res,* 171**,** 1-5.

HU, C. C., DOUGAN, S. K., MCGEHEE, A. M., LOVE, J. C. & PLOEGH, H. L. 2009. XBP-1 regulates signal transduction, transcription factors and bone marrow colonization in B cells. *EMBO J,* 28**,** 1624-36.

HU, H., WANG, B., BORDE, M., NARDONE, J., MAIKA, S., ALLRED, L., TUCKER, P. W. & RAO, A. 2006. Foxp1 is an essential transcriptional regulator of B cell development. *Nat Immunol,* 7**,** 819-26.

HUANG, J., BRAMESHUBER, M., ZENG, X., XIE, J., LI, Q. J., CHIEN, Y. H., VALITUTTI, S. & DAVIS, M. M. 2013. A single peptide-major histocompatibility complex ligand triggers digital cytokine secretion in CD4(+) T cells. *Immunity,* 39**,** 846-57.

IGLESIA, M. D., VINCENT, B. G., PARKER, J. S., HOADLEY, K. A., CAREY, L. A., PEROU, C. M. & SERODY, J. S. 2014. Prognostic B-cell signatures using mRNA-seq in patients with subtype-specific breast and ovarian cancer. *Clin Cancer Res,* 20**,** 3818-29.

ISHIKAWA, H. & BARBER, G. N. 2008. STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. *Nature,* 455**,** 674-8.

IWATA, Y., MATSUSHITA, T., HORIKAWA, M., DILILLO, D. J., YANABA, K., VENTURI, G. M., SZABOLCS, P. M., BERNSTEIN, S. H., MAGRO, C. M., WILLIAMS, A. D., HALL, R. P., ST CLAIR, E. W. & TEDDER, T. F. 2011. Characterization of a rare IL-10-competent B-cell subset in humans that parallels mouse regulatory B10 cells. *Blood,* 117**,** 530-41.

JAMAL-HANJANI, M., QUEZADA, S. A., LARKIN, J. & SWANTON, C. 2015. Translational implications of tumor heterogeneity. *Clin Cancer Res,* 21**,** 1258-66.

JHOU, J. P., CHEN, S. J., HUANG, H. Y., LIN, W. W., HUANG, D. Y. & TZENG, S. J. 2017. Upregulation of FcgammaRIIB by resveratrol via NF-kappaB activation reduces B-cell numbers and ameliorates lupus. *Exp Mol Med,* 49**,** e381.

JONES, S. A., WHITE, C. A., ROBB, L., ALEXANDER, W. S. & TARLINTON, D. M. 2011. SOCS3 deletion in B cells alters cytokine responses and germinal center output. *J Immunol,* 187**,** 6318-26.

JUNE, C. H., WARSHAUER, J. T. & BLUESTONE, J. A. 2017. Is autoimmunity the Achilles' heel of cancer immunotherapy? *Nat Med,* 23**,** 540-547.

JUNG, A. C., BRIOLAT, J., MILLON, R., DE REYNIES, A., RICKMAN, D., THOMAS, E., ABECASSIS, J., CLAVEL, C. & WASYLYK, B. 2010. Biological and clinical relevance of transcriptionally active human papillomavirus (HPV) infection in oropharynx squamous cell carcinoma. *Int J Cancer,* 126**,** 1882-94.

KAMBUROV, A., PENTCHEV, K., GALICKA, H., WIERLING, C., LEHRACH, H. & HERWIG, R. 2011. ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Res,* 39**,** D712-7.

KAMINSKI, D. A., WEI, C., QIAN, Y., ROSENBERG, A. F. & SANZ, I. 2012. Advances in human B cell phenotypic profiling. *Front Immunol,* 3**,** 302.

KARAGIANNIS, P., GILBERT, A. E., JOSEPHS, D. H., ALI, N., DODEV, T., SAUL, L., CORREA, I., ROBERTS, L., BEDDOWES, E., KOERS, A., HOBBS, C., FERREIRA, S., GEH, J. L., HEALY, C., HARRIES, M., ACLAND, K. M., BLOWER, P. J., MITCHELL, T., FEAR, D. J., SPICER, J. F., LACY, K. E., NESTLE, F. O. & KARAGIANNIS, S. N. 2013. IgG4 subclass antibodies impair antitumor immunity in melanoma. *J Clin Invest,* 123**,** 1457-74.

KATIKANENI, D. S. & JIN, L. 2019. B cell MHC class II signaling: A story of life and death. *Hum Immunol,* 80**,** 37-43.

KAWAI, T. & AKIRA, S. 2010. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat Immunol,* 11**,** 373-84.

KEIR, M. E., BUTTE, M. J., FREEMAN, G. J. & SHARPE, A. H. 2008. PD-1 and its ligands in tolerance and immunity. *Annu Rev Immunol,* 26**,** 677-704.

KERFOOT, S. M., YAARI, G., PATEL, J. R., JOHNSON, K. L., GONZALEZ, D. G., KLEINSTEIN, S. H. & HABERMAN, A. M. 2011. Germinal center B cell and T follicular helper cell development initiates in the interfollicular zone. *Immunity,* 34**,** 947-60.

KILIC, A., LANDRENEAU, R. J., LUKETICH, J. D., PENNATHUR, A. & SCHUCHERT, M. J. 2011. Density of tumor-infiltrating lymphocytes correlates with disease recurrence and survival in patients with large non-small-cell lung cancer tumors. *J Surg Res,* 167**,** 207-10.

KING, E. V., OTTENSMEIER, C. H. & THOMAS, G. J. 2014. The immune response in HPV+ oropharyngeal cancer. *Oncoimmunology,* 3**,** e27254.

KLINGELHUTZ, A. J., FOSTER, S. A. & MCDOUGALL, J. K. 1996. Telomerase activation by the E6 gene product of human papillomavirus type 16. *Nature,* 380**,** 79-82.

KOBAYASHI, A., DARRAGH, T., HERNDIER, B., ANASTOS, K., MINKOFF, H., COHEN, M., YOUNG, M., LEVINE, A., GRANT, L. A., HYUN, W., WEINBERG, V., GREENBLATT, R. & SMITH-MCCUNE, K. 2002. Lymphoid follicles are generated in high-grade cervical dysplasia and have differing characteristics depending on HIV status. *Am J Pathol,* 160**,** 151-64.

KOLSET, S. O. & PEJLER, G. 2011. Serglycin: a structural and functional chameleon with wide impact on immune cells. *J Immunol,* 187**,** 4927-33.

KOMETANI, K., NAKAGAWA, R., SHINNAKASU, R., KAJI, T., RYBOUCHKIN, A., MORIYAMA, S., FURUKAWA, K., KOSEKI, H., TAKEMORI, T. & KUROSAKI, T. 2013. Repression of the transcription factor Bach2 contributes to predisposition of IgG1 memory B cells toward plasma cell differentiation. *Immunity,* 39**,** 136-47.

KONDO, M., SAKUTA, K., NOGUCHI, A., ARIYOSHI, N., SATO, K., SATO, S., SATO, K., HOSOI, A., NAKAJIMA, J., YOSHIDA, Y., SHIRAISHI, K., NAKAGAWA, K. & KAKIMI, K. 2008. Zoledronate facilitates large-scale ex vivo expansion of functional gammadelta T cells from cancer patients for use in adoptive immunotherapy. *Cytotherapy,* 10**,** 842-56.

KORPETINOU, A., SKANDALIS, S. S., LABROPOULOU, V. T., SMIRLAKI, G., NOULAS, A., KARAMANOS, N. K. & THEOCHARIS, A. D. 2014. Serglycin: at the crossroad of inflammation and malignancy. *Front Oncol,* 3**,** 327.

KORZENIEWSKI, N., SPARDY, N., DUENSING, A. & DUENSING, S. 2011. Genomic instability and cancer: lessons learned from human papillomaviruses. *Cancer Lett,* 305**,** 113-22.

KROEGER, D. R., MILNE, K. & NELSON, B. H. 2016. Tumor-Infiltrating Plasma Cells Are Associated with Tertiary Lymphoid Structures, Cytolytic T-Cell Responses, and Superior Prognosis in Ovarian Cancer. *Clin Cancer Res,* 22**,** 3005-15.

KVISTBORG, P., PHILIPS, D., KELDERMAN, S., HAGEMAN, L., OTTENSMEIER, C., JOSEPH-PIETRAS, D., WELTERS, M. J., VAN DER BURG, S., KAPITEIJN, E., MICHIELIN, O., ROMANO, E., LINNEMANN, C., SPEISER, D., BLANK, C., HAANEN, J. B. & SCHUMACHER, T. N. 2014. Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. *Sci Transl Med,* 6**,** 254ra128.

LAIBLE, M., SCHLOMBS, K., KAISER, K., VELTRUP, E., HERLEIN, S., LAKIS, S., STOHR, R., EIDT, S., HARTMANN, A., WIRTZ, R. M. & SAHIN, U. 2016. Technical validation of an RT-qPCR in vitro diagnostic test system for the determination of breast cancer molecular subtypes by quantification of ERBB2, ESR1, PGR and MKI67 mRNA levels from formalin-fixed paraffin-embedded breast tumor specimens. *BMC Cancer,* 16**,** 398.

LAMBRECHTS, D., WAUTERS, E., BOECKX, B., AIBAR, S., NITTNER, D., BURTON, O., BASSEZ, A., DECALUWE, H., PIRCHER, A., VAN DEN EYNDE, K., WEYNAND, B., VERBEKEN, E., DE LEYN, P., LISTON, A., VANSTEENKISTE, J., CARMELIET, P., AERTS, S. & THIENPONT, B. 2018. Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat Med,* 24**,** 1277-1289.

LASSEN, P., ERIKSEN, J. G., HAMILTON-DUTOIT, S., TRAMM, T., ALSNER, J. & OVERGAARD, J. 2009. Effect of HPV-associated p16INK4A expression on response to radiotherapy and survival in squamous cell carcinoma of the head and neck. *J Clin Oncol,* 27**,** 1992-8.

LAWRENCE, C. W., REAM, R. M. & BRACIALE, T. J. 2005. Frequency, specificity, and sites of expansion of CD8+ T cells during primary pulmonary influenza virus infection. *J Immunol,* 174**,** 5332-40.

LE NAOUR, F., BRICHORY, F., MISEK, D. E., BRECHOT, C., HANASH, S. M. & BERETTA, L. 2002. A distinct repertoire of autoantibodies in hepatocellular carcinoma identified by proteomic analysis. *Mol Cell Proteomics,* 1**,** 197-203.

LECHNER, A., SCHLOSSER, H., ROTHSCHILD, S. I., THELEN, M., REUTER, S., ZENTIS, P., SHIMABUKURO-VORNHAGEN, A., THEURICH, S., WENNHOLD, K., GARCIA-MARQUEZ, M., THARUN, L., QUAAS, A., SCHAUSS, A., ISENSEE, J., HUCHO, T., HUEBBERS, C., VON BERGWELT-BAILDON, M. & BEUTNER, D. 2017. Characterization of tumor-associated T-lymphocyte subsets and immune checkpoint molecules in head and neck squamous cell carcinoma. *Oncotarget,* 8**,** 44418-44433.

LEDGERWOOD, L. G., KUMAR, D., ETEROVIC, A. K., WICK, J., CHEN, K., ZHAO, H., TAZI, L., MANNA, P., KERLEY, S., JOSHI, R., WANG, L., CHIOSEA, S. I., GARNETT, J. D., TSUE, T. T., CHIEN, J.,

MILLS, G. B., GRANDIS, J. R. & THOMAS, S. M. 2016. The degree of intratumor mutational heterogeneity varies by primary tumor sub-site. *Oncotarget,* 7**,** 27185-98.

LEEK, J. T., JOHNSON, W. E., PARKER, H. S., JAFFE, A. E. & STOREY, J. D. 2012. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics,* 28**,** 882-3.

LI, X. J., ONG, C. K., CAO, Y., XIANG, Y. Q., SHAO, J. Y., OOI, A., PENG, L. X., LU, W. H., ZHANG, Z., PETILLO, D., QIN, L., BAO, Y. N., ZHENG, F. J., CHIA, C. S., IYER, N. G., KANG, T. B., ZENG, Y. X., SOO, K. C., TRENT, J. M., TEH, B. T. & QIAN, C. N. 2011. Serglycin is a theranostic target in nasopharyngeal carcinoma that promotes metastasis. *Cancer Res,* 71**,** 3162-72.

LIANG, C., MARSIT, C. J., MCCLEAN, M. D., NELSON, H. H., CHRISTENSEN, B. C., HADDAD, R. I., CLARK, J. R., WEIN, R. O., GRILLONE, G. A., HOUSEMAN, E. A., HALEC, G., WATERBOER, T., PAWLITA, M., KRANE, J. F. & KELSEY, K. T. 2012. Biomarkers of HPV in head and neck squamous cell carcinoma. *Cancer Res,* 72**,** 5004-13.

LINDNER, S., DAHLKE, K., SONTHEIMER, K., HAGN, M., KALTENMEIER, C., BARTH, T. F., BEYER, T., REISTER, F., FABRICIUS, D., LOTFI, R., LUNOV, O., NIENHAUS, G. U., SIMMET, T., KREIENBERG, R., MOLLER, P., SCHREZENMEIER, H. & JAHRSDORFER, B. 2013. Interleukin 21-induced granzyme B-expressing B cells infiltrate tumors and regulate T cells. *Cancer Res,* 73**,** 2468-79.

LINSLEY, P. S., CLARK, E. A. & LEDBETTER, J. A. 1990. T-cell antigen CD28 mediates adhesion with B cells by interacting with activation antigen B7/BB-1. *Proc Natl Acad Sci U S A,* 87**,** 5031-5.

LIU, D., XU, H., SHIH, C., WAN, Z., MA, X., MA, W., LUO, D. & QI, H. 2014. T-B-cell entanglement and ICOSL-driven feed-forward regulation of germinal centre reaction. *Nature*.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods,* 25**,** 402-8.

LIZOTTE, P. H., IVANOVA, E. V., AWAD, M. M., JONES, R. E., KEOGH, L., LIU, H., DRIES, R., ALMONTE, C., HERTER-SPRIE, G. S., SANTOS, A., FEENEY, N. B., PAWELETZ, C. P., KULKARNI, M. M., BASS, A. J., RUSTGI, A. K., YUAN, G. C., KUFE, D. W., JANNE, P. A., HAMMERMAN, P. S., SHOLL, L. M., HODI, F. S., RICHARDS, W. G., BUENO, R., ENGLISH, J. M., BITTINGER, M. A. & WONG, K. K. 2016. Multiparametric profiling of non-small-cell lung cancers reveals distinct immunophenotypes. *JCI Insight,* 1**,** e89014.

MACKAY, F. & SCHNEIDER, P. 2008. TACI, an enigmatic BAFF/APRIL receptor, with new unappreciated biochemical and biological properties. *Cytokine Growth Factor Rev,* 19**,** 263-76.

MACKAY, L. K., RAHIMPOUR, A., MA, J. Z., COLLINS, N., STOCK, A. T., HAFON, M. L., VEGA-RAMOS, J., LAUZURICA, P., MUELLER, S. N., STEFANOVIC, T., TSCHARKE, D. C., HEATH, W. R., INOUYE, M., CARBONE, F. R. & GEBHARDT, T. 2013. The developmental pathway for CD103(+)CD8+ tissue-resident memory T cells of skin. *Nat Immunol,* 14**,** 1294-301.

MACLENNAN, I. C., TOELLNER, K. M., CUNNINGHAM, A. F., SERRE, K., SZE, D. M., ZUNIGA, E., COOK, M. C. & VINUESA, C. G. 2003. Extrafollicular antibody responses. *Immunol Rev,* 194**,** 8-18.

MACOSKO, E. Z., BASU, A., SATIJA, R., NEMESH, J., SHEKHAR, K., GOLDMAN, M., TIROSH, I., BIALAS, A. R., KAMITAKI, N., MARTERSTECK, E. M., TROMBETTA, J. J., WEITZ, D. A., SANES, J. R., SHALEK, A. K., REGEV, A. & MCCARROLL, S. A. 2015. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell,* 161**,** 1202-1214.

MAECKER, H. T., MCCOY, J. P. & NUSSENBLATT, R. 2012. Standardizing immunophenotyping for the Human Immunology Project. *Nat Rev Immunol,* 12**,** 191-200.

MAHMOUD, S. M., PAISH, E. C., POWE, D. G., MACMILLAN, R. D., GRAINGE, M. J., LEE, A. H., ELLIS, I. O. & GREEN, A. R. 2011. Tumor-infiltrating CD8+ lymphocytes predict clinical outcome in breast cancer. *J Clin Oncol,* 29**,** 1949-55.

MAJCHRZAK, K., NELSON, M. H., BAILEY, S. R., BOWERS, J. S., YU, X. Z., RUBINSTEIN, M. P., HIMES, R. A. & PAULOS, C. M. 2016. Exploiting IL-17-producing CD4+ and CD8+ T cells to improve cancer immunotherapy in the clinic. *Cancer Immunol Immunother,* 65**,** 247-59.

MALEK, T. R. & CASTRO, I. 2010. Interleukin-2 receptor signaling: at the interface between tolerance and immunity. *Immunity,* 33**,** 153-65.

MALISSEN, B., GREGOIRE, C., MALISSEN, M. & RONCAGALLI, R. 2014. Integrative biology of T cell activation. *Nat Immunol,* 15**,** 790-7.

MANTOVANI, A., SOZZANI, S., LOCATI, M., ALLAVENA, P. & SICA, A. 2002. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol,* 23**,** 549-55.

MARSH, D., SUCHAK, K., MOUTASIM, K. A., VALLATH, S., HOPPER, C., JERJES, W., UPILE, T., KALAVREZOS, N., VIOLETTE, S. M., WEINREB, P. H., CHESTER, K. A., CHANA, J. S., MARSHALL, J. F., HART, I. R., HACKSHAW, A. K., PIPER, K. & THOMAS, G. J. 2011. Stromal features are predictive of disease mortality in oral cancer patients. *J Pathol,* 223**,** 470-81.

MARTIN, F. & CHAN, A. C. 2006. B cell immunobiology in disease: evolving concepts from the clinic. *Annu Rev Immunol,* 24**,** 467-96.

MARTIN, F., OLIVER, A. M. & KEARNEY, J. F. 2001. Marginal zone and B1 B cells unite in the early response against T-independent blood-borne particulate antigens. *Immunity,* 14**,** 617-29.

MARTIN-FONTECHA, A., THOMSEN, L. L., BRETT, S., GERARD, C., LIPP, M., LANZAVECCHIA, A. & SALLUSTO, F. 2004. Induced recruitment of NK cells to lymph nodes provides IFN-gamma for T(H)1 priming. *Nat Immunol,* 5**,** 1260-5.

MARTIN-OROZCO, N., LI, Y., WANG, Y., LIU, S., HWU, P., LIU, Y. J., DONG, C. & RADVANYI, L. 2010. Melanoma cells express ICOS ligand to promote the activation and expansion of T-regulatory cells. *Cancer Res,* 70**,** 9581-90.

MATHIS, D. & BENOIST, C. 2009. Aire. *Annu Rev Immunol,* 27**,** 287-312.

MATSUI, K., BONIFACE, J. J., REAY, P. A., SCHILD, H., FAZEKAS DE ST GROTH, B. & DAVIS, M. M. 1991. Low affinity interaction of peptide-MHC complexes with T cell receptors. *Science,* 254**,** 1788-91.

MATSUI, K., BONIFACE, J. J., STEFFNER, P., REAY, P. A. & DAVIS, M. M. 1994. Kinetics of T-cell receptor binding to peptide/I-Ek complexes: correlation of the dissociation rate with T-cell responsiveness. *Proc Natl Acad Sci U S A,* 91**,** 12862-6.

MATSUMOTO, M., BABA, A., YOKOTA, T., NISHIKAWA, H., OHKAWA, Y., KAYAMA, H., KALLIES, A., NUTT, S. L., SAKAGUCHI, S., TAKEDA, K., KUROSAKI, T. & BABA, Y. 2014. Interleukin-10-producing plasmablasts exert regulatory function in autoimmune inflammation. *Immunity,* 41**,** 1040-51.

MAURI, C. & MENON, M. 2015. The expanding family of regulatory B cells. *Int Immunol,* 27**,** 479-86.

MCBRYAN, J., FAGAN, A., MCCARTAN, D., BANE, F. T., VARESLIJA, D., COCCHIGLIA, S., BYRNE, C., BOLGER, J., MCILROY, M., HUDSON, L., TIBBITTS, P., GAORA, P. O., HILL, A. D. & YOUNG, L. S. 2015. Transcriptomic Profiling of Sequential Tumors from Breast Cancer Patients Provides a Global View of Metastatic Expression Changes Following Endocrine Therapy. *Clin Cancer Res,* 21**,** 5371-9.

MCCANN, K. J., MANDER, A., CAZALY, A., CHUDLEY, L., STASAKOVA, J., THIRDBOROUGH, S., KING, A., LLOYD-EVANS, P., BUXTON, E., EDWARDS, C., HALFORD, S., BATEMAN, A., O'CALLAGHAN, A., CLIVE, S., ANTHONEY, A., JODRELL, D. I., WEINSCHENK, T., SIMON, P., SAHIN, U., THOMAS, G. J., STEVENSON, F. K. & OTTENSMEIER, C. H. 2016. Targeting Carcinoembryonic Antigen with DNA Vaccination: On-Target Adverse Events Link with Immunologic and Clinical Outcomes. *Clin Cancer Res,* 22**,** 4827-4836.

MCGRANAHAN, N., FURNESS, A. J., ROSENTHAL, R., RAMSKOV, S., LYNGAA, R., SAINI, S. K., JAMAL-HANJANI, M., WILSON, G. A., BIRKBAK, N. J., HILEY, C. T., WATKINS, T. B., SHAFI, S., MURUGAESU, N., MITTER, R., AKARCA, A. U., LINARES, J., MARAFIOTI, T., HENRY, J. Y., VAN ALLEN, E. M., MIAO, D., SCHILLING, B., SCHADENDORF, D., GARRAWAY, L. A., MAKAROV, V., RIZVI, N. A., SNYDER, A., HELLMANN, M. D., MERGHOUB, T., WOLCHOK, J. D., SHUKLA, S. A., WU, C. J., PEGGS, K. S., CHAN, T. A., HADRUP, S. R., QUEZADA, S. A. & SWANTON, C. 2016. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science,* 351**,** 1463-9.

MCHEYZER-WILLIAMS, L. J., MILPIED, P. J., OKITSU, S. L. & MCHEYZER-WILLIAMS, M. G. 2015. Class-switched memory B cells remodel BCRs within secondary germinal centers. *Nat Immunol,* 16**,** 296-305.

MEI, Z., HUANG, J., QIAO, B. & LAM, A. K. 2020. Immune checkpoint pathways in immunotherapy for head and neck squamous cell carcinoma. *Int J Oral Sci,* 12**,** 16.

MELCHERS, F. 2005. The pre-B-cell receptor: selector of fitting immunoglobulin heavy chains for the B-cell repertoire. *Nat Rev Immunol,* 5**,** 578-84.

MELERO, I., GAUDERNACK, G., GERRITSEN, W., HUBER, C., PARMIANI, G., SCHOLL, S., THATCHER, N., WAGSTAFF, J., ZIELINSKI, C., FAULKNER, I. & MELLSTEDT, H. 2014. Therapeutic vaccines for cancer: an overview of clinical trials. *Nat Rev Clin Oncol,* 11**,** 509-24.

MENDEZ-HUERGO, S. P., BLIDNER, A. G. & RABINOVICH, G. A. 2017. Galectins: emerging regulatory checkpoints linking tumor immunity and angiogenesis. *Curr Opin Immunol,* 45**,** 8-15.

MILANI, V., NOESSNER, E., GHOSE, S., KUPPNER, M., AHRENS, B., SCHARNER, A., GASTPAR, R. & ISSELS, R. D. 2002. Heat shock protein 70: role in antigen presentation and immune stimulation. *Int J Hyperthermia,* 18**,** 563-75.

MIRSADRAEE, S., OSWAL, D., ALIZADEH, Y., CAULO, A. & VAN BEEK, E., JR. 2012. The 7th lung cancer TNM classification and staging system: Review of the changes and implications. *World J Radiol,* 4**,** 128-34.

MIURA, Y., MOROOKA, M., SAX, N., ROYCHOUDHURI, R., ITOH-NAKADAI, A., BRYDUN, A., FUNAYAMA, R., NAKAYAMA, K., SATOMI, S., MATSUMOTO, M., IGARASHI, K. & MUTO, A. 2018. Bach2 Promotes B Cell Receptor-Induced Proliferation of B Lymphocytes and Represses Cyclin-Dependent Kinase Inhibitors. *J Immunol,* 200**,** 2882-2893.

MOENS, E., BROUWER, M., DIMOVA, T., GOLDMAN, M., WILLEMS, F. & VERMIJLEN, D. 2011. IL-23R and TCR signaling drives the generation of neonatal Vgamma9Vdelta2 T cells

expressing high levels of cytotoxic mediators and producing IFN-gamma and IL-17. *J Leukoc Biol,* 89**,** 743-52.

MOIR, S., HO, J., MALASPINA, A., WANG, W., DIPOTO, A. C., O'SHEA, M. A., ROBY, G., KOTTILIL, S., ARTHOS, J., PROSCHAN, M. A., CHUN, T. W. & FAUCI, A. S. 2008. Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals. *J Exp Med,* 205**,** 1797-805.

MOLDVAY, J., SCHEID, P., WILD, P., NABIL, K., SIAT, J., BORRELLY, J., MARIE, B., FARRE, G., LABIB, T., POTTIER, G., SESBOUE, R., BRONNER, C., VIGNAUD, J. M., MARTINET, Y. & MARTINET, N. 2000. Predictive survival markers in patients with surgically resected non-small cell lung carcinoma. *Clin Cancer Res,* 6**,** 1125-34.

MONTGOMERY, R. B., MAKARY, E., SCHIFFMAN, K., GOODELL, V. & DISIS, M. L. 2005. Endogenous anti-HER2 antibodies block HER2 phosphorylation and signaling through extracellular signal-regulated kinase. *Cancer Res,* 65**,** 650-6.

MOODY, C. A. & LAIMINS, L. A. 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer,* 10**,** 550-60.

MORENO-GALINDO, C., HERMSEN, M., GARCIA-PEDRERO, J. M., FRESNO, M. F., SUAREZ, C. & RODRIGO, J. P. 2014. p27 and BCL2 expression predicts response to chemotherapy in head and neck squamous cell carcinomas. *Oral Oncol,* 50**,** 128-34.

MOVAHEDI, K., LAOUI, D., GYSEMANS, C., BAETEN, M., STANGE, G., VAN DEN BOSSCHE, J., MACK, M., PIPELEERS, D., IN'T VELD, P., DE BAETSELIER, P. & VAN GINDERACHTER, J. A. 2010. Different tumor microenvironments contain functionally distinct subsets of macrophages derived from Ly6C(high) monocytes. *Cancer Res,* 70**,** 5728-39.

MUTO, A., OCHIAI, K., KIMURA, Y., ITOH-NAKADAI, A., CALAME, K. L., IKEBE, D., TASHIRO, S. & IGARASHI, K. 2010. Bach2 represses plasma cell gene regulatory network in B cells to promote antibody class switch. *EMBO J,* 29**,** 4048-61.

NAM, K. O., KANG, H., SHIN, S. M., CHO, K. H., KWON, B., KWON, B. S., KIM, S. J. & LEE, H. W. 2005. Cross-linking of 4-1BB activates TCR-signaling pathways in CD8+ T lymphocytes. *J Immunol,* 174**,** 1898-905.

NASMAN, A., ROMANITAN, M., NORDFORS, C., GRUN, N., JOHANSSON, H., HAMMARSTEDT, L., MARKLUND, L., MUNCK-WIKLAND, E., DALIANIS, T. & RAMQVIST, T. 2012. Tumor infiltrating CD8+ and Foxp3+ lymphocytes correlate to clinical outcome and human papillomavirus (HPV) status in tonsillar cancer. *PLoS One,* 7**,** e38711.

NEEFJES, J., JONGSMA, M. L., PAUL, P. & BAKKE, O. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nat Rev Immunol,* 11**,** 823-36.

NELSON, B. H. 2010. CD20+ B cells: the other tumor-infiltrating lymphocytes. *J Immunol,* 185**,** 4977-82.

NIKOLAYEVA, O. & ROBINSON, M. D. 2014. edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. *Methods Mol Biol,* 1150**,** 45-79.

NIMMO, R. A., MAY, G. E. & ENVER, T. 2015. Primed and ready: understanding lineage commitment through single cell analysis. *Trends Cell Biol,* 25**,** 459-67.

NORDFORS, C., GRUN, N., TERTIPIS, N., AHRLUND-RICHTER, A., HAEGGBLOM, L., SIVARS, L., DU, J., NYBERG, T., MARKLUND, L., MUNCK-WIKLAND, E., NASMAN, A., RAMQVIST, T. & DALIANIS, T. 2013. CD8+ and CD4+ tumour infiltrating lymphocytes in relation to human

papillomavirus status and clinical outcome in tonsillar and base of tongue squamous cell carcinoma. *Eur J Cancer,* 49**,** 2522-30.

NOVELLINO, L., CASTELLI, C. & PARMIANI, G. 2005. A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer Immunol Immunother,* 54**,** 187-207.

NUTT, S. L., HODGKIN, P. D., TARLINTON, D. M. & CORCORAN, L. M. 2015. The generation of antibody-secreting plasma cells. *Nat Rev Immunol,* 15**,** 160-71.

O'GARRA, A. & ARAI, N. 2000. The molecular basis of T helper 1 and T helper 2 cell differentiation. *Trends Cell Biol,* 10**,** 542-50.

OBEID, J. M., WAGES, N. A., HU, Y., DEACON, D. H. & SLINGLUFF, C. L., JR. 2017. Heterogeneity of CD8(+) tumor-infiltrating lymphocytes in non-small-cell lung cancer: impact on patient prognostic assessments and comparison of quantification by different sampling strategies. *Cancer Immunol Immunother,* 66**,** 33-43.

OLKHANUD, P. B., DAMDINSUREN, B., BODOGAI, M., GRESS, R. E., SEN, R., WEJKSZA, K., MALCHINKHUU, E., WERSTO, R. P. & BIRAGYN, A. 2011. Tumor-evoked regulatory B cells promote breast cancer metastasis by converting resting CD4(+) T cells to T-regulatory cells. *Cancer Res,* 71**,** 3505-15.

OLTHOF, N. C., SPEEL, E. J., KOLLIGS, J., HAESEVOETS, A., HENFLING, M., RAMAEKERS, F. C., PREUSS, S. F., DREBBER, U., WIELAND, U., SILLING, S., LAM, W. L., VUCIC, E. A., KREMER, B., KLUSSMANN, J. P. & HUEBBERS, C. U. 2014. Comprehensive analysis of HPV16 integration in OSCC reveals no significant impact of physical status on viral oncogene and virally disrupted human gene expression. *PLoS One,* 9**,** e88718.

OTT, P. A., HU, Z., KESKIN, D. B., SHUKLA, S. A., SUN, J., BOZYM, D. J., ZHANG, W., LUOMA, A., GIOBBIE-HURDER, A., PETER, L., CHEN, C., OLIVE, O., CARTER, T. A., LI, S., LIEB, D. J., EISENHAURE, T., GJINI, E., STEVENS, J., LANE, W. J., JAVERI, I., NELLAIAPPAN, K., SALAZAR, A. M., DALEY, H., SEAMAN, M., BUCHBINDER, E. I., YOON, C. H., HARDEN, M., LENNON, N., GABRIEL, S., RODIG, S. J., BAROUCH, D. H., ASTER, J. C., GETZ, G., WUCHERPFENNIG, K., NEUBERG, D., RITZ, J., LANDER, E. S., FRITSCH, E. F., HACOHEN, N. & WU, C. J. 2017. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature,* 547**,** 217-221.

OTTENSMEIER, C. H., PERRY, K. L., HARDEN, E. L., STASAKOVA, J., JENEI, V., FLEMING, J., WOOD, O., WOO, J., WOELK, C. H., THOMAS, G. J. & THIRDBOROUGH, S. M. 2016a. Upregulated Glucose Metabolism Correlates Inversely with CD8+ T-cell Infiltration and Survival in Squamous Cell Carcinoma. *Cancer Res,* 76**,** 4136-48.

OTTENSMEIER, C. H., PERRY, K. L., HARDEN, E. L., STASAKOVA, J., JENEI, V., FLEMING, J., WOOD, O., WOO, J., WOELK, C. H., THOMAS, G. J. & THIRDBOROUGH, S. M. 2016b. Upregulated Glucose Metabolism Correlates Inversely with CD8+ T-cell Infiltration and Survival in Squamous Cell Carcinoma. *Cancer Res*.

PAI, S. I. & WESTRA, W. H. 2009. Molecular pathology of head and neck cancer: implications for diagnosis, prognosis, and treatment. *Annu Rev Pathol,* 4**,** 49-70.

PAINTER, M. W., DAVIS, S., HARDY, R. R., MATHIS, D., BENOIST, C. & IMMUNOLOGICAL GENOME PROJECT, C. 2011. Transcriptomes of the B and T lineages compared by multiplatform microarray profiling. *J Immunol,* 186**,** 3047-57.

PALMER, C., DIEHN, M., ALIZADEH, A. A. & BROWN, P. O. 2006. Cell-type specific gene expression profiles of leukocytes in human peripheral blood. *BMC Genomics,* 7**,** 115.

PAPE, K. A., TAYLOR, J. J., MAUL, R. W., GEARHART, P. J. & JENKINS, M. K. 2011. Different B cell populations mediate early and late memory during an endogenous immune response. *Science,* 331**,** 1203-7.

PARDOLL, D. M. 2012. The blockade of immune checkpoints in cancer immunotherapy. *Nat Rev Cancer,* 12**,** 252-64.

PARK, S., JIANG, Z., MORTENSON, E. D., DENG, L., RADKEVICH-BROWN, O., YANG, X., SATTAR, H., WANG, Y., BROWN, N. K., GREENE, M., LIU, Y., TANG, J., WANG, S. & FU, Y. X. 2010. The therapeutic effect of anti-HER2/neu antibody depends on both innate and adaptive immunity. *Cancer Cell,* 18**,** 160-70.

PARKIN, D. M., BRAY, F., FERLAY, J. & PISANI, P. 2005. Global cancer statistics, 2002. *CA Cancer J Clin,* 55**,** 74-108.

PATEL, D., HUANG, S. M., BAGLIA, L. A. & MCCANCE, D. J. 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *EMBO J,* 18**,** 5061-72.

PATZELT, T., KEPPLER, S. J., GORKA, O., THOENE, S., WARTEWIG, T., RETH, M., FORSTER, I., LANG, R., BUCHNER, M. & RULAND, J. 2018. Foxp1 controls mature B cell survival and the development of follicular and B-1 B cells. *Proc Natl Acad Sci U S A,* 115**,** 3120-3125.

PAZ, H., JOO, E. J., CHOU, C. H., FEI, F., MAYO, K. H., ABDEL-AZIM, H., GHAZARIAN, H., GROFFEN, J. & HEISTERKAMP, N. 2018. Treatment of B-cell precursor acute lymphoblastic leukemia with the Galectin-1 inhibitor PTX008. *J Exp Clin Cancer Res,* 37**,** 67.

PEARCE, D. A., ARTHUR, L. M., TURNBULL, A. K., RENSHAW, L., SABINE, V. S., THOMAS, J. S., BARTLETT, J. M., DIXON, J. M. & SIMS, A. H. 2016. Tumour sampling method can significantly influence gene expression profiles derived from neoadjuvant window studies. *Sci Rep,* 6**,** 29434.

PEGHINI, B. C., ABDALLA, D. R., BARCELOS, A. C., TEODORO, L., MURTA, E. F. & MICHELIN, M. A. 2012a. Local cytokine profiles of patients with cervical intraepithelial and invasive neoplasia. *Hum Immunol,* 73**,** 920-6.

PEGHINI, B. C., ABDALLA, D. R., BARCELOS, A. C. M., TEODORO, L. D. G. V. L., MURTA, E. F. C. & MICHELIN, M. A. 2012b. Local cytokine profiles of patients with cervical intraepithelial and invasive neoplasia. *Hum Immunol,* 73**,** 920-6.

PENG, J., SUN, B. F., CHEN, C. Y., ZHOU, J. Y., CHEN, Y. S., CHEN, H., LIU, L., HUANG, D., JIANG, J., CUI, G. S., YANG, Y., WANG, W., GUO, D., DAI, M., GUO, J., ZHANG, T., LIAO, Q., LIU, Y., ZHAO, Y. L., HAN, D. L., ZHAO, Y., YANG, Y. G. & WU, W. 2019. Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Res*.

PENTCHEVA-HOANG, T., CHEN, L., PARDOLL, D. M. & ALLISON, J. P. 2007. Programmed death-1 concentration at the immunological synapse is determined by ligand affinity and availability. *Proc Natl Acad Sci U S A,* 104**,** 17765-70.

PEREIRA, J. P., KELLY, L. M. & CYSTER, J. G. 2010. Finding the right niche: B-cell migration in the early phases of T-dependent antibody responses. *Int Immunol,* 22**,** 413-9.

PERILLO, N. L., PACE, K. E., SEILHAMER, J. J. & BAUM, L. G. 1995. Apoptosis of T cells mediated by galectin-1. *Nature,* 378**,** 736-9.

PETRELLI, A., MIJNHEER, G., HOYTEMA VAN KONIJNENBURG, D. P., VAN DER WAL, M. M., GIOVANNONE, B., MOCHOLI, E., VAZIRPANAH, N., BROEN, J. C., HIJNEN, D., OLDENBURG, B., COFFER, P. J., VASTERT, S. J., PRAKKEN, B. J., SPIERINGS, E., PANDIT, A., MOKRY, M. & VAN WIJK, F. 2018. PD-1+CD8+ T cells are clonally expanding effectors in human chronic inflammation. *J Clin Invest,* 128**,** 4669-4681.

PHILIP, M., FAIRCHILD, L., SUN, L., HORSTE, E. L., CAMARA, S., SHAKIBA, M., SCOTT, A. C., VIALE, A., LAUER, P., MERGHOUB, T., HELLMANN, M. D., WOLCHOK, J. D., LESLIE, C. S. & SCHIETINGER, A. 2017. Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature,* 545**,** 452-456.

PIPKIN, M. E., SACKS, J. A., CRUZ-GUILLOTY, F., LICHTENHELD, M. G., BEVAN, M. J. & RAO, A. 2010. Interleukin-2 and inflammation induce distinct transcriptional programs that promote the differentiation of effector cytolytic T cells. *Immunity,* 32**,** 79-90.

PRIETO, I., SUJA, J. A., PEZZI, N., KREMER, L., MARTINEZ, A. C., RUFAS, J. S. & BARBERO, J. L. 2001. Mammalian STAG3 is a cohesin specific to sister chromatid arms in meiosis I. *Nat Cell Biol,* 3**,** 761-6.

PURAM, S. V., TIROSH, I., PARIKH, A. S., PATEL, A. P., YIZHAK, K., GILLESPIE, S., RODMAN, C., LUO, C. L., MROZ, E. A., EMERICK, K. S., DESCHLER, D. G., VARVARES, M. A., MYLVAGANAM, R., ROZENBLATT-ROSEN, O., ROCCO, J. W., FAQUIN, W. C., LIN, D. T., REGEV, A. & BERNSTEIN, B. E. 2017. Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell,* 171**,** 1611-1624 e24.

PYEON, D., NEWTON, M. A., LAMBERT, P. F., DEN BOON, J. A., SENGUPTA, S., MARSIT, C. J., WOODWORTH, C. D., CONNOR, J. P., HAUGEN, T. H., SMITH, E. M., KELSEY, K. T., TUREK, L. P. & AHLQUIST, P. 2007. Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. *Cancer Res,* 67**,** 4605-19.

QIAN, B. Z. & POLLARD, J. W. 2010. Macrophage diversity enhances tumor progression and metastasis. *Cell,* 141**,** 39-51.

QUACH, T. D., MANJARREZ-ORDUNO, N., ADLOWITZ, D. G., SILVER, L., YANG, H., WEI, C., MILNER, E. C. & SANZ, I. 2011. Anergic responses characterize a large fraction of human autoreactive naive B cells expressing low levels of surface IgM. *J Immunol,* 186**,** 4640-8.

QUATROMONI, J. G. & ERUSLANOV, E. 2012. Tumor-associated macrophages: function, phenotype, and link to prognosis in human lung cancer. *Am J Transl Res,* 4**,** 376-89.

RAO, R. R., LI, Q., ODUNSI, K. & SHRIKANT, P. A. 2010. The mTOR kinase determines effector versus memory CD8+ T cell fate by regulating the expression of transcription factors T-bet and Eomesodermin. *Immunity,* 32**,** 67-78.

REDMOND, W. L., MARINCEK, B. C. & SHERMAN, L. A. 2005. Distinct requirements for deletion versus anergy during CD8 T cell peripheral tolerance in vivo. *J Immunol,* 174**,** 2046-53.

REHWINKEL, J., TAN, C. P., GOUBAU, D., SCHULZ, O., PICHLMAIR, A., BIER, K., ROBB, N., VREEDE, F., BARCLAY, W., FODOR, E. & REIS E SOUSA, C. 2010. RIG-I detects viral genomic RNA during negative-strand RNA virus infection. *Cell,* 140**,** 397-408.

RESTIFO, N. P., DUDLEY, M. E. & ROSENBERG, S. A. 2012. Adoptive immunotherapy for cancer: harnessing the T cell response. *Nat Rev Immunol,* 12**,** 269-81.

RIZVI, N. A., HELLMANN, M. D., SNYDER, A., KVISTBORG, P., MAKAROV, V., HAVEL, J. J., LEE, W., YUAN, J., WONG, P., HO, T. S., MILLER, M. L., REKHTMAN, N., MOREIRA, A. L., IBRAHIM, F.,

BRUGGEMAN, C., GASMI, B., ZAPPASODI, R., MAEDA, Y., SANDER, C., GARON, E. B., MERGHOUB, T., WOLCHOK, J. D., SCHUMACHER, T. N. & CHAN, T. A. 2015. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science,* 348**,** 124-8.

ROBERT, C., SCHACHTER, J., LONG, G. V., ARANCE, A., GROB, J. J., MORTIER, L., DAUD, A., CARLINO, M. S., MCNEIL, C., LOTEM, M., LARKIN, J., LORIGAN, P., NEYNS, B., BLANK, C. U., HAMID, O., MATEUS, C., SHAPIRA-FROMMER, R., KOSH, M., ZHOU, H., IBRAHIM, N., EBBINGHAUS, S., RIBAS, A. & INVESTIGATORS, K.-. 2015. Pembrolizumab versus Ipilimumab in Advanced Melanoma. *N Engl J Med,* 372**,** 2521-32.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

ROBINSON, M. D. & OSHLACK, A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol,* 11**,** R25.

RODRIGUEZ-PINTO, D. 2005. B cells as antigen presenting cells. *Cell Immunol,* 238**,** 67-75.

ROONEY, M. S., SHUKLA, S. A., WU, C. J., GETZ, G. & HACOHEN, N. 2015. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell,* 160**,** 48-61.

ROSALES, S. L., LIANG, S., ENGEL, I., SCHMIEDEL, B. J., KRONENBERG, M., VIJAYANAND, P. & SEUMOIS, G. 2018. A Sensitive and Integrated Approach to Profile Messenger RNA from Samples with Low Cell Numbers. *Methods Mol Biol,* 1799**,** 275-302.

ROSENBERG, H. F., PHIPPS, S. & FOSTER, P. S. 2007. Eosinophil trafficking in allergy and asthma. *J Allergy Clin Immunol,* 119**,** 1303-10; quiz 1311-2.

ROSENBERG, S. A. & RESTIFO, N. P. 2015. Adoptive cell transfer as personalized immunotherapy for human cancer. *Science,* 348**,** 62-8.

ROSSER, E. C. & MAURI, C. 2015. Regulatory B cells: origin, phenotype, and function. *Immunity,* 42**,** 607-12.

ROSZER, T. 2015. Understanding the Mysterious M2 Macrophage through Activation Markers and Effector Mechanisms. *Mediators Inflamm,* 2015**,** 816460.

RUDD, C. E., TAYLOR, A. & SCHNEIDER, H. 2009. CD28 and CTLA-4 coreceptor expression and signal transduction. *Immunol Rev,* 229**,** 12-26.

RUSSELL, S., ANGELL, T., LECHNER, M., LIEBERTZ, D., CORREA, A., SINHA, U., KOKOT, N. & EPSTEIN, A. 2013. Immune cell infiltration patterns and survival in head and neck squamous cell carcinoma. *Head Neck Oncol,* 5**,** 24.

RYGIEL, T. P. & MEYAARD, L. 2012. CD200R signaling in tumor tolerance and inflammation: A tricky balance. *Curr Opin Immunol,* 24**,** 233-8.

SADE-FELDMAN, M., YIZHAK, K., BJORGAARD, S. L., RAY, J. P., DE BOER, C. G., JENKINS, R. W., LIEB, D. J., CHEN, J. H., FREDERICK, D. T., BARZILY-ROKNI, M., FREEMAN, S. S., REUBEN, A., HOOVER, P. J., VILLANI, A. C., IVANOVA, E., PORTELL, A., LIZOTTE, P. H., AREF, A. R., ELIANE, J. P., HAMMOND, M. R., VITZTHUM, H., BLACKMON, S. M., LI, B., GOPALAKRISHNAN, V., REDDY, S. M., COOPER, Z. A., PAWELETZ, C. P., BARBIE, D. A., STEMMER-RACHAMIMOV, A., FLAHERTY, K. T., WARGO, J. A., BOLAND, G. M., SULLIVAN, R. J., GETZ, G. & HACOHEN, N. 2018. Defining T Cell States Associated with Response to Checkpoint Immunotherapy in Melanoma. *Cell,* 175**,** 998-1013 e20.

SAHIN, U., DERHOVANESSIAN, E., MILLER, M., KLOKE, B. P., SIMON, P., LOWER, M., BUKUR, V., TADMOR, A. D., LUXEMBURGER, U., SCHRORS, B., OMOKOKO, T., VORMEHR, M., ALBRECHT, C., PARUZYNSKI, A., KUHN, A. N., BUCK, J., HEESCH, S., SCHREEB, K. H., MULLER, F., ORTSEIFER, I., VOGLER, I., GODEHARDT, E., ATTIG, S., RAE, R., BREITKREUZ, A., TOLLIVER, C., SUCHAN, M., MARTIC, G., HOHBERGER, A., SORN, P., DIEKMANN, J., CIESLA, J., WAKSMANN, O., BRUCK, A. K., WITT, M., ZILLGEN, M., ROTHERMEL, A., KASEMANN, B., LANGER, D., BOLTE, S., DIKEN, M., KREITER, S., NEMECEK, R., GEBHARDT, C., GRABBE, S., HOLLER, C., UTIKAL, J., HUBER, C., LOQUAI, C. & TURECI, O. 2017. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature,* 547**,** 222-226.

SAKIMURA, C., TANAKA, H., OKUNO, T., HIRAMATSU, S., MUGURUMA, K., HIRAKAWA, K., WANIBUCHI, H. & OHIRA, M. 2017. B cells in tertiary lymphoid structures are associated with favorable prognosis in gastric cancer. *J Surg Res,* 215**,** 74-82.

SAKUISHI, K., APETOH, L., SULLIVAN, J. M., BLAZAR, B. R., KUCHROO, V. K. & ANDERSON, A. C. 2010. Targeting Tim-3 and PD-1 pathways to reverse T cell exhaustion and restore anti-tumor immunity. *J Exp Med,* 207**,** 2187-94.

SALAVOURA, K., KOLIALEXI, A., TSANGARIS, G. & MAVROU, A. 2008. Development of cancer in patients with primary immunodeficiencies. *Anticancer Res,* 28**,** 1263-9.

SALEK, M., MCGOWAN, S., TRUDGIAN, D. C., DUSHEK, O., DE WET, B., EFSTATHIOU, G. & ACUTO, O. 2013. Quantitative phosphoproteome analysis unveils LAT as a modulator of CD3zeta and ZAP-70 tyrosine phosphorylation. *PLoS One,* 8**,** e77423.

SALLUSTO, F., GEGINAT, J. & LANZAVECCHIA, A. 2004. Central memory and effector memory T cell subsets: function, generation, and maintenance. *Annu Rev Immunol,* 22**,** 745-63.

SANCHEZ-PAULETE, A. R., TEIJEIRA, A., CUETO, F. J., GARASA, S., PEREZ-GRACIA, J. L., SANCHEZ-ARRAEZ, A., SANCHO, D. & MELERO, I. 2017. Antigen cross-presentation and T-cell cross-priming in cancer immunology and immunotherapy. *Ann Oncol,* 28**,** xii44-xii55.

SARVARIA, A., MADRIGAL, J. A. & SAUDEMONT, A. 2017. B cell regulation in cancer and anti-tumor immunity. *Cell Mol Immunol,* 14**,** 662-674.

SATHALIYAWALA, T., KUBOTA, M., YUDANIN, N., TURNER, D., CAMP, P., THOME, J. J., BICKHAM, K. L., LERNER, H., GOLDSTEIN, M., SYKES, M., KATO, T. & FARBER, D. L. 2013. Distribution and compartmentalization of human circulating and tissue-resident memory T cell subsets. *Immunity,* 38**,** 187-97.

SCHACHTER, J., RIBAS, A., LONG, G. V., ARANCE, A., GROB, J. J., MORTIER, L., DAUD, A., CARLINO, M. S., MCNEIL, C., LOTEM, M., LARKIN, J., LORIGAN, P., NEYNS, B., BLANK, C., PETRELLA, T. M., HAMID, O., ZHOU, H., EBBINGHAUS, S., IBRAHIM, N. & ROBERT, C. 2017. Pembrolizumab versus ipilimumab for advanced melanoma: final overall survival results of a multicentre, randomised, open-label phase 3 study (KEYNOTE-006). *Lancet,* 390**,** 1853-1862.

SCHENKEL, J. M., FRASER, K. A., BEURA, L. K., PAUKEN, K. E., VEZYS, V. & MASOPUST, D. 2014. T cell memory. Resident memory CD8 T cells trigger protective innate and adaptive immune responses. *Science,* 346**,** 98-101.

SCHENKEL, J. M. & MASOPUST, D. 2014. Tissue-resident memory T cells. *Immunity,* 41**,** 886-97.

SCHIER, W. W. 1954. Cutaneous anergy and Hodgkin's disease. *N Engl J Med,* 250**,** 353-61.

SCHILDBERG, F. A., KLEIN, S. R., FREEMAN, G. J. & SHARPE, A. H. 2016. Coinhibitory Pathways in the B7-CD28 Ligand-Receptor Family. *Immunity,* 44**,** 955-72.

SCHNEIDER, W. M., CHEVILLOTTE, M. D. & RICE, C. M. 2014. Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol,* 32**,** 513-45.

SCHRAML, B. U. & REIS E SOUSA, C. 2015. Defining dendritic cells. *Curr Opin Immunol,* 32**,** 13-20.

SCHREIBELT, G., BOL, K. F., WESTDORP, H., WIMMERS, F., AARNTZEN, E. H., DUIVEMAN-DE BOER, T., VAN DE RAKT, M. W., SCHARENBORG, N. M., DE BOER, A. J., POTS, J. M., OLDE NORDKAMP, M. A., VAN OORSCHOT, T. G., TEL, J., WINKELS, G., PETRY, K., BLOKX, W. A., VAN ROSSUM, M. M., WELZEN, M. E., MUS, R. D., CROOCKEWIT, S. A., KOORNSTRA, R. H., JACOBS, J. F., KELDERMAN, S., BLANK, C. U., GERRITSEN, W. R., PUNT, C. J., FIGDOR, C. G. & DE VRIES, I. J. 2016. Effective Clinical Responses in Metastatic Melanoma Patients after Vaccination with Primary Myeloid Dendritic Cells. *Clin Cancer Res,* 22**,** 2155-66.

SCHROEDER, A., MUELLER, O., STOCKER, S., SALOWSKY, R., LEIBER, M., GASSMANN, M., LIGHTFOOT, S., MENZEL, W., GRANZOW, M. & RAGG, T. 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol,* 7**,** 3.

SCHROEDER, H. W., JR. & CAVACINI, L. 2010. Structure and function of immunoglobulins. *J Allergy Clin Immunol,* 125**,** S41-52.

SCHWARTZENTRUBER, D. J., LAWSON, D. H., RICHARDS, J. M., CONRY, R. M., MILLER, D. M., TREISMAN, J., GAILANI, F., RILEY, L., CONLON, K., POCKAJ, B., KENDRA, K. L., WHITE, R. L., GONZALEZ, R., KUZEL, T. M., CURTI, B., LEMING, P. D., WHITMAN, E. D., BALKISSOON, J., REINTGEN, D. S., KAUFMAN, H., MARINCOLA, F. M., MERINO, M. J., ROSENBERG, S. A., CHOYKE, P., VENA, D. & HWU, P. 2011. gp100 peptide vaccine and interleukin-2 in patients with advanced melanoma. *N Engl J Med,* 364**,** 2119-27.

SCIARANGHELLA, G., TONG, N., MAHAN, A. E., SUSCOVICH, T. J. & ALTER, G. 2013. Decoupling activation and exhaustion of B cells in spontaneous controllers of HIV infection. *AIDS,* 27**,** 175-80.

SEIWERT, T. Y., BURTNESS, B., MEHRA, R., WEISS, J., BERGER, R., EDER, J. P., HEATH, K., MCCLANAHAN, T., LUNCEFORD, J., GAUSE, C., CHENG, J. D. & CHOW, L. Q. 2016. Safety and clinical activity of pembrolizumab for treatment of recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-012): an open-label, multicentre, phase 1b trial. *Lancet Oncol,* 17**,** 956-65.

SELIKTAR-OFIR, S., MERHAVI-SHOHAM, E., ITZHAKI, O., YUNGER, S., MARKEL, G., SCHACHTER, J. & BESSER, M. J. 2017. Selection of Shared and Neoantigen-Reactive T Cells for Adoptive Cell Therapy Based on CD137 Separation. *Front Immunol,* 8**,** 1211.

SEUMOIS, G., CHAVEZ, L., GERASIMOVA, A., LIENHARD, M., OMRAN, N., KALINKE, L., VEDANAYAGAM, M., GANESAN, A. P., CHAWLA, A., DJUKANOVIC, R., ANSEL, K. M., PETERS, B., RAO, A. & VIJAYANAND, P. 2014. Epigenomic analysis of primary human T cells reveals enhancers associated with TH2 memory cell differentiation and asthma susceptibility. *Nat Immunol,* 15**,** 777-88.

SEUMOIS, G., VIJAYANAND, P., EISLEY, C. J., OMRAN, N., KALINKE, L., NORTH, M., GANESAN, A. P., SIMPSON, L. J., HUNKAPILLER, N., MOLTZAHN, F., WOODRUFF, P. G., FAHY, J. V., ERLE, D. J., DJUKANOVIC, R., BLELLOCH, R. & ANSEL, K. M. 2012. An integrated nano-scale approach to profile miRNAs in limited clinical samples. *Am J Clin Exp Immunol,* 1**,** 70-89.

SHALAPOUR, S., FONT-BURGADA, J., DI CARO, G., ZHONG, Z., SANCHEZ-LOPEZ, E., DHAR, D., WILLIMSKY, G., AMMIRANTE, M., STRASNER, A., HANSEL, D. E., JAMIESON, C., KANE, C. J.,

KLATTE, T., BIRNER, P., KENNER, L. & KARIN, M. 2015. Immunosuppressive plasma cells impede T-cell-dependent immunogenic chemotherapy. *Nature,* 521**,** 94-8.

SHALAPOUR, S., LIN, X. J., BASTIAN, I. N., BRAIN, J., BURT, A. D., AKSENOV, A. A., VRBANAC, A. F., LI, W., PERKINS, A., MATSUTANI, T., ZHONG, Z., DHAR, D., NAVAS-MOLINA, J. A., XU, J., LOOMBA, R., DOWNES, M., YU, R. T., EVANS, R. M., DORRESTEIN, P. C., KNIGHT, R., BENNER, C., ANSTEE, Q. M. & KARIN, M. 2017. Inflammation-induced IgA+ cells dismantle anti-liver cancer immunity. *Nature,* 551**,** 340-345.

SHANKARAN, V., IKEDA, H., BRUCE, A. T., WHITE, J. M., SWANSON, P. E., OLD, L. J. & SCHREIBER, R. D. 2001. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity. *Nature,* 410**,** 1107-11.

SHARMA, P. & ALLISON, J. P. 2015. Immune checkpoint targeting in cancer therapy: toward combination strategies with curative potential. *Cell,* 161**,** 205-14.

SHEN, P. & FILLATREAU, S. 2015. Antibody-independent functions of B cells: a focus on cytokines. *Nat Rev Immunol,* 15**,** 441-51.

SHETTY, K. & SCHATZ, D. G. 2015. Recruitment of RAG1 and RAG2 to Chromatinized DNA during V(D)J Recombination. *Mol Cell Biol,* 35**,** 3701-13.

SHI, C. & PAMER, E. G. 2011. Monocyte recruitment during infection and inflammation. *Nat Rev Immunol,* 11**,** 762-74.

SHI, Y., FAN, X., DENG, H., BREZSKI, R. J., RYCYZYN, M., JORDAN, R. E., STROHL, W. R., ZOU, Q., ZHANG, N. & AN, Z. 2015. Trastuzumab triggers phagocytic killing of high HER2 cancer cells in vitro and in vivo by interaction with Fcgamma receptors on macrophages. *J Immunol,* 194**,** 4379-86.

SHLOMCHIK, M. J. & WEISEL, F. 2012. Germinal center selection and the development of memory B and plasma cells. *Immunol Rev,* 247**,** 52-63.

SHOEMAKER, J. E., LOPES, T. J., GHOSH, S., MATSUOKA, Y., KAWAOKA, Y. & KITANO, H. 2012. CTen: a web-based platform for identifying enriched cell types from heterogeneous microarray data. *BMC Genomics,* 13**,** 460.

SIEGEL, R. L., MILLER, K. D. & JEMAL, A. 2018. Cancer statistics, 2018. *CA Cancer J Clin,* 68**,** 7-30.

SIGEL, K., WISNIVESKY, J., GORDON, K., DUBROW, R., JUSTICE, A., BROWN, S. T., GOULET, J., BUTT, A. A., CRYSTAL, S., RIMLAND, D., RODRIGUEZ-BARRADAS, M., GIBERT, C., PARK, L. S. & CROTHERS, K. 2012. HIV as an independent risk factor for incident lung cancer. *AIDS,* 26**,** 1017-25.

SLEBOS, R. J., YI, Y., ELY, K., CARTER, J., EVJEN, A., ZHANG, X., SHYR, Y., MURPHY, B. M., CMELAK, A. J., BURKEY, B. B., NETTERVILLE, J. L., LEVY, S., YARBROUGH, W. G. & CHUNG, C. H. 2006. Gene expression differences associated with human papillomavirus status in head and neck squamous cell carcinoma. *Clin Cancer Res,* 12**,** 701-9.

SMULSKI, C. R., KURY, P., SEIDEL, L. M., STAIGER, H. S., EDINGER, A. K., WILLEN, L., SEIDL, M., HESS, H., SALZER, U., ROLINK, A. G., RIZZI, M., SCHNEIDER, P. & EIBEL, H. 2017. BAFF- and TACI-Dependent Processing of BAFFR by ADAM Proteases Regulates the Survival of B Cells. *Cell Rep,* 18**,** 2189-2202.

SMYTH, M. J., CRETNEY, E., KELLY, J. M., WESTWOOD, J. A., STREET, S. E., YAGITA, H., TAKEDA, K., VAN DOMMELEN, S. L., DEGLI-ESPOSTI, M. A. & HAYAKAWA, Y. 2005. Activation of NK cell cytotoxicity. *Mol Immunol,* 42**,** 501-10.

STANKOVIC, B., BJORHOVDE, H. A. K., SKARSHAUG, R., AAMODT, H., FRAFJORD, A., MULLER, E., HAMMARSTROM, C., BERAKI, K., BAEKKEVOLD, E. S., WOLDBAEK, P. R., HELLAND, A., BRUSTUGUN, O. T., OYNEBRATEN, I. & CORTHAY, A. 2018. Immune Cell Composition in Human Non-small Cell Lung Cancer. *Front Immunol,* 9**,** 3101.

STANLEY, M. 2010. HPV - immune response to infection and vaccination. *Infect Agent Cancer,* 5**,** 19.

STANLEY, M. A. 2012. Epithelial cell responses to infection with human papillomavirus. *Clin Microbiol Rev,* 25**,** 215-22.

STEELE, J. C., MANN, C. H., ROOKES, S., ROLLASON, T., MURPHY, D., FREETH, M. G., GALLIMORE, P. H. & ROBERTS, S. 2005. T-cell responses to human papillomavirus type 16 among women with different grades of cervical neoplasia. *Br J Cancer,* 93**,** 248-59.

STONE, K. D., PRUSSIN, C. & METCALFE, D. D. 2010. IgE, mast cells, basophils, and eosinophils. *J Allergy Clin Immunol,* 125**,** S73-80.

STRANSKY, N., EGLOFF, A. M., TWARD, A. D., KOSTIC, A. D., CIBULSKIS, K., SIVACHENKO, A., KRYUKOV, G. V., LAWRENCE, M. S., SOUGNEZ, C., MCKENNA, A., SHEFLER, E., RAMOS, A. H., STOJANOV, P., CARTER, S. L., VOET, D., CORTES, M. L., AUCLAIR, D., BERGER, M. F., SAKSENA, G., GUIDUCCI, C., ONOFRIO, R. C., PARKIN, M., ROMKES, M., WEISSFELD, J. L., SEETHALA, R. R., WANG, L., RANGEL-ESCARENO, C., FERNANDEZ-LOPEZ, J. C., HIDALGO-MIRANDA, A., MELENDEZ-ZAJGLA, J., WINCKLER, W., ARDLIE, K., GABRIEL, S. B., MEYERSON, M., LANDER, E. S., GETZ, G., GOLUB, T. R., GARRAWAY, L. A. & GRANDIS, J. R. 2011. The mutational landscape of head and neck squamous cell carcinoma. *Science,* 333**,** 1157-60.

STRUTT, T. M., MCKINSTRY, K. K., MARSHALL, N. B., VONG, A. M., DUTTON, R. W. & SWAIN, S. L. 2013. Multipronged CD4(+) T-cell effector and memory responses cooperate to provide potent immunity against respiratory virus. *Immunol Rev,* 255**,** 149-64.

SU, G. H., IP, H. S., COBB, B. S., LU, M. M., CHEN, H. M. & SIMON, M. C. 1996. The Ets protein Spi-B is expressed exclusively in B cells and T cells during development. *J Exp Med,* 184**,** 203-14.

SULLIVAN, B. M. & LOCKSLEY, R. M. 2009. Basophils: a nonredundant contributor to host immunity. *Immunity,* 30**,** 12-20.

SUNDBLAD, V., MOROSI, L. G., GEFFNER, J. R. & RABINOVICH, G. A. 2017. Galectin-1: A Jack-of-All-Trades in the Resolution of Acute and Chronic Inflammation. *J Immunol,* 199**,** 3721-3730.

SUNDD, P., POSPIESZALSKA, M. K. & LEY, K. 2013. Neutrophil rolling at high shear: flattening, catch bond behavior, tethers and slings. *Mol Immunol,* 55**,** 59-69.

SUPEK, F., BOSNJAK, M., SKUNCA, N. & SMUC, T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One,* 6**,** e21800.

TAKAHASHI, Y., GLEBER-NETTO, F. O., BELL, D., ROBERTS, D., XIE, T. X., ABDELMEGUID, A. S., PICKERING, C., MYERS, J. N. & HANNA, E. Y. 2019. Identification of novel diagnostic markers for sinonasal undifferentiated carcinoma. *Head Neck,* 41**,** 2688-2695.

TAKEUCHI, O. & AKIRA, S. 2009. Innate immunity to virus infection. *Immunol Rev,* 227**,** 75-86.

TANG, X., MO, C., WANG, Y., WEI, D. & XIAO, H. 2013. Anti-tumour strategies aiming to target tumour-associated macrophages. *Immunology,* 138**,** 93-104.

TARTE, K., ZHAN, F., DE VOS, J., KLEIN, B. & SHAUGHNESSY, J., JR. 2003. Gene expression profiling of plasma cells and plasmablasts: toward a better understanding of the late stages of B-cell differentiation. *Blood,* 102**,** 592-600.

TERABE, M., MATSUI, S., NOBEN-TRAUTH, N., CHEN, H., WATSON, C., DONALDSON, D. D., CARBONE, D. P., PAUL, W. E. & BERZOFSKY, J. A. 2000. NKT cell-mediated repression of tumor immunosurveillance by IL-13 and the IL-4R-STAT6 pathway. *Nat Immunol,* 1**,** 515-20.

THOMMEN, D. S., KOELZER, V. H., HERZIG, P., ROLLER, A., TREFNY, M., DIMELOE, S., KIIALAINEN, A., HANHART, J., SCHILL, C., HESS, C., SAVIC PRINCE, S., WIESE, M., LARDINOIS, D., HO, P. C., KLEIN, C., KARANIKAS, V., MERTZ, K. D., SCHUMACHER, T. N. & ZIPPELIUS, A. 2018. A transcriptionally and functionally distinct PD-1(+) CD8(+) T cell pool with predictive potential in non-small-cell lung cancer treated with PD-1 blockade. *Nat Med,* 24**,** 994-1004.

THORSSON, V., GIBBS, D. L., BROWN, S. D., WOLF, D., BORTONE, D. S., OU YANG, T. H., PORTA-PARDO, E., GAO, G. F., PLAISIER, C. L., EDDY, J. A., ZIV, E., CULHANE, A. C., PAULL, E. O., SIVAKUMAR, I. K. A., GENTLES, A. J., MALHOTRA, R., FARSHIDFAR, F., COLAPRICO, A., PARKER, J. S., MOSE, L. E., VO, N. S., LIU, J., LIU, Y., RADER, J., DHANKANI, V., REYNOLDS, S. M., BOWLBY, R., CALIFANO, A., CHERNIACK, A. D., ANASTASSIOU, D., BEDOGNETTI, D., RAO, A., CHEN, K., KRASNITZ, A., HU, H., MALTA, T. M., NOUSHMEHR, H., PEDAMALLU, C. S., BULLMAN, S., OJESINA, A. I., LAMB, A., ZHOU, W., SHEN, H., CHOUEIRI, T. K., WEINSTEIN, J. N., GUINNEY, J., SALTZ, J., HOLT, R. A., RABKIN, C. E., CANCER GENOME ATLAS RESEARCH, N., LAZAR, A. J., SERODY, J. S., DEMICCO, E. G., DISIS, M. L., VINCENT, B. G. & SHMULEVICH, L. 2018. The Immune Landscape of Cancer. *Immunity,* 48**,** 812-830 e14.

THUNNISSEN, E., BOERS, E., HEIDEMAN, D. A., GRUNBERG, K., KUIK, D. J., NOORDUIN, A., VAN OOSTERHOUT, M., PRONK, D., SELDENRIJK, C., SIETSMA, H., SMIT, E. F., VAN SUYLEN, R., VON DER THUSEN, J., VRUGT, B., WIERSMA, A., WITTE, B. I. & DEN BAKKER, M. 2012. Correlation of immunohistochemical staining p63 and TTF-1 with EGFR and K-ras mutational spectrum and diagnostic reproducibility in non small cell lung carcinoma. *Virchows Arch,* 461**,** 629-38.

THURLOW, J. K., PENA MURILLO, C. L., HUNTER, K. D., BUFFA, F. M., PATIAR, S., BETTS, G., WEST, C. M., HARRIS, A. L., PARKINSON, E. K., HARRISON, P. R., OZANNE, B. W., PARTRIDGE, M. & KALNA, G. 2010. Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. *J Clin Oncol,* 28**,** 2881-8.

TIROSH, I., IZAR, B., PRAKADAN, S. M., WADSWORTH, M. H., 2ND, TREACY, D., TROMBETTA, J. J., ROTEM, A., RODMAN, C., LIAN, C., MURPHY, G., FALLAHI-SICHANI, M., DUTTON-REGESTER, K., LIN, J. R., COHEN, O., SHAH, P., LU, D., GENSHAFT, A. S., HUGHES, T. K., ZIEGLER, C. G., KAZER, S. W., GAILLARD, A., KOLB, K. E., VILLANI, A. C., JOHANNESSEN, C. M., ANDREEV, A. Y., VAN ALLEN, E. M., BERTAGNOLLI, M., SORGER, P. K., SULLIVAN, R. J., FLAHERTY, K. T., FREDERICK, D. T., JANE-VALBUENA, J., YOON, C. H., ROZENBLATT-ROSEN, O., SHALEK, A. K., REGEV, A. & GARRAWAY, L. A. 2016. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science,* 352**,** 189-96.

TRAPNELL, C., CACCHIARELLI, D., GRIMSBY, J., POKHAREL, P., LI, S., MORSE, M., LENNON, N. J., LIVAK, K. J., MIKKELSEN, T. S. & RINN, J. L. 2014. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol,* 32**,** 381-386.

TRAPNELL, C., PACHTER, L. & SALZBERG, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics,* 25**,** 1105-11.

TRINCHIERI, G. 2003. Interleukin-12 and the regulation of innate resistance and adaptive immunity. *Nat Rev Immunol,* 3**,** 133-46.

TRINCHIERI, G. & SHER, A. 2007. Cooperation of Toll-like receptor signals in innate immune defence. *Nat Rev Immunol,* 7**,** 179-90.

TRIPATHI, S. K. & LAHESMAA, R. 2014. Transcriptional and epigenetic regulation of T-helper lineage specification. *Immunol Rev,* 261**,** 62-83.

TSAI, C. M., WU, H. Y., SU, T. H., KUO, C. W., HUANG, H. W., CHUNG, C. H., CHEN, C. S., KHOO, K. H., CHEN, Y. J. & LIN, K. I. 2014. Phosphoproteomic analyses reveal that galectin-1 augments the dynamics of B-cell receptor signaling. *J Proteomics,* 103**,** 241-53.

TUMEH, P. C., HARVIEW, C. L., YEARLEY, J. H., SHINTAKU, I. P., TAYLOR, E. J., ROBERT, L., CHMIELOWSKI, B., SPASIC, M., HENRY, G., CIOBANU, V., WEST, A. N., CARMONA, M., KIVORK, C., SEJA, E., CHERRY, G., GUTIERREZ, A. J., GROGAN, T. R., MATEUS, C., TOMASIC, G., GLASPY, J. A., EMERSON, R. O., ROBINS, H., PIERCE, R. H., ELASHOFF, D. A., ROBERT, C. & RIBAS, A. 2014. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature,* 515**,** 568-71.

TURVEY, S. E. & BROIDE, D. H. 2010. Innate immunity. *J Allergy Clin Immunol,* 125**,** S24-32.

UHM, T. G., KIM, B. S. & CHUNG, I. Y. 2012. Eosinophil development, regulation of eosinophil-specific genes, and role of eosinophils in the pathogenesis of asthma. *Allergy Asthma Immunol Res,* 4**,** 68-79.

VAN LAETHEM, F., TIKHONOVA, A. N., POBEZINSKY, L. A., TAI, X., KIMURA, M. Y., LE SAOUT, C., GUINTER, T. I., ADAMS, A., SHARROW, S. O., BERNHARDT, G., FEIGENBAUM, L. & SINGER, A. 2013. Lck availability during thymic selection determines the recognition specificity of the T cell repertoire. *Cell,* 154**,** 1326-41.

VAN PANHUYS, N., KLAUSCHEN, F. & GERMAIN, R. N. 2014. T-cell-receptor-dependent signal intensity dominantly controls CD4(+) T cell polarization In Vivo. *Immunity,* 41**,** 63-74.

VARGA, Z., LEBEAU, A., BU, H., HARTMANN, A., PENAULT-LLORCA, F., GUERINI-ROCCO, E., SCHRAML, P., SYMMANS, F., STOEHR, R., TENG, X., TURZYNSKI, A., VON WASIELEWSKI, R., GURTLER, C., LAIBLE, M., SCHLOMBS, K., JOENSUU, H., KELLER, T., SINN, P., SAHIN, U., BARTLETT, J. & VIALE, G. 2017. An international reproducibility study validating quantitative determination of ERBB2, ESR1, PGR, and MKI67 mRNA in breast cancer using MammaTyper(R). *Breast Cancer Res,* 19**,** 55.

VAZQUEZ, M. I., CATALAN-DIBENE, J. & ZLOTNIK, A. 2015. B cells responses and cytokine production are regulated by their immune microenvironment. *Cytokine,* 74**,** 318-26.

VESELY, M. D., KERSHAW, M. H., SCHREIBER, R. D. & SMYTH, M. J. 2011. Natural innate and adaptive immunity to cancer. *Annu Rev Immunol,* 29**,** 235-71.

VETHANAYAGAM, R. R., ALMYROUDIS, N. G., GRIMM, M. J., LEWANDOWSKI, D. C., PHAM, C. T., BLACKWELL, T. S., PETRAITIENE, R., PETRAITIS, V., WALSH, T. J., URBAN, C. F. & SEGAL, B. H. 2011. Role of NADPH oxidase versus neutrophil proteases in antimicrobial host defense. *PLoS One,* 6**,** e28149.

VIDARSSON, G., DEKKERS, G. & RISPENS, T. 2014. IgG subclasses and allotypes: from structure to effector functions. *Front Immunol,* 5**,** 520.

WANG, H., JAIN, S., LI, P., LIN, J. X., OH, J., QI, C., GAO, Y., SUN, J., SAKAI, T., NAGHASHFAR, Z., ABBASI, S., KOVALCHUK, A. L., BOLLAND, S., NUTT, S. L., LEONARD, W. J. & MORSE, H. C., 3RD 2019. Transcription factors IRF8 and PU.1 are required for follicular B cell development and BCL6-driven germinal center responses. *Proc Natl Acad Sci U S A,* 116**,** 9511-9520.

WANG, L., LIU, J. Q., TALEBIAN, F., EL-OMRANI, H. Y., KHATTABI, M., YU, L. & BAI, X. F. 2010. Tumor expression of CD200 inhibits IL-10 production by tumor-associated myeloid cells and prevents tumor immune evasion of CTL therapy. *Eur J Immunol,* 40**,** 2569-79.

WANG, X., RODDA, L. B., BANNARD, O. & CYSTER, J. G. 2014. Integrin-mediated interactions between B cells and follicular dendritic cells influence germinal center B cell fitness. *J Immunol,* 192**,** 4601-9.

WANG, Z. Q., MILNE, K., DEROCHER, H., WEBB, J. R., NELSON, B. H. & WATSON, P. H. 2016. CD103 and Intratumoral Immune Response in Breast Cancer. *Clin Cancer Res,* 22**,** 6290-6297.

WARD, M. J., MELLOWS, T., HARRIS, S., WEBB, A., PATEL, N. N., COX, H. J., PIPER, K., OTTENSMEIER, C. H., THOMAS, G. J. & KING, E. V. 2014a. Staging and treatment of oropharyngeal cancer in the human papillomavirus era. *Head Neck*.

WARD, M. J., MELLOWS, T., HARRIS, S., WEBB, A., PATEL, N. N., COX, H. J., PIPER, K., OTTENSMEIER, C. H., THOMAS, G. J. & KING, E. V. 2015. Staging and treatment of oropharyngeal cancer in the human papillomavirus era. *Head Neck,* 37**,** 1002-13.

WARD, M. J., THIRDBOROUGH, S. M., MELLOWS, T., RILEY, C., HARRIS, S., SUCHAK, K., WEBB, A., HAMPTON, C., PATEL, N. N., RANDALL, C. J., COX, H. J., JOGAI, S., PRIMROSE, J., PIPER, K., OTTENSMEIER, C. H., KING, E. V. & THOMAS, G. J. 2014b. Tumour-infiltrating lymphocytes predict for outcome in HPV-positive oropharyngeal cancer. *Br J Cancer,* 110**,** 489-500.

WATKINS, N. A., GUSNANTO, A., DE BONO, B., DE, S., MIRANDA-SAAVEDRA, D., HARDIE, D. L., ANGENENT, W. G., ATTWOOD, A. P., ELLIS, P. D., ERBER, W., FOAD, N. S., GARNER, S. F., ISACKE, C. M., JOLLEY, J., KOCH, K., MACAULAY, I. C., MORLEY, S. L., RENDON, A., RICE, K. M., TAYLOR, N., THIJSSEN-TIMMER, D. C., TIJSSEN, M. R., VAN DER SCHOOT, C. E., WERNISCH, L., WINZER, T., DUDBRIDGE, F., BUCKLEY, C. D., LANGFORD, C. F., TEICHMANN, S., GOTTGENS, B., OUWEHAND, W. H. & BLOODOMICS, C. 2009. A HaemAtlas: characterizing gene expression in differentiated human blood cells. *Blood,* 113**,** e1-9.

WEBER, T. S. 2018. Cell Cycle-Associated CXCR4 Expression in Germinal Center B Cells and Its Implications on Affinity Maturation. *Front Immunol,* 9**,** 1313.

WHERRY, E. J. 2011a. T cell exhaustion. *Nature Immunology,* 12**,** 492-499.

WHERRY, E. J. 2011b. T cell exhaustion. *Nat Immunol,* 12**,** 492-9.

WHERRY, E. J., BLATTMAN, J. N., MURALI-KRISHNA, K., VAN DER MOST, R. & AHMED, R. 2003. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *J Virol,* 77**,** 4911-27.

WHERRY EJ, H. S., KAECH SM, HAINING WN, SARKAR S, KALIA V, SUBRAMANIAM S, BLATTMAN JN, BARBER DL, AHMED R. 2007. Molecular signature of CD8+ T cell exhaustion during chronic viral infection. *Immunity,* 27**,** 670-84.

WHITNEY, A. R., DIEHN, M., POPPER, S. J., ALIZADEH, A. A., BOLDRICK, J. C., RELMAN, D. A. & BROWN, P. O. 2003. Individuality and variation in gene expression patterns in human blood. *Proc Natl Acad Sci U S A,* 100**,** 1896-901.

WIDEROFF, L., SCHIFFMAN, M. H., NONNENMACHER, B., HUBBERT, N., KIRNBAUER, R., GREER, C. E., LOWY, D., LORINCZ, A. T., MANOS, M. M., GLASS, A. G. & ET AL. 1995. Evaluation of seroreactivity to human papillomavirus type 16 virus-like particles in an incident case-control study of cervical neoplasia. *J Infect Dis,* 172**,** 1425-30.

WILLOUGHBY, J. E., KERR, J. P., ROGEL, A., TARABAN, V. Y., BUCHAN, S. L., JOHNSON, P. W. & AL-SHAMKHANI, A. 2014. Differential impact of CD27 and 4-1BB costimulation on effector and memory CD8 T cell generation following peptide immunization. *J Immunol,* 193**,** 244-51.

WILSON, T. J., FUCHS, A. & COLONNA, M. 2012. Cutting edge: human FcRL4 and FcRL5 are receptors for IgA and IgG. *J Immunol,* 188**,** 4741-5.

WING, K. & SAKAGUCHI, S. 2010. Regulatory T cells exert checks and balances on self tolerance and autoimmunity. *Nat Immunol,* 11**,** 7-13.

WIRSING, A. M., RIKARDSEN, O. G., STEIGEN, S. E., UHLIN-HANSEN, L. & HADLER-OLSEN, E. 2014. Characterisation and prognostic value of tertiary lymphoid structures in oral squamous cell carcinoma. *BMC Clin Pathol,* 14**,** 38.

WOLCHOK, J. D., KLUGER, H., CALLAHAN, M. K., POSTOW, M. A., RIZVI, N. A., LESOKHIN, A. M., SEGAL, N. H., ARIYAN, C. E., GORDON, R. A., REED, K., BURKE, M. M., CALDWELL, A., KRONENBERG, S. A., AGUNWAMBA, B. U., ZHANG, X., LOWY, I., INZUNZA, H. D., FEELY, W., HORAK, C. E., HONG, Q., KORMAN, A. J., WIGGINTON, J. M., GUPTA, A. & SZNOL, M. 2013. Nivolumab plus ipilimumab in advanced melanoma. *N Engl J Med,* 369**,** 122-33.

WOOD, O., CLARKE, J., WOO, J., MIRZA, A. H., WOELK, C. H., THOMAS, G. J., VIJAYANAND, P., KING, E. & OTTENSMEIER, C. H. 2017. Head and Neck Squamous Cell Carcinomas Are Characterized by a Stable Immune Signature Within the Primary Tumor Over Time and Space. *Clin Cancer Res,* 23**,** 7641-7649.

WOOD, O., WOO, J., SEUMOIS, G., SAVELYEVA, N., MCCANN, K. J., SINGH, D., JONES, T., PEEL, L., BREEN, M. S., WARD, M., GARRIDO MARTIN, E., SANCHEZ-ELSNER, T., THOMAS, G., VIJAYANAND, P., WOELK, C. H., KING, E. & OTTENSMEIER, C. 2016a. Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget*.

WOOD, O., WOO, J., SEUMOIS, G., SAVELYEVA, N., MCCANN, K. J., SINGH, D., JONES, T., PEEL, L., BREEN, M. S., WARD, M., GARRIDO MARTIN, E., SANCHEZ-ELSNER, T., THOMAS, G., VIJAYANAND, P., WOELK, C. H., KING, E., OTTENSMEIER, C. & CONSORTIUM, S. 2016b. Gene expression analysis of TIL rich HPV-driven head and neck tumors reveals a distinct B-cell signature when compared to HPV independent tumors. *Oncotarget,* 7**,** 56781-56797.

WOYACH, J. A., JOHNSON, A. J. & BYRD, J. C. 2012. The B-cell receptor signaling pathway as a therapeutic target in CLL. *Blood,* 120**,** 1175-84.

WU, D. Y., SEGAL, N. H., SIDOBRE, S., KRONENBERG, M. & CHAPMAN, P. B. 2003. Cross-presentation of disialoganglioside GD3 to natural killer T cells. *J Exp Med,* 198**,** 173-81.

WYATT, A. W., MO, F., WANG, K., MCCONEGHY, B., BRAHMBHATT, S., JONG, L., MITCHELL, D. M., JOHNSTON, R. L., HAEGERT, A., LI, E., LIEW, J., YEUNG, J., SHRESTHA, R., LAPUK, A. V., MCPHERSON, A., SHUKIN, R., BELL, R. H., ANDERSON, S., BISHOP, J., HURTADO-COLL, A., XIAO, H., CHINNAIYAN, A. M., MEHRA, R., LIN, D., WANG, Y., FAZLI, L., GLEAVE, M. E., VOLIK, S. V. & COLLINS, C. C. 2014. Heterogeneity in the inter-tumor transcriptome of high risk prostate cancer. *Genome Biol,* 15**,** 426.

XU, Y., XU, J., YANG, Y., ZHU, L., LI, X. & ZHAO, W. 2018. SRGN Promotes Colorectal Cancer Metastasis as a Critical Downstream Target of HIF-1alpha. *Cell Physiol Biochem,* 48**,** 2429-2440.

YANG, X., CHENG, Y. & LI, C. 2017. The role of TLRs in cervical cancer with HPV infection: a review. *Signal Transduct Target Ther,* 2**,** 17055.

YUAN, D., GUO, Y. & THET, S. 2013. Enhancement of antigen-specific immunoglobulin G responses by anti-CD48. *J Innate Immun,* 5**,** 174-84.

YUSEFF, M. I., PIEROBON, P., REVERSAT, A. & LENNON-DUMENIL, A. M. 2013. How B cells capture, process and present antigens: a crucial role for cell polarity. *Nat Rev Immunol,* 13**,** 475-86.

ZHANG, H. & CHEN, J. 2018. Current status and future directions of cancer immunotherapy. *J Cancer,* 9**,** 1773-1781.

ZHANG, N. & BEVAN, M. J. 2011. CD8(+) T cells: foot soldiers of the immune system. *Immunity,* 35**,** 161-8.

ZHANG, P., MIRANI, N., BAISRE, A. & FERNANDES, H. 2014. Molecular heterogeneity of head and neck squamous cell carcinoma defined by next-generation sequencing. *Am J Pathol,* 184**,** 1323-30.

ZHANG, Y., MEYER-HERMANN, M., GEORGE, L. A., FIGGE, M. T., KHAN, M., GOODALL, M., YOUNG, S. P., REYNOLDS, A., FALCIANI, F., WAISMAN, A., NOTLEY, C. A., EHRENSTEIN, M. R., KOSCO-VILBOIS, M. & TOELLNER, K. M. 2013. Germinal center B cells govern their own fate via antibody feedback. *J Exp Med,* 210**,** 457-64.

ZHANG, Y., TECH, L., GEORGE, L. A., ACS, A., DURRETT, R. E., HESS, H., WALKER, L. S. K., TARLINTON, D. M., FLETCHER, A. L., HAUSER, A. E. & TOELLNER, K. M. 2018. Plasma cell output from germinal centers is regulated by signals from Tfh and stromal cells. *J Exp Med,* 215**,** 1227-1243.

ZHANG, Z., DENG, Y., ZHENG, G., JIA, X., XIONG, Y., LUO, K., QIU, Q., QIU, N., YIN, J., LU, M., LIU, H., GU, Y. & HE, Z. 2017. SRGN-TGFbeta2 regulatory loop confers invasion and metastasis in triple-negative breast cancer. *Oncogenesis,* 6**,** e360.

ZHENG, Y., FENG, W., WANG, Y. J., SUN, Y., SHI, G. & YU, Q. 2019. Galectins as potential emerging key targets in different types of leukemia. *Eur J Pharmacol,* 844**,** 73-78.

ZHENG, Y., ZHA, Y., DRIESSENS, G., LOCKE, F. & GAJEWSKI, T. F. 2012. Transcriptional regulator early growth response gene 2 (Egr2) is required for T cell anergy in vitro and in vivo. *J Exp Med,* 209**,** 2157-63.

ZHOU, L., CHONG, M. M. & LITTMAN, D. R. 2009. Plasticity of CD4+ T cell lineage differentiation. *Immunity,* 30**,** 646-55.

ZHOU, X., SU, Y. X., LAO, X. M., LIANG, Y. J. & LIAO, G. Q. 2016. 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.

ZHU, C., ANDERSON, A. C., SCHUBART, A., XIONG, H., IMITOLA, J., KHOURY, S. J., ZHENG, X. X., STROM, T. B. & KUCHROO, V. K. 2005. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. *Nat Immunol,* 6**,** 1245-52.

ZHU, W., GERMAIN, C., LIU, Z., SEBASTIAN, Y., DEVI, P., KNOCKAERT, S., BROHAWN, P., LEHMANN, K., DAMOTTE, D., VALIDIRE, P., YAO, Y., VALGE-ARCHER, V., HAMMOND, S. A., DIEU-

NOSJEAN, M. C. & HIGGS, B. W. 2015. A high density of tertiary lymphoid structure B cells in lung tumors is associated with increased CD4(+) T cell receptor repertoire clonality. *Oncoimmunology,* 4**,** e1051922.

ZIKHERMAN, J., PARAMESWARAN, R. & WEISS, A. 2012. Endogenous antigen tunes the responsiveness of naive B cells but not T cells. *Nature,* 489**,** 160-4.

ZININGA, T., RAMATSUI, L. & SHONHAI, A. 2018. Heat Shock Proteins as Immunomodulants. *Molecules,* 23.

ZUNIGA, E., RABINOVICH, G. A., IGLESIAS, M. M. & GRUPPI, A. 2001. Regulated expression of galectin-1 during B-cell activation and implications for T-cell apoptosis. *J Leukoc Biol,* 70**,** 73-9.