Identification of distinct markers to differentiate natural and induced T regulatory cells in cancer

Abdullah Al-Omari

A thesis submitted in partial fulfilment of the requirements of Nottingham Trent University for the degree of Doctor of Philosophy

November 2019

Copyright statement

Dedication

This thesis is dedicated to my **parents**, my wife **Dr. Yasmeen** and my daughter **Sema** who supported me through my studies.

Acknowledgements

First of all, I would like to thank my director of studies **Dr David Boocock**, the Group Leader of the Biological Mass Spectrometry & Clinical Proteomics group in the John van Geest Cancer Research Centre, for giving me the opportunity to complete my PhD in the exciting field of Cancer Immunology and Immuno-proteomics. I am deeply grateful for his constant support, ideas and discussion I have with him during the last four years of my PhD. I would never forget his effort with me to make my project successful. I would never find such a scientist who loves the technology like him.

I would like to thank the **Applied Science Private University** represented by **Dr Haitham Abu Khadijah** and Faculty of Pharmacy represented by **Prof Iman Basheti** for granting me the scholarship to complete my PhD.

I would like to thank **Prof Robert Rees**, Founder of the John van Geest Cancer Research Centre, and **Prof Graham Pockley**, Director of the the John van Geest Cancer Research Centre, for giving me the opportunity to work in this great research environment and for their valued advice and support.

I would like to thank my co-supervisor **Dr Jayakumar Vadakekolathu**, the master of cancer research. I have never found such a passionate cancer scientist who is always committed to find a suitable treatment for cancer. He has an outstanding brain which allows him to be always the first. He improved my laboratory skills and let me to be inspired more deeply in the cancer molecular pathology. I have spent the most beautiful years working with him during my PhD. I would never forget his support, brilliant ideas, critical discussion, effort to make my project meaningful. I have not found any technique or an assay he is not familiar with. He knows everything in the cancer molecular pathology. **I could not find any Arabic or English word to describe Dr Jayakumar, but all of my respect to this great scientist**. It was an honour to work with you and Dr David.

I would like to thank my co-supervisor **Prof Sergio Rutella**, for his support during my study. I would never forget his meaningful ideas in cancer immunotherapy. He is very determined in his field, working hard to find suitable treatment for Leukaemias. Also, I would like to thank **Stephen Reeder**, the master of cell sorting, for his effort with me fixing the issues of cell sorter machine during the last four years. My project was mainly based on that machine, which means that without Stephen I would never finish my project. I would never forget how many times he helped me during these years. I don't know how the centre would be without him.

I would like to thank the great proteomics team **Dr Amanda Miles** and **Dr Clare Coveney** for their help in the analysis of proteomics data, **Dr Gemma Foulds, Dr Simon Hood** for their help in the field of flow cytometry. Many thanks to **Anne Schneider** for her help in the lab, especially finding the hidden reagents.

A special thanks to **Dr Tarik Regad** for his valued ideas, knowledge, and his support during my studying. I would never forget his words "you **can do it Abdullah**"; these words encouraged me to solve any problem I have faced during my studying.

All my thanks to **Dr Murrium Ahmad**, **Dr Stephanie McArdle, Andrew Marr** and other staffs of the John van Geest Cancer Research Centre for being great colleagues and friends.

A BIG thanks to the John van Geest students, Pauline Le Vu, Sarra Idri, Rukaia, Divya, Melissa, Jenny, Sarah Wagner, Joshua, Anna Di Biase, Arif Surani, Luisa Barbato, Shaymaa and Payton. I enjoyed my days working together in the lab.

Finally, I would like to thank my brothers, sisters and relatives who support and encourage me to complete my studies.

Abstract

Regulatory T cells (Tregs) are a distinct subset of CD4+ T cells that play a vital role in maintaining immune homeostasis and peripheral tolerance, thereby preventing autoimmunity. Tregs are generally categorised into two main subsets; natural and induced Tregs. In cancer, Tregs are found extremely enriched in the tumour microenvironment and contribute to the inhibition of anti-tumour response and tumour progression. The origin of tumour-infiltrating Tregs (whether it is nTregs or iTregs) is still enigmatic, since there are no distinct biomarkers which can differentiate between the two subsets. Therefore, the aim of this study is to identify cell surface biomarkers that can differentiate phenotypic features of iTregs from nTregs in the context of cancer. With this aim, an *in vitro* murine model was successfully developed to generate $CD4^{+}CD25^{++}Foxp3^{+}$ iTregs from purely sorted naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells in the presence of TGF-β1. The induction of iTregs was assessed using flow cytometry. Methylation status of Foxp3-TSDR and Foxp3 stability was assessed. Naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells and $CD4^{+}CD25^{+}Foxp3^{+}$ nTregs were purely sorted using cell sorting. Five biologically different subsets of $CD4^{+}$ cells including naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells, activated $CD4^{+}CD25^{-}Foxp3^{-}$ T cells, naïve $CD4^{+}CD25^{+}Foxp3^{+}$ nTregs, activated $CD4^{+}CD25^{+}Foxp3^{+}$ nTregs and $CD4^{+}CD25^{++}Foxp3^{+}$ iTregs were subjected to quantitative proteomic profiling using SWATH-MS. Subcellular fractionation methods were employed to isolate membrane and cytoplasmic proteins from each of the subsets. Quantitative proteomic data were analysed using artificial neural networks. The results revealed that 4 distinct membrane biomarkers (PLP2, ITIH4, HEM6 and MAVS) were differentially up-regulated in iTregs compared to other subsets. EPHX1 (HYEP) was identified upregulated only in naïve nTregs and downregulated in iTregs and other subsets. The biomarkers were further tested. Pathway enrichment analysis of iTregs showed a distinct metabolic pathway enrichment in iTregs indicating a mechanistic insight into the iTreg development. Once validated in humans these proteins could be used as a biomarker for iTreg or as a drug target for the selective depletion for better immunotherapeutic outcome in cancer patients.

Contents

List of Figures

List of Tables

List of abbreviations

µL	microliter
µM	micromolar
ACN	Acetonitrile
ANN	Artificial neural network
ANOVA	Analysis of variance
APCs	Antigen-presenting cells
BCR	B-cell receptor
BSA	Bovine serum albumin
BU	Bottom-up
CD	Cluster of differentiation
cDCs	Conventional dendritic cells
CE	Capillary electrophoresis
CFSE	Carboxyfluorescein succinimidyl ester
CLPs	Common committed lymphoid progenitors
CNS1	Conserved non-coding sequence 1
Conf	Confidence
CPOX	Oxygen-dependent coproporphyrinogen-III oxidase
CTLA-4	Cytotoxic T-lymphocyte-antigen 4
Cyto	Cytoplasmic
DCs	Dendritic cells
DDA	Data dependent acquisition
DIA	Data independent acquisition
DPBS	Dulbecco's phosphate buffered saline
DTT	Dithiothreitol
EB	Erika's Buffer
ELISA	Enzyme linked immunosorbent assays
ELPs	Early lymphoid progenitors
EMT	epithelial-mesenchymal transition
EPHX-1	Epoxide hydrolase-1
ESI	electrospray ionisation
ETPs	T-lineage progenitors
EXP	Experiment
FA	Formic Acid
FDR	False Discovery rate
Foxp3	Forkhead box P3
GBM	glioblastoma
h	hour
HCC	hepatocellular carcinoma
HIF-1α	Hypoxia-inducible factor 1α
HNSCC	head and neck squamous cell carcinomas
HPLC	High performance liquid chromatography
HSCs	Haematopoietic stem cells

IAA	iodoacetamide
IDA	information dependent acquisition
IFN-γ	Interferon gamma
IL-	Interleukin -
ITIH4	Inter-alpha-trypsin inhibitor heavy chain H4
iTRAQ	isobaric tag for absolute and relative quantification
iTregs	Induced T regulatory cells
LC	Liquid chromatography
LFQ	Label-free quantification
LPS	Lipopolysaccharide
M	Molar
m/z	Mass to charge ratio
mAbs	Monoclonal antibodies
MALDI	Matrix assisted laser desorption ionisation
MAVS	Mitochondrial antiviral signalling protein
MD	Middle-down
mg	milligram
MHC	Major histocompatibility complex
min	minute
mL	milliliter
MLPs	Committed myeloid-lymphoid progenitors
MMP-	Matrix metalloproteinase-
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS-MS	Tandem MS
mTECs	Medullary thymic epithelial cells
n	number
ng	nanogram
NK cells	Natural Killer cells
NSCLCs	non-small cell lung carcinomas
nTregs	natural T regulatory cells
Nuc	Nuclear
p	p-value
P.Max	Protease Max
PAMPs	pathogen-associated molecular patterns
PDAC	pancreatic ductal adenocarcinoma
pDCs	plasmacytoid dendritic cells
PLP2	Proteolipid protein 2
PRM	parallel reaction monitoring
PRRs	Pattern recognition receptors
pTregs	peripheral T regulatory cells
RBCs	Red blood cells
sec	second
SEM	Standard error of the mean

SRM	selection reaction monitoring
SWATH	Sequential window acquisition of all theoretical mass spectra
TAMs	Tumour-associated macrophages
Tc cells	T cytotoxic cells
TCR	T-cell receptor
TD	Top-down
TEAB	triethyl ammonium bicarbonate
TfH cells	T follicular helper cells
TGF-β	Transforming growth factor-beta
Th cells	T helper cells
Th17	T helper 17
TLRs	Toll-Like receptors
TMT	Tandem mass tag
TNBC	triple negative breast cancer
TNF-α	tumour necrosis alpha
TOF	Time of flight
Tr1	Type 1 regulatory cells
Tregs	T regulatory cells
TSDR	Treg-specific demethylation region
U	unite
VEGF	Vascular endothelial growth factor

Chapter 1.

General introduction

1.1. Cancer – A brief overview

Cancer is the second most common cause of death worldwide. In 2018, 9.6 million deaths have been reported due to cancer, which is about 1 in 6 deaths worldwide (WHO, 2018, www.who.com, Accessed on 20 Oct 2019). In UK, cancer incidence has been estimated about 363,000 new cancer cases in 2016; 178,000 (545 per 100,000) cases are female and 185,000 (664 per 100,000) cases are male. Moreover, 53% of all new cases have been diagnosed with breast, prostate, lung and bowel cancer. In 2017, 28% of all deaths in UK have been reported due to cancer (Cancer Research UK, www.cancerresearchuk.org, Accessed on 20 Oct 2019).

Cancer is a group of diseases defined by uncontrolled cell growth through which cells acquire the ability to invade adjacent tissues and then disseminate (metastasise) from the primary site of origin to other sites in the body. More than 100 types of cancer have been identified which are characterised based on the tissue of origin where tumour develops (Knowles and Selby 2005). For instance, cancers derived from epithelial tissue are categorised as "carcinomas" and constitute around 85% of cancers. This include cancers come from glandular epithelial tissues which are called "adenocarcinomas". Whereas cancers developed in mesoderm tissues (bone, muscle and connective tissues) are called "sarcomas"(Pecorino 2012).

Several external and lifestyle factors can contribute to the development of cancer including smoking, infection, obesity, alcohol consumption, ionising and ultraviolet radiation and chemical reagents (Hesketh 2013). There is a broad consensus that cancer is a genetic disease in which accumulation of molecular lesions in the genetic material of somatic cells (DeVita 2015). In normal circumstances these molecular aberrations are corrected by our own DNA repair machinery or eliminated along with mutation bearing cells by the immune system (Bernstein, et al. 2013). Once the mutation/s evaded the repair/immune system favours the bearing cell proliferative and oncogenic advantage resulting in malignancy (Messerschmidt, Prendergast and Messerschmidt 2016).

1.2. The immune system – A brief overview

The main function of the immune system is to protect multicellular organisms from pathogens. The immune system generates an enormous array of cells and molecules which act together in a dynamic network to recognise and eliminate foreign invaders. The immune system can protect the body by two related activities; recognition and response. Immune recognition is characterised by the ability to discriminate foreign invaders from self-components. It can even perceive slight chemical differences that differentiate one foreign pathogen from another (Janeway, et al. 1996). In addition, the immune system can also recognise host cells that are abnormally transformed or on their way to become cancerous cells. Recognition of a pathogen by the immune system induces an effector response that eliminates or overcomes the invader. Each immune cell has a different pattern of response and is specialised to eliminate or neutralise a particular type of pathogen. Some exposures trigger a memory response that is characterised by a more rapid and powerful immune reaction upon later attack. This type of response is a remarkable feature of the mammalian immune system which protects the body from catching certain diseases a second time (Kindt, et al. 2007).

In general, immunity is divided into two main categories; “innate” and “adaptive” immunity which both collaborate to protect the body. The innate immune system is known as the first line of defence which can effectively eliminate pathogens within hours of encounter with innate immune cells. The adaptive immunity provides a second, comprehensive line of defence in which adaptive immune cells recognise, eliminate and memorise the invading pathogens that evade the innate response (Roitt, Brostoff and Male 2001).

Innate and adaptive immune cells are generated by a highly regulated process called “haematopoiesis” which takes place in the bone marrow. This process involves the production and development of red and white blood cells and starts when self-renewing haematopoietic stem cells give rise to lymphoid and myeloid progenitors. Lymphoid progenitors give rise to T-cell progenitors, B-cell progenitors, dendritic cells and natural killer (NK) cells, whereas myeloid progenitor give rise to granulocyte-monocyte progenitor, eosinophil progenitor, basophil progenitor,

megakaryocyte and erythroid progenitor (figure 1.2) (Kindt, et al. 2007) (Taniuchi 2018)

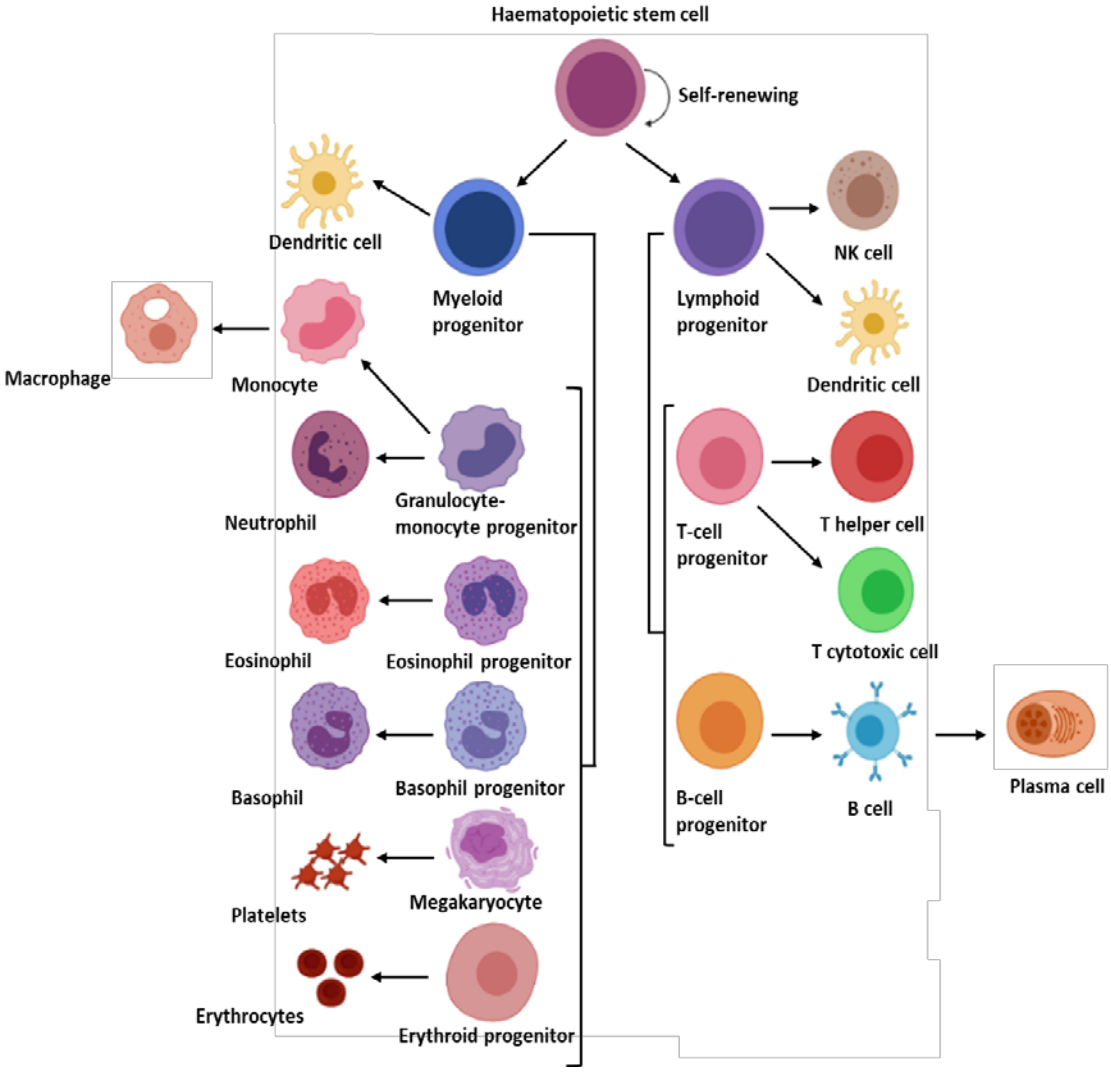


Figure 1.1. Schematic representation of haematopoiesis process by which red and white blood cells are produced in the bone marrow. The left panel shows innate immune cells that are developed from lineage committed progenitors derived from myeloid-lineage progenitors. The right panel shows adaptive immune cells that are developed from lineage committed progenitors derived from lymphoid-lineage progenitors. This figure was adaptive from (Kindt, et al. 2007).

1.2.1. Innate immunity

The innate immunity provides the first primitive type of defence mechanism which is ready for immediate activation prior to invasion by a pathogen. The innate immune system consists of three different types of barriers to protect the body, including physical, chemical and cellular barriers. The key physical barriers are skin and mucous membranes which prevent the entry of invading pathogens into the body. Chemical barriers involve the acidity of the stomach contents and specialised soluble molecules such as lysozyme that have antimicrobial activities (Tosi 2005). Once a pathogen breaches these primary barriers, it will encounter the cellular barrier which consists of various types of cells that can initiate a rapid response within minutes of invasion (Gallo and Nizet 2008).

As shown in figure 1.1, innate immune cells develop from myeloid progenitor in the bone marrow and are broadly grouped int two main classes. Mononuclear phagocytes include dendritic cells (DCs), monocytes and macrophages. polymorphonuclear phagocytes include neutrophils, basophils and eosinophils based on the shape of their nucleus (De Kleer, et al. 2014).

Innate immune cells generally express pattern recognition receptors (PRRs) and Toll-like receptors (TLRs). PRRs allow innate cells to specifically recognise broad structural motifs (pathogen-associated molecular patterns (PAMPs)) that highly conserved within pathogens but are completely absent from the host. TLRs are a family of 11 versatile receptors, each reacts with a specific microbial product, allowing innate cells to detect different types of pathogens (Mogensen 2009).

Innate immune responses are mediated by the participation of various cells. The major functional innate cells are neutrophils, macrophages, dendritic cells and NK cells (Beutler 2004). Neutrophils are the first cells to travel from the blood into the site of inflammation (pathogen invasion) where they kill and eliminate pathogen by phagocytosis (Kobayashi and DeLeo 2009). They are also important for the innate defence against bacteria and fungi. Neutrophils express several TLRs on their surfaces. The two key receptors are TLR2 and TLR4 which enables neutrophils to recognise the peptidoglycans of Gram-positive bacteria and the lipopolysaccharide

(LPS) expressed on the wall of Gram-negative microbes, respectively (Sabroe, et al. 2003).

Macrophages and dendritic cells are professional antigen presenting (Ito, et al. 2013) cells that link innate and adaptive responses via antigen presentation on MHC molecules. Macrophages are generated in the bone marrow as immature monocytes that circulate in the blood to become mature monocytes. Mature monocytes further differentiate into macrophages once they enter peripheral tissues (Mosser and Edwards 2008). Interaction between PRRs expressed on macrophages and pathogen components (PAMPs) activates macrophages which induce phagocytosis. Activated macrophages then secrete proinflammatory cytokines such as interleukin-1 (IL-1), IL-6 and tumour necrosis alpha (TNF-α) which induce and support inflammatory responses (Arango Duque and Descoteaux 2014).

DCs act as a bridge between innate and adaptive immunity as they can interact with T helper (Th) and T cytotoxic (Tc) cells. Immature DCs as innate cells recognise pathogens via PRRs and TLRs (Steinman 2006). This recognition induces maturation of DCs that results in the upregulation of MHC class II on their surfaces. In addition, as nucleated cells, DCs also express MHC class I (Dalod, et al. 2014). Upon maturation, DCs internalise pathogens components and then migrate to the lymphoid tissue (e.g. lymph nodes) where they introduce antigen (pathogens components) to MHC class II-dependent Th cells and MHC class I-dependent Tc cells. DCs can also produce IL-12, IL-6 and TNF-α (Guermonprez, et al. 2002).

NK cells play an important role in the innate immunity by their ability to kill virus-infected cells and pre-malignant transformed cells. NK cells lack membrane molecules and antigen-specific receptors (Hamerman, Ogasawara and Lanier 2005). Instead, NK cells express specific receptors which allow them to distinguish defects such as a reduction in the expression of MHC class I molecules, the unusual phenotypes expressed by tumour cells and cells infected by some viruses. NK cells have an essential role in the regulation of both innate and adaptive immunity (Brilot, Strowig and Munz 2008, Bellora, et al. 2010). They produce Interferon gamma (IFN-γ) which activates macrophages and enhances their phagocytic and anti-microbial activities. IFN-γ produced by NK cells can also affect the Th1 versus Th2 lineage by

inhibition Th2 expansion and stimulation development of Th1 in the presence of IL-12 derived from macrophages and DCs (Mailliard, et al. 2003).

One of the hallmarks of innate immunity is the lack of antigenic specificity and immunogenic memory. Therefore, if there is a recognised antigenic challenge to the pathogen which may exceed the ability of innate immune cells during inflammation, adaptive immunity will be initiated to respond to the challenge with a high degree of specificity and immunologic memory (Kindt, et al. 2007). However, a recent evidence has revealed that NK cells develop antigenic specificity and immunological memory against viruses although they are involved in the innate immunity (Peng and Tian 2017).

1.2.2. Adaptive immunity

The adaptive immunity has four different features including antigenic specificity, diversity, immunologic memory and self-nonself recognition. The antigenic specificity allows adaptive immune system to discriminate finer differences among antigens. The adaptive system can produce tremendous highly-diverse receptors which can recognise many of unique structures with minor genetic variations of foreign antigens (Bonilla and Oettgen 2010). The adaptive system can also memorise the eliminated antigens by generating memory cells that specifically react to the same antigens for future exposure, conferring lifelong immunity to many pathogens after an initial exposure. Finally, the adaptive system is educated to distinguish between self and non-self molecules and initiate immune response against non-self molecules only. Failure of this ability leads to autoimmune diseases and can be fatal (Kurtz 2004).

Adaptive immune responses are mediated by specialised cells including B and T cells in addition to other specialised antigen-presenting cells (macrophages and DCs). B cells are generated and mature in the bone marrow and express unique antigen-binding receptor (B-cell receptor (BCR)) which is a membrane-bound antibody molecule. B cells are responsible for mediating humoral immune response (LeBien and Tedder 2008). Upon antigen binding to BCR, B cells undergo activation and differentiation into effector B cells (plasma cells) and memory B cells. A single plasma cell can secrete hundreds to thousands of antibodies per second. Secreted

antibodies are considered as key effector molecules of humoral immunity (LeBien and Tedder 2008).

Unlike B cells, T cells are generated in the bone marrow as progenitor T cells that migrate to thymus where they become mature. T cell maturation includes rearrangements of the germ line TCR genes and the expression of various membrane markers such as CD3, CD4, CD25 and CD8 (Koch and Radtke 2011). Developing T cells in the thymus are defined as thymocytes that undergo a highly regulated selection process, called thymic selection. Positive thymic selection allows only T cells with TCRs that can recognise self-MHC molecules to continue as mature self-MHC-restricted T cells. T cells that react too strongly with self-peptide mounted on MHC molecules are eliminated by negative selection (Klein, et al. 2014a). Negative selection is responsible for the creation of primary T cells that are self-tolerant (Takaba and Takayanagi 2017). Mature T cells are grouped into two main subpopulations; Th and Tc cells. Th cells express CD4, a membrane glycoprotein receptor that recognises antigens bound to MHC class II molecules. Tc cells express CD8 that recognises antigens bound to MHC class I molecules (Murphy and Weaver 2016).

The key event in the development of cell-mediated immune responses is the activation and clonal expansion of T cells. T cell activation is triggered by interaction of the TCR-CD3 complex with a processed antigenic peptide bound to either MHC class I ($CD8^+$ T cells) or class II ($CD4^+$ T cells) molecule expressed on the surface of antigen-presenting cells (APCs) (Smith-Garvin, Koretzky and Jordan 2009). However, this interaction is not sufficient to fully activate naïve T cells. For subsequent proliferation into effector cells, naïve T cells receive two essential activation signals. The first signal is initiated by interaction of an antigenic peptide with the TCR-CD3 complex. The second signal is a consecutive antigen-nonspecific costimulatory signal provided by interactions between CD28 on T cells and receptors of the B7 family expressed on the surface of APCs; B7-1 (CD80) and B7-2 (CD86) (figure 1.2) (Chen and Flies 2013). B7 receptors are mainly expressed on dendritic cells and induced on activated macrophages and B cells (Mbongue, et al. 2017). Costimulatory signals derived by CD28 are regulated by CTLA-4 (CD152), a glycoprotein which is structurally similar to CD28. CD152 provides inhibitory

signals that down-regulate the activation of T cells (Rudd, Taylor and Schneider 2009). Activation of T cells by the interaction of TCR-CD3 with an antigenic peptide in the absence of CD28 costimulatory signalling leads to T cell anergy, which a state of nonresponsive characterised by failure of T cells to proliferate (Schwartz 2003). This confirms the importance of both signals in the activation of T cells. However, the strength of activation signals can also affect the function of T cells. It is found that persistent antigen-derived TCR activation signals lead to T cell exhaustion which is a state when T cells lose their functional ability as effector cells following prolonged antigen stimulation (Wherry 2011).

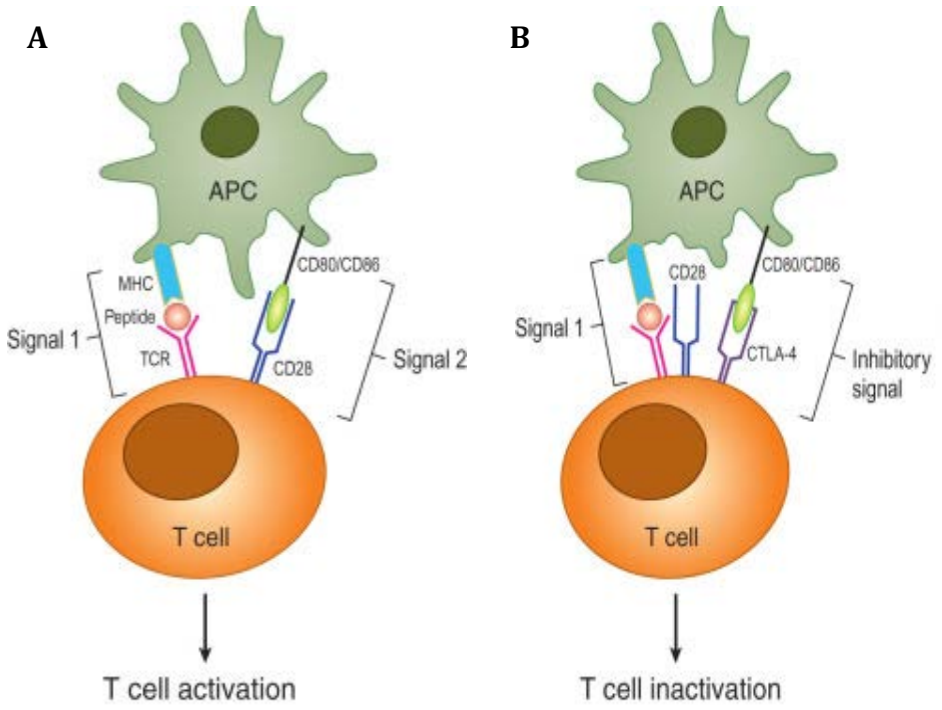


Figure 1.2. Signals of T cells activation and inactivation mediated by CD80/86 receptors expressed on APC. Signa1 is generated by the interaction of antigenic peptide attached with MHC molecule. Signal 2 is a costimulatory signal derived by the interaction of B7 receptors family (CD80/CD80) with CD28 expressed on the surface of T cells. Costimulatory signal is regulated by CTLA-4 which induce inhibitory signals to deactivate T cells. This figure was adapted from (Sharma, et al. 2019). Copyright permission was obtained from the publisher (Elsevier, license number 4855390275342, license date on 24 Jun 2020 by Copyright Clearance Center).

Upon activation, naïve T cells differentiate into effector T cells. Naïve $CD8^+$ T cells differentiate into $CD8^+$ effector T cells with cytotoxic killing activity and memory $CD8^+$ T cells. Naïve $CD4^+$ T cells differentiate into several subpopulations based on secreted cytokines in the local microenvironment. Figure 1.3 shows differentiation of naive $CD4^+$ T cells in the presence of cytokines that are required for lineage commitment of each subsets of $CD4^+$ T cells.

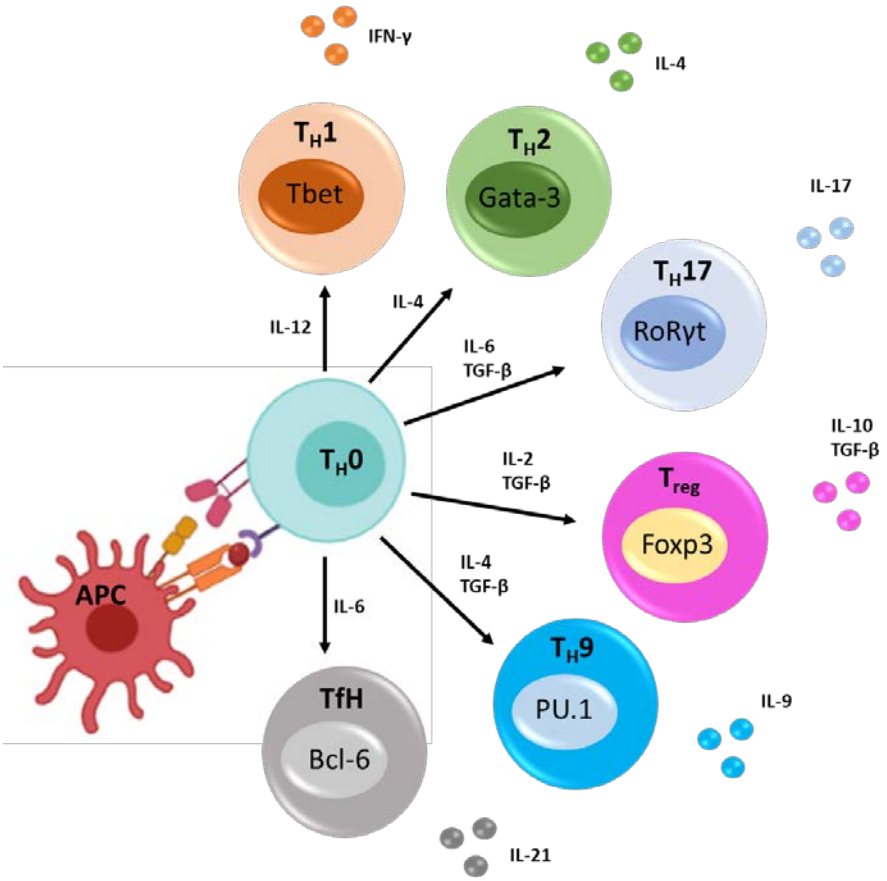


Figure 1.3. Differentiation of naïve CD4⁺ T cells (Th0) upon activation by antigenic peptides represented by MHC class II molecule expressed on the surface of an antigen-presenting cell (APC) in the presence of other required cytokines. Th1 cells are differentiated in the presence of IL-12 cytokines and secrete IFN-γ and IL-2 after differentiation. The differentiation of Th1 is regulated by the transcription factor T-bet that maintains their lineage-specific development. The differentiation of Th2 cells is regulated by the transcription factor Gata-3 and induced in the presence of IL-4. Th2 cells are characterised by the secretion of various interleukins such as IL-4, IL-10, IL-5 and IL-13. Th17 cells are differentiated in the presence of TGF-β and IL-6, and their differentiation is controlled by the transcription factor RoRγt. Upon differentiation, Th17 cells secrete IL-17, IL-22 and TNF-α. Tregs are characterised by the expression of the transcription factor Foxp3 and differentiated in the presence of IL-2 and TGF-β. Tregs secrete TGF-β and IL-10 immunosuppressive cytokines. Th9 cells are characterised by the expression of PU.1 transcription factor and differentiated by the effect of TGFβ and IL-4. Th9 cells are found to secrete IL-9 and IL-10 to mediate their function. TfH (T follicular helper) cells are defined by the expression of Bcl-6 transcription factor and differentiated in the presence of IL-6. TfH cells produce high level of IL-21. This figure was adapted from (Russ, et al. 2013).

Th1 cells secrete several cytokines including IL-2, IFN-γ and TNF-α/β which can induce cell-mediated immune response by activating $CD8^+$ T cytotoxic cells against invading pathogens or altered host cells. Production of IL-4 by Th2 can effectively induce humoral immune response by activating B cells to produce antibodies against pathogens (Zhang, et al. 2014). The function of other subsets of $CD4^+$ T cells (Th17 and Th9 cells) is mentioned in the section (1.5.1.3).

The innate and adaptive immune cells play pivotal role in the recognition and elimination host transformed-malignant cells. However, another evidence confirms that some of these immune cells contribute to the progression of tumour (Vesely, et al. 2011). The cellular interactions between immune and tumour cells which result in tumour escape from immune response will be discussed in the next section.

1.3. Cancer and immune system

The interaction between immune system and cancer is widely acknowledged in the literature. Inflammation in the context of cancer plays a significant role in the progression of cancer. Inflammation is a complex physiological process triggered by various immune cells in response to infection or tissue injury, thereby providing a full protection against dangerous pathogens (Barton 2008). This highly regulated response can be detrimental if it becomes dysregulated which results in pathological consequences including chronic inflammatory diseases, autoimmunity, systemic sepsis, tissue damage and fibrosis, neoplasia (Medzhitov 2008). In cancer, it has been reported that inflammation is an essential process at different stages of tumour growth and progression as it can promote tumorigenesis by altering the host immune response to tumour (Grivennikov, Greten and Karin 2010). For instance, chronic *Helicobacter pylori* inflammation can lead to gastric cancer and mucosa-associated lymphoid tissue lymphoma, whereas persistent infection with hepatitis B and C viruses is associated with hepatocellular carcinoma (Kew 2013, Correa and Piazuelo 2011, J. B. Park and Koo 2014). In cancer-related inflammation, at the initial stage of tumorigenesis, the tumour microenvironment contains various immune cells which recognise tumour-expressed antigens and produce inflammatory and potentially protective anti-tumour responses. However, the subversion of these responses by tumour cells enables the establishment of a persistent inflammatory

microenvironment (low grade chronic inflammation) which further promote tumour development (Grivennikov, et al. 2010). A wide array of immune cells infiltrating with the tumour microenvironment, which include macrophages, neutrophils, mast cells, myeloid-derived suppressor cells (MDSCs), dendritic cells (DCs), natural killer (NK) cells, T and B lymphocytes (Mantovani, et al. 2008). The inflammatory tumour microenvironment can also be regulated by certain oncogenes such as MYC and RAS which can recruit leukocytes, lymphocytes, and maintaining a proinflammatory microenvironment tumour, thus remodelling the tumour microenvironment (Soucek, et al. 2007).

Accumulating evidence confirms that there is a dynamic cellular communication between tumours and immune cells which can regulate tumour growth and progression. This evidence has been translated into a concept of "cancer immunoediting" which postulated that the host's immune system can inhibit cancer growth and promote cancer progression simultaneously. Cancer immunoediting consists of three different phases including elimination, equilibrium and escape (Dunn, Old and Schreiber 2004). The concept of immunoediting which is characterised by three distinct phases is discussed in detail in the following sections.

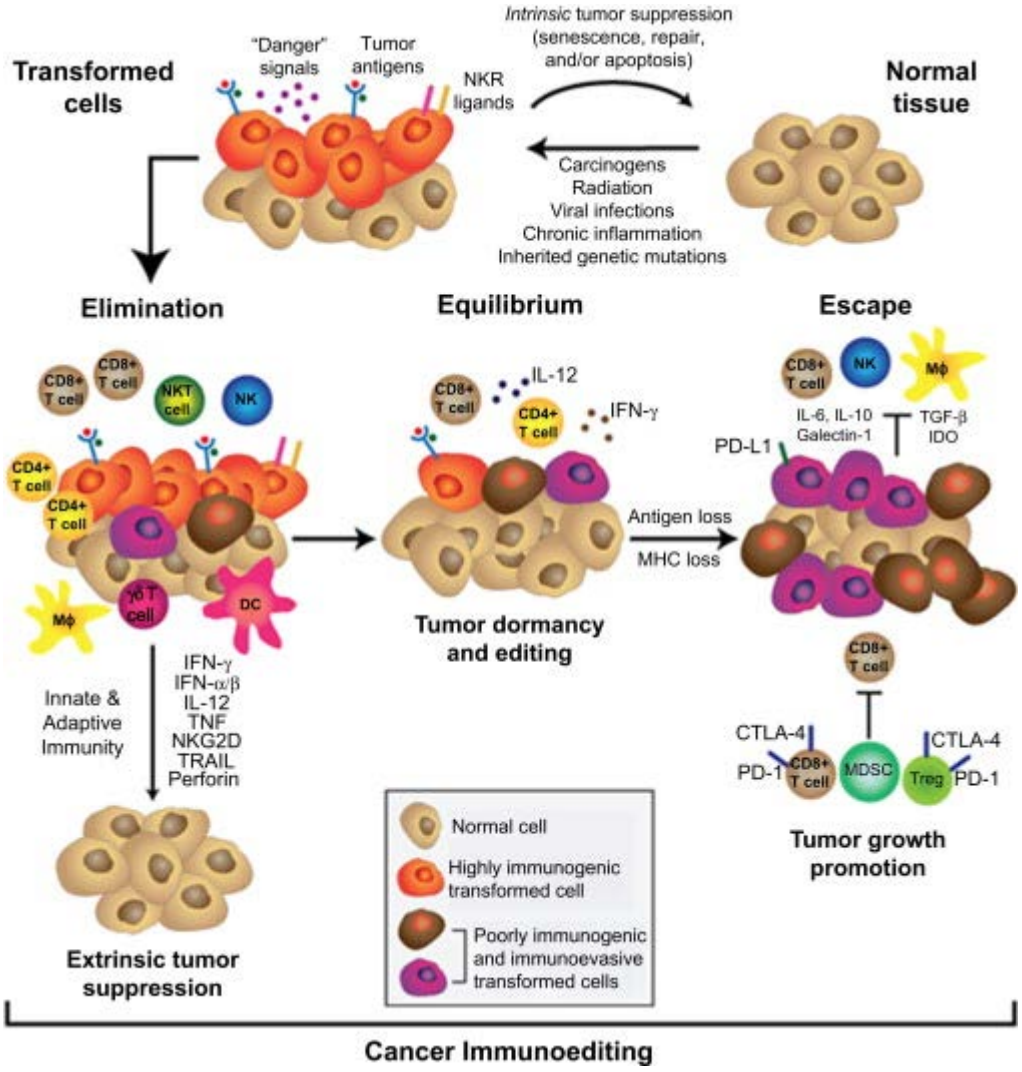


Figure 1.4. Schematic representation showing the stages of cancer immunoediting. In the first phase (elimination), innate (NK, macrophages, DCs) and adaptive immune cells ($CD4^+$ and $CD8^+$ and γδ T cells) recognise pre-malignant lesions (newly transformed cells and kill growing tumours by producing several cytokines, thus maintaining protection against cancer. In the second phase (equilibrium), poorly-immunogenic tumour cells are thought to be specifically selected by the tumour to maintain its progression by avoiding the intensive, constant attack from the immune cells ($CD4^+$ and $CD8^+$, NK cells) that still have potential to eliminate the newly growing cells. In the third phase (escape), tumour cells that have gained the competency to evade immune recognition and destruction sprout progressively by creating immunosuppressive tumour microenvironment through recruiting Tregs, MDSCs that can inhibit the function of $CD8^+$ T cells and NK cells and several other immune evasion mechanisms. This graph was adapted from (Schreiber, Old and Smyth 2011). Copyright permission was obtained from the publisher (The American Association for the Advancement of Science, license number 4855370109321, License date on 24 Jun 2020 by Copyright Clearance Center).

1.3.1. The phase of tumour elimination (Cancer immunosurveillance)

The elimination phase (which is also known as "immunosurveillance") starts when the innate and adaptive immune cells recognise pre-malignant lesions (newly transformed cells and kill growing tumours, thus maintaining protection against cancer (Shankaran, et al. 2001). It has been proposed that the initial anti-tumour immune response is induced by innate immune cells triggered due to the local tissue disruption. The tissue disruption occurs as a results of angiogenesis or tissue-invasive growth (Carmeliet and Jain 2000, Hanahan and Weinberg 2000). These two processes are believed to promote the tissue-stromal remodelling which produces proinflammatory molecules that direct the innate immune cells to the site of tissue damage. These proinflammatory molecules include interleukin (IL-1), TNF-α (tumour necrosis factor alpha), type 1 IFNs (interferon), GM-CSF (granulocyte-macrophage colony-stimulating factor) and IL-15 (Smyth, Dunn and Schreiber 2006). These proinflammatory molecules mediate the differentiation and function of DCs which have a direct cellular communication with NK cells and T cells (Blanco, et al. 2008). DCs are thought to serve as sentinel cells that monitor tissue stress, damage and transformation, secretion of proinflammatory molecules derived from extracellular matrix breakdown, and secretion of heat-shock proteins (HSPs) as a result of tumour cell necrosis and damage (Hanke, et al. 2013).

Several cells have been proposed to participate in the phase of cancer immunosurveillance. NK cells, macrophages, NKT cells, $CD4^+$ and $CD8^+$ T cells and $\gamma\delta^+$ T cells are thought to be recruited to the site of pre-malignant transformed tissue in which they become fully activated and exert their protective response (Chow, Möller and Smyth 2012) (figure 1.4). NK cells are thought to be activated through their activating receptor NKG2D (natural killer group 2, member D) which can bind to the stress ligands expressed by the developing tumour cells, infected, stressed or transformed cells (Waldhauer and Steinle 2008). These stress ligands include MHC-class I-polypeptide-related sequence A (MICA) and MICB. Upon activation, NK cells secrete perforin and various granzymes which can permeabilise the plasma membrane of tumour cells and initiate cascades of apoptosis (Zitvogel, Tesniere and Kroemer 2006). Secretion or expression of tumour necrosis factor (TNF)-related

apoptosis-inducing ligand (TRAIL) by NK cells can kill tumour cells which express the corresponding receptor "TRAIL receptor" (Falschlehner, et al. 2009). The developing tumours can also be recognised by T cells and NKT cells via the cellular interaction between T-cell receptors (TCRs) and MHC-tumour associated peptide complexes and glycolipid-CD1 complexes which are expressed by the growing tumour cells. Upon activation, cytotoxic effector cells, mainly $CD8^+$ T cells and NKT cells, exert their cytotoxic mechanisms to eliminate tumour or premalignant-transformed cells by secreting various types of IFNs including type I IFNs (IFN-α and IFN-β) and type II IFN (IFN-γ) (Krijgsman, Hokland and Kuppen 2018). The production of IFNs, mainly IFN-γ, can mediate the recruitment of other innate immune cells to the site of growing tumour. The production of IFN-γ inhibits angiogenesis mediated by tumour cells and induce proapoptotic mechanisms within tumour cells which inhibit the growth of tumour cells (Sun, et al. 2014).

Apoptotic or lysed tumour cells release intracellular molecules and antigens that recruit adaptive immune cells (DCs, macrophages, $CD4^+$ and $CD8^+$ T cells) and initiate the development of tumour-specific adaptive immune responses. Signals derived from tumour infiltrating innate immune cells can also induce the recruitment and activation of adaptive immune cells (Topfer, et al. 2011). For instance, immature DCs can be recruited to the site of growing tumour and activated by the cytokine milieu of innate immunity (Palucka, et al. 2011). Upon activation, immature DCs encounter and present tumour antigens on their cell surface in two different ways either directly by ingesting tumour cell apoptotic bodies or indirectly when tumour antigens become bound and chaperoned with heat shock proteins which are taken up by DCs (McDonnell, Robinson and Currie 2010, Murshid, Gong and Calderwood 2012). After antigen internalisation, DCs acquire a highly activated-mature phenotype by which they can activate naïve tumour-specific T helper 1 (Th1) $CD4^+$ T cells via MHC (major histocompatibility complex) class II signalling. Th1 mediate the activation of naïve tumour-specific $CD8^+$ T cells induced through cross-presentation of antigenic tumour peptides on DC-MHC class I complexes (Sánchez-Paulete, et al. 2017).

The activation of tumour-specific adaptive immune response provides the host with a potential to effectively eliminate the growing tumour. At the site of developing

tumour, $CD4^+$ T cells secrete IL-2 that maintains the function and viability of $CD8^+$ T cells (Gajewski, Schreiber and Fu 2013). Production of IL-15 due to tissue-stromal remodelling can also maintain cytotoxic activity of tumour-specific $CD8^+$ T cells which carefully recognise and attack antigen-positive tumour cells (Doedens, et al. 2016). After interaction with tumour cells, tumour-specific $CD8^+$ T cells can also secrete large amount of IFN-γ and thus kill tumour cell via IFN-γ-dependent mechanisms of apoptosis and angiogenesis inhibition (Bruno, et al. 2014). The contribution of tumour-specific adaptive immunity may vary among different tumours. Therefore, the elimination phase of cancer immunoediting is mainly based on the immunogenicity of antigens expressed by newly-growing tumour and must be repeated each time when new antigenically distinct tumour cells develop (Vesely and Schreiber 2013).

1.3.2. The phase of equilibrium (cancer persistence)

Tumour cells which have escaped the elimination phase can persist and resist the activated tumour-specific adaptive immune cells via expression of non-immunogenic antigens (Quezada, et al. 2011) (figure 1.4). These non-immunogenic tumour cells are thought to be specifically selected by the tumour to maintains its progression by avoiding the intensive, constant attack from the immune cells that still have potential to eliminate the newly growing cells. This process has been called "immunoselection" (Zitvogel, et al. 2006). The equilibrium phase (immunoselection) is a period of cancer persistence in which tumour cells are prevented from expanding by immune resistance. In this stage, tumour cells show a functional state of dormancy and may reprogram their intrinsic mechanisms which block antigen-processing machinery and therefore inhibit expression of immunogenic molecules that can be recognised by immune cells (Teng, et al. 2008). Loss of MHC class I molecules, TAP1 (transporter associated with antigen processing 1), LMP2 (low-molecular mass protein 2) and IFN-γR (IFN-γ receptor) helps tumour cells to avoid adaptive immune response which is mediated via a process called immune-mediated tumour dormancy (Starling 2017). The molecular mechanisms of immune-mediated tumour dormancy in the equilibrium phase are still not comprehensively explained due to the difficulties of developing a mouse model that can represent the equilibrium phase (Mittal, et al. 2014). An evidence

confirming the existence of equilibrium phase has been reported by Koebel et al. The authors have showed that immunocompetent mice failed to develop any palpable tumours after receiving low-dose carcinogen (3'-methylcholanthrene (MCA)) which induced occult cancer cells within mice. However, 50% of these mice rapidly developed tumours at the site of MCA injection when their immune system was inhibited by administrating monoclonal antibodies targeting T cells and IFN-γ (Koebel, et al. 2007). The authors have also demonstrated that adaptive immune cells, specifically $CD4^+$, $CD8^+$, IL-12-secreting and IFN-γ-secreting T cells, but not innate immune cells were responsible for preserving occult tumour cells in a dormant state (Koebel, et al. 2007). Another study has also used the MCA-induced occult tumour mouse model to investigate the equilibrium phase of immunoediting (Teng, et al. 2012). The authors suggested that this specific phase could be the longest in cancer immunoediting as primary tumours were apparent within mice after 10-20 weeks of carcinogen inoculation, while occult tumour cells in a state of dormancy for 200 days before T-cell depletion which promoted the tumour progression (Teng, et al. 2012). The authors have also demonstrated that IL-12 and IL-13 cytokines might be the main regulators that coordinate the molecular mechanisms of immune-mediated tumour dormancy in the equilibrium phase. IL-23 has been shown to maintain tumour growth and persistence, whereas IL-12 was efficient to inhibit and control the outgrowth of occult primary tumour cells. Whereas, IL-4, IL-17A, TNF, and IFN-α and β cytokines were not shown to have any role in mediating the equilibrium phase (Teng, et al. 2012). Wu et al. have profiled immune cells enriched in the microenvironment of tumours at the equilibrium and escape phases. The authors have found that $CD8^+$ T cells, NK cells and $\gamma\delta^+$ T cells were predominantly enriched in tumours undergoing immune-equilibrium phase, whereas NKT cells, Tregs and MDCSs were present in very low percentages. However, the microenvironment of tumours in the immune-escape phase was heavily enriched with immunosuppressive cells, mainly T regulatory cells (Tregs) and MDSCs (Wu, et al. 2013).

1.3.3. The phase of immune escape (cancer progression)

Tumour cells that have gained the competency to evade immune recognition and destruction sprout progressively and therefore enter the escape phase, which is the

last phase of cancer immunoediting (Kim, Emi and Tanabe 2007). Tumour escape is mediated through different mechanisms, including development of immunosuppressive tumour microenvironment by the recruitment and/or induction of immunosuppressive/immunoregulatory T cell populations, reduced immune recognition of T cells due to the absence of immunogenic antigens/MHCs expressed on tumour cells, overexpression of anti-apoptotic molecules by tumour cells and immune exhaustion (Mittal, et al. 2014, Schreiber, Old and Smyth 2011) (figure 1.4). Understanding the molecular biology of these mechanisms has provided new insights for the development of immunotherapy against cancer (Beatty and Gladney 2015). These mechanisms will be discussed in more details in the next section.

1.4. Immune escape mechanisms in the tumour microenvironment

Accumulative evidence has suggested that immune evasion is mediated by tumour cell-mediated mechanisms and immune cell-mediated mechanisms; both of which contribute to resistance to cancer immunotherapies and promote tumour invasiveness, progression and metastasis (Bhutia, Mallick and Maiti 2010).

1.4.1. Tumour cell-mediated mechanisms

Several studies have reported that the cellular microenvironment may vary between tumours. Some tumours are heavily infiltrated by immune cells which have been called T-cell-inflamed tumours, while other that lack a baseline of immune-cell infiltration are called non T cell-inflamed tumours in which T cells and other immune cells are precluded from the tumour microenvironment (Trujillo, et al. 2018). It has been reported that tumour cells can employ several intrinsic mechanisms that mediate T cell exclusion (Trujillo, et al. 2018).

1.4.1.1. WNT/β-catenin pathway

Tumour cell-intrinsic upregulation of the WNT/β-catenin pathway has been correlated with absence of T cell infiltration in the tumour microenvironment. Spranger et al. (2015) have found that genes related WNT/β-catenin pathway are significantly upregulated in non-T cell-inflamed metastatic melanoma tumours,

compared to T cell-inflamed tumours. By using inducible genetically engineered autochthonous mouse model, the authors have demonstrated that activation of WNT/β-catenin signalling due to gain-of-function mutations could inhibit the initial steps of T cell priming by tumour antigens. Moreover, the authors have revealed that tumours with active β-catenin failed to recruit and activate Batf3-encoded $CD103^+$ DCs in the microenvironment, which was correlated with lacking T cell infiltration within the tumour microenvironment (Spranger, Bao and Gajewski 2015). Batf3 DCs are found to be indispensable mediators for priming and activation of $CD8^+$ T cells in the tumour microenvironment via cross-presentation of tumour-associated antigens (Engelhardt, et al. 2012, Fuertes, et al. 2011, Hildner, et al. 2008). T cell exclusion due to activation of the WNT/β-catenin signalling has been found in the tumour microenvironment of other cancers including bladder cancer, head and neck and other types of cancer (Sweis, et al. 2016, Luke, et al. 2016). Abnormal activation of the WNT/β-catenin signalling has been found to suppress the expression of Eomesoderin (EOMES) which is the master regulator of the function of $CD8^+$ T effector cells (Gattinoni, Ji and Restifo 2010). Activation of the WNT/β-catenin signalling due to functional mutations has found to affect the role of DCs and condition them to a regulatory state which inactivates anti-tumour immunity (Castañeda-Patlán, Fuentes-García and Robles-Flores 2018). Activation of β-catenin has also found to enhance the expression of the transcription repressor ATF3 that inhibits the transcription of CCL4. The insufficient production of CCL4 impairs the maturation and activation of Batf3-$CD103^+$ DCs that are essential for $CD8^+$ T cells priming and differentiation (Pai, et al. 2017).

1.4.1.2. PTEN/PI3K pathway

Another pathway found to mediate T cell exclusion within the microenvironment of non-T cell-inflamed tumours is loss of PTEN/PI3K activation pathway. PTEN (the phosphatase and tensin homologue) is a well-known tumour suppressor which negatively regulates the PI3K/Akt signalling pathway (Song, Salmena and Pandolfi 2012). PI3K/Akt signalling pathway plays a pivotal role in maintaining cell growth and survival and has been found significantly upregulated in tumour cells, thereby maintaining the tumour progression (Porta, Paglino and Mosca 2014). In preclinical models of melanoma, loss of PTEN in tumour cells is found to promote the expression of immunosuppressive cytokines, suppress T cell-mediated anti-tumour

response and reduce T cell infiltration into the tumour microenvironment (Peng, et al. 2016). Loss of PTEN pathway in prostate tumour cells has also been found to induce immunosuppressive tumour microenvironment via downregulation of PTPN11/SHP2 which therefore maintains the activation of the JAK2/STAT3 pathway (Toso, et al. 2014). The activation of STAT3 has been found to prevent anti-tumour responses by blocking the activation of DCs, NK cells and T cells, thereby promoting tumour growth. Kortylewski et al. have demonstrated that blocking STAT3 signalling in haematopoietic cells could markedly activate and maintain the function of DCs, T cells, NK cells and neutrophils, and suppress the tumour growth (Kortylewski, et al. 2005). Loss of functional PTEN pathway leads to persistent activation of PI3K/Akt signalling pathway which has been correlated with non T cell-inflamed tumours (Carnevalli, et al. 2018). In muscle-invasive bladder cancer, upregulation of PI3K signalling due to PIK3CA (phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha) mutations is found to be significantly correlated with the downregulation of 57 relevant immune genes, compared to unmutated-PIK3CA bearing tumours. Inhibiting PI3K signalling has been shown to promote T cell activation and trafficking within the tumour microenvironment of muscle-invasive bladder cancer (Borcoman, et al. 2019). In chronic myeloid leukaemia, activation of PI3K signalling is also found to inhibit the expression of NKG2D receptor on NK cells, thereby preventing NK cells activation within tumour microenvironment (Groh, et al. 2002, Boissel, et al. 2006). In melanoma, the PI3K signalling is found to promote the production of IL-10 following prolonged exposure to TGF-β through a crosstalk between Smad and PI3K/Akt signalling pathway. IL-10 as an immunosuppressive cytokine could inhibit the activation and function of NK cells and DCs by reducing MICA expression on melanoma cell surface (Serrano, et al. 2011). Loss of PTEN has been found to inhibit the maturation and activation of DCs and macrophages via blocking the expression of LC3I and LC3II genes that regulate the process of autophagy (Peng, et al. 2016). Autophagy facilitates cross-presentation of tumour-derived antigens by antigen-presenting cells, and thus activate innate and adaptive immune response (Zhong, Sanchez-Lopez and Karin 2016).

1.4.1.3. c-MYC signalling pathway

c-MYC signalling as a tumour cell-intrinsic mechanism has also been found to affect anti-tumour immunity within the tumour microenvironment. c-MYC is one of three related proto-oncogenes that normally regulate cell cycle, cell growth and survival (Lin, et al. 2012). The expression of c-MYC is found significantly upregulated in many types of cancer and promote tumour initiation and maintenance (Gabay, Li and Felsher 2014). In leukaemia/lymphoma mouse model, inactivation of c-MYC signalling contributes to significantly reduction in the size of tumour in wide-type animals (Jain, et al. 2002), whereas, minimal reduction in the size of tumour with tumour recurrence has been noticed in immunocompromised mice, confirming the cellular link between c-MYC signalling and immune cells within the tumour milieu (Felsher 2010, Casey, Bellovin and Felsher 2013). For instance, Rakhra et al. have showed that $CD4^+$ T cells, but not $CD8^+$ T cells, are required for sustained tumour regression upon MYC inactivation. The authors have demonstrated that $CD4^+$ T cells are found to be rapidly localised to the site of tumour after 4 days of MYC inactivation. The presence of $CD4^+$ T cells is also found to significantly inhibit angiogenesis upon MYC inactivation via the expression of TSP-1 (thrombospondin-1), a robust anti-angiogenic protein (Rakhra, et al. 2010). This study suggests that the activation of MYC by tumour cells may inhibit the activation and recruitment of innate immune cells to the site of tumour and prevents them activating adaptive immune cells, and therefore facilitate tumour progression (Rakhra, et al. 2010).

MYC has also been found to regulate the expression of CD47 and PD-L1 (Casey, et al. 2016). CD47 is found significantly expressed on the surface of tumour cells and acts as innate immune checkpoint which regulate immune response derived from innate immune cells such as macrophages and DCs (antigen-presenting cells (APCs)) (Jaiswal, et al. 2009). Upregulation CD47 by tumour cells inhibits the recruitment and activation of APCs and T cells into the site of tumour, thereby facilitating T cell exclusion within the tumour microenvironment through inhibition of tumour cells phagocytosis by APCs (McCracken, Cha and Weissman 2015). This finding is also supported by Kauder et al. The authors have demonstrated that blocking CD47 enhances tumour cells phagocytosis by macrophages and promotes the activation of DCs and adaptive immune cells, thereby boosting anti-tumour immunity (Kauder, et al. 2018). PD-L1 (programmed death-ligand 1) is also found to be significantly

expressed on the surface of tumour cells and mediates coinhibitory mechanism via binding to its receptor PD-1 expressed on T cells to inhibit the activation of T cells, leading to T cell exhaustion within the tumour microenvironment (Wang, et al. 2016). Casey et al. have proved that the expression of both CD47 and PD-L1 is correlated with the expression of MYC oncogene. Using MYC-induced T cell acute lymphoblastic leukaemia (MYC T-ALL) transgenic mouse model, the authors have showed that MYC inactivation could significantly reduce the expression of CD47 and PD-L1 on tumour cells at the mRNA and protein level *in vitro* and *in vivo*. The authors have also confirmed that MYC could induce the expression of both CD47 and PD-L1 by binding directly to their gene promoters (Casey, et al. 2016). These finding have also been supported by Kim et al. The authors have confirmed the positive correlation between the expression of MYC and PD-L1 in non-small cell lung carcinoma (NSCLC). NSCLC tumours expressing both MYC and PD-L1 are significantly correlated with worse prognosis and decreased disease-free and overall survival rates, compared to $MYC^{negative}$ $PD\text{-}L1^{negative}$ tumours (Kim, et al. 2017). Another mechanism by which MYC controls anti-tumour immunity has been demonstrated by Kortlever et al. In this study, using $Kras^{G12D}$-derive lung adenocarcinoma mouse model, MYC amplification could inhibit the infiltration of B, T and NK cells into the tumour microenvironment by inducing the production of IL-23 from tumour cells. The authors have also found that MYC amplification could increase the induction of CCL9 from adenoma epithelial compartment, which promotes stromal angiogenesis (Kortlever, et al. 2017).

1.4.1.4. Loss of antigenicity

Loss of antigenicity is also another tumour-cell intrinsic mechanism which is operated by tumour cells to avoid anti-tumour immune responses. Loss of antigenicity occurs when tumour cells fail to express immunogenic antigens that can be recognised by immune cells (Coulie, et al. 2014). Loss of antigenicity is either due to the selection of less immunogenic variants of tumour cell, down-regulation of Class I MHC molecule or defects in antigen processing machinery within tumour cells, which hinders cross-presentation of antigens by tumour cells. Loss of MHC class I expression is found to be one of main mechanisms by which malignant cells evade adaptive immune response (Schreiber, et al. 2011). Even if a tumour produces enough immunogenic antigens, immune recognition is based on the ability of

tumour cells to present antigen in the context of a peptide-MHC complex (Beatty and Gladney 2015). Loss of MHC class I expression is found in about 18-60% of common solid cancers including breast, head and neck squamous cell carcinoma, lung, colon, renal cell carcinoma, cervical cancer, bladder cancer, prostate and melanoma (Campoli and Ferrone 2008). Loss of MHC class I expression can be due to genetic alterations at the gene level or regulatory alteration at the transcriptional level. These alterations result in total loss of all class I molecules or selective downregulation of HLA class I alleles (Leone, et al. 2013). Mutations in the β2-microglobulin gene are found highly correlated with the total loss of MHC class I expression. In addition, aberrations in the components of antigen processing machinery which include low-molecular weight proteins (LMPs), transporter associated with antigen presentation (TAP) isoforms and tapasin (TPN) contribute to the loss of MHC class I expression on tumour cells (Thuring, et al. 2015). Hypermethylation of MHC gene promoters may also affect cross-presentation of MHC class I expression on tumour cells (Leone, et al. 2013).

1.4.2. Immune cell-mediated mechanisms

Tumour escape is not only mediated by tumour cells, immune cells infiltrating into the tumour microenvironment can facilitate tumour evasion and progression through establishing an immunosuppressive tumour microenvironment (Croci, et al. 2007). Several immune cells are found to play an essential role in the development of immunosuppressive microenvironment which promotes tumour progression, including MDSCs, T regulatory cells (Tregs), tumour-associated macrophages (TAMs), DCs, T helper 17 (Th17), mast cells and B regulatory cells (Kitamura, Qian and Pollard 2015).

1.4.2.1. Tumour-associated macrophages

Growing evidence has confirmed the role of tumour-associated macrophages (TAMs) in promoting the tumour microenvironment. TAMs are found to be the more abundant among the inflammatory cells within the tumour microenvironment (Vitale, et al. 2014). TAMs are generated from circulating monocytes that are recruited to the site of tumour via tumour-derived chemokine CCL2, or from tissue-resident macrophages (Sica, et al. 2007). Macrophages are conventionally classified into two main subsets including classical and alternative macrophages. “classical”

M1 macrophages are thought to boost anti-tumour response and regulate the functional activity of T helper 1 cells (Th1) by producing Th1-related cytokines such as IL-12 and TNF-α (Sica, et al. 2008). Whereas "alternative" macrophages, known as M2 macrophages, are involved in many mechanisms such as wound healing, tissue remodelling and angiogenesis, and secrete Th2-related cytokines such as TGF-β, IL-6, and IL-10 that regulate the function of T helper 2 (Th2) cells (Solinas, et al. 2009). Many studies have confirmed that TAMs are unstable in tumour microenvironment in that they can switch their phenotypic signature from a M1 to a M2 phenotype, thus providing a favourable milieu for tumour growth (Sica, et al. 2006). This polarisation of M2-TAMs is favoured by other immune cells including Tregs and dendritic cells (DCs), and also by tumour-related factors such as colony-stimulating factor (CSF-1) and hypoxia-inducible factor 1 (HIF 1-α) (Noy and Pollard 2014). Unlike classical macrophages, M2-TAMs are deficient in antigen-presenting capabilities and express PD-L1, and thus suppress anti-tumour immunity (Chanmee, et al. 2014). The immunosuppressive role of TAMs has been reported in many solid tumours including breast, hepatocellular and renal cell carcinomas, and thus several clinical trials have been designed for targeting TAMs for the treatment of cancer (Williams, Yeh and Soloff 2016, Shirabe, et al. 2012, Behnes, et al. 2014).

1.4.2.2. Tumour-infiltrating dendritic cells

Dendritic cells (DCs) also play an important role in mediating anti-tumour responses due to their functional ability as 'professional' antigen-presenting cells which have the capacity to activate adaptive immune cells ($CD4^+$ and $CD8^+$ T cells) (Moser 2003). However, a robust activation of anti-tumour responses requires only mature DCs which can cross-present antigens effectively. The maturation of DCs is regulated by several factors depending on the local milieu (Munz, Steinman and Fujii 2005). For instance, the majority of infiltrating DCs in the tumour microenvironment are immature or "paralyzed" DCs that have become tolerogenic due to many factors such as IL-10 and vascular endothelial growth factor (VEGF) secreted by TAMs (Suciu-Foca, Berloco and Cortesini 2009). Immature DCs can also secrete IL-10 and TGF-β and induce the expression of IDO, and thereby can promote the expansion of Tregs. Thus, immature DCs elicit an immunosuppressive role which can suppress the activation of effector T cells, induce T-cell anergy, and then promote tumour progression (Benencia, Muccioli and Alnaeeli 2014). The

immunosuppressive function of immature DCs can also be mediated by Tregs through cell-cell contact mechanisms involving cytotoxic T-lymphocyte antigen-4 (CTLA-4) expressed on Tregs, and CD80/CD86 expressed on DCs. This upregulates the expression of IDO in DCs, and thus promotes the immunosuppressive milieu (Hubert, et al. 2007). Furthermore, Tregs are thought to influence the maturation of DCs via IL-10 and TGF-β signalling and affect the ability for DCs to stimulate effector T cells via inhibition of co-stimulatory molecules expressed on DCs (Tran Janco, et al. 2015).

1.4.2.3. Myeloid-derived suppressor cells (MDSCs)

MDSCs are a heterogenous subpopulation derived from myeloid progenitor and immature myeloid cells. MDSCs are highly immunosuppressive cells which act as negative regulators of immune response in various pathological conditions such as chronic inflammation and cancer (Gabrilovich and Nagaraj 2009). MDSCs can be grouped into two main subsets; monocytic-MDSCs (M-MDSCs) and polymorphonuclear-MDSCs (PMN-MDSCs). MDSCs are found highly enriched in the tumour microenvironment. Many studies have studied molecular features by which MDSCs are recruited and regulated in the tumour bed and found that several factors could mediate the expansion and activation of MDSCs in the tumour milieu. These factors include granulocytes-macrophage colony-stimulating factor (GM-CSF), G-CSF, IL-6, prostaglandin-E2 (PGE2), VEGF and cyclooxygenase-2 (COX-2) (Obermajer, et al. 2011, Sinha, et al. 2007, Ostrand-Rosenberg, et al. 2012). Two main transcription factors are also found to play a significant role in the development and function of MDSCs including STAT transcription factors and C/EBP-β transcription factor, respectively (Qu, Wang and Lin 2016). The secretion of GM-CSF and G-CSF could regulate the development of MDSCs through downregulation of interferon-regulatory factor 8 (IRF-8) via STAT3 and STAT5-dependent pathways, confirming the role of IRF-8 as a negative regulator of MDSCs (Waight, et al. 2013). In addition, C/EBP-β transcription factor is also found to regulate the suppressive activity of tumour-induced MDSCs. Blocking C/EBP-β could significantly inhibit the activity of MDSCs, affect the differentiation of immature myeloid precursors and enhance the activity of $CD8^+$ effector T cells (Marigo, et al. 2010). The immunosuppressive activity of MDSCs is mediated by several factors including inducible nitric oxide synthase (iNOS), Cox-2, PGE-2, TGF-β and arginase-

1 (Arg-1) (Li, et al. 2009, Bogdan 2010, Yang, et al. 2008, Parker, Beury and Ostrand-Rosenberg 2015). In general, PMN-MDSCs are found more enriched in the tumour microenvironment than M-MDSCs, whereas the latter can elicit stronger immunosuppressive activity (Marvel and Gabrilovich 2015). The suppressive activity of M-MDSCs is linked with the metabolism of L-arginine which is used as a common substrate by Arg-1 and iNOS to produce urea or nitric oxide to inhibit proliferation and differentiation of T cells (Bogdan 2010). M-MDSCs are more plastic cells and found to have the potential to differentiate into tumour-associated macrophages (TAMs) within the tumour microenvironment. CCL2-CCR2 signalling pathway is found to mediate the recruitment of M-MDSCs into the tumour site where they can become TAMs (Tcyganov, et al. 2018). Unlike M-MDSCs, PMN-MDSCs can subdue anti-tumour immunity in an antigen-specific manner. These cells can take up and present antigens to antigen-specific T cells and then induce nitration of T cell receptors which makes T cells unresponsive to antigen stimulation (Dilek, et al. 2012).

Reactive oxygen species (ROS) plays an essential role in PMN-MDSCs mediated immune suppression. The production of ROS by PMN-MDSCs is found tightly regulated by NADPH oxidase (NOX2). Therefore, inhibition of NOX2 could suppress PMN-MDSCs mediated immune suppression, confirming the pivotal role of ROS in the functional activity of PMN-MDSCs (Corzo, et al. 2009). The presence of both ROS and nitric oxide (NO) could induce the production of peroxynitrite which mediates the nitration of TCR and thus blocks $CD8^+$ T cell response (Feng, et al. 2018).

MDSCs are also found to promote tumour progression by inducing Tregs expansion within the tumour microenvironment. The induction of Tregs by MDSCs is mediated via secretion of IL-10 and TGF-β (Nakamura and Ushigome 2018). The production of IFN-γ by activated T cells induces expression of immunoregulatory molecules PD-L1 and CD40 on MDSCs (Lu, et al. 2016, Pan, et al. 2010). Upregulation of CD40 by MDSCs is found to mediate the induction of Tregs within the tumour microenvironment (Pan, et al. 2010).

In addition, MDSCs are found to be implicated in the process of EMT (epithelial-to-mesenchymal transition) which mediates tumour invasiveness and metastasis. The production of NO and TGF-β by MDSCs could induce upregulation of COX-2 in

tumour cells and then activate β-catenin/TCF-4 pathway (Li, et al. 2015). This pathway is found to play a critical role in mediating the process of EMT (Chang, et al. 2015). Moreover, MDSCs play a vital role in regulating the process of angiogenesis which facilitates tumour progression. MDSCs express matrix metalloproteinase-9 (MMP-9) which can release VEGF stored in extracellular matrix, thereby promoting angiogenesis (Lee, et al. 2018). In addition, MDSCs produce VEGF and fibroblast growth factor (FGF) via STAT3-dependent mechanism which boosts angiogenesis within the tumour milieu (Kujawski, et al. 2008).

1.4.2.4. T regulatory cells (Tregs)

Regulatory T cells (Tregs) are a distinct subset of CD4+ T cells that play a vital role in maintaining immune homeostasis and peripheral tolerance therby preventing autoimmunity (Fehervari and Sakaguchi 2004). Tregs are characterised by the expression of the transcription factor forkhead-box protein P3 (Foxp3) that is fundamental for their development and suppressive activity. In human, mutation of the foxp3 gene leads to develop a severe X-linked organ-specific autoimmune disease known as immune-dysregulation polyendocrinopathy enteropathy X-linked syndrome (IPEX) (Gambineri, Torgerson and Ochs 2003). A foxp3-mutant mouse strain called the Scurfy mouse is also found to show multi-organ autoimmunity (Le Bras and Geha 2006). Tregs are generally categorised into two main subsets according to their ontogeny including naturally thymus-derived Tregs (nTregs) and peripherally-derived or induced Tregs (p or iTregs) (de Lafaille, Maria A Curotto and Lafaille 2009). Both of these subsets of Tregs are phenotypically defined by the expression of $CD4^{+}CD25^{++}Foxp3^{+}CD127^{-}$ and based on the methylation status of their Treg-specific demethylated region (TSDR) within the Foxp3 locus. The TSDR is found demethylated in nTregs, whereas it is methylated in i/pTregs (Floess, et al. 2007).

nTregs are developed in the thymic medulla from immature CD4+ T cells in response to high-affinity interaction with a self-peptide: major histocompatibility (MHC) complex expressed on medullary thymic epithelial cells (mTECs) and dendritic cells (DCs) (Bettini and Vignali 2010). Whereas, peripheral or induced Tregs are developed extra-thymically from mature naïve CD4+ T cells circulating in the

periphery in response to sub-optimal TCR signalling in the presence of several cytokines such as IL-2 and TGF-β (Lee, Bautista and Hsieh 2011). iTregs are also found to be generated at sites of inflammation or in a tolerogenic environment and have been characterised to have a various recognition of antigens (Workman, et al. 2009).

In cancer, Tregs are found extremely enriched in the tumour microenvironment and mediate several mechanisms that contribute to the inhibition of anti-tumour response and tumour progression (Facciabene, Motz and Coukos 2012). Several mechanisms are identified to mediate the recruitment of Tregs to the site of tumour. The production of CC-chemokines ligand 22 (CCL22) by tumour cells and macrophages can attract Tregs to the tumour bed via CC-chemokine receptor 4 (CCR4) expressed on the surface of Tregs (Gobert, et al. 2009). Upregulation of CCL17 by tumour cells is also found to regulate recruitment of Tregs via CCR4 (Mizukami, et al. 2008). Moreover, the overexpression of CCL28 by hypoxic tumour cells is found to recruit Tregs to the tumour milieu via binding to the cognate receptor CCR10 expressed on Tregs. Hypoxia-inducible factor 1α (HIF-1α) secreted by tumour cells is found to mediate Tregs recruitment by regulating upregulation of CCL28 in tumour cells (Ren, et al. 2016). CCL5 is also found to mediate Tregs recruitment to the tumour microenvironment via its corresponding receptor CCR5 which is detected on the surface of Tregs. CCL5 can be expressed by macrophages, T lymphocytes, platelets, tubular epithelium and tumour cells (Aldinucci and Colombatti 2014). Tregs are also found to be recruited to the tumour microenvironment through CXCR3 (C-X-C motif chemokine receptor 3) expression which is a cognate receptor for CXCL9, CXCL10 and CXCL11. For instance, the expression of CXCL10 by tumour cells could recruit Tregs to the site of tumour in hepatocellular carcinoma (Li, et al. 2016).

Upon recruitment to the tumour microenvironment, Tregs exert their immunosuppressive activities via various mechanisms including: (1) secretion of soluble immunosuppressive cytokines, (2) modulation of dendritic cells, (3) metabolic disruption and (4) suppression by direct cytolysis (Vignali, Collison and Workman 2008). Tregs are found to secret TGF-β cytokine which acts as a negative regulator for $CD8^+$ T cells by affecting their proliferation and differentiation (Strauss,

et al. 2007). In addition, Tregs can also produce IL-10 which is an immunoregulatory cytokine and plays a pivotal role in the regulation of adaptive immunity by suppressing prolonged activation of $CD8^+$ T cells (Ng, et al. 2013). Tregs can also inhibit anti-tumour response via production and secretion of IL-35 cytokine which is found to be required for Tregs to maintain their suppressive activity (Wei, et al. 2017). IL-35 has showed the ability to arrest the proliferation of T helper 1 and 17 cells at the G1 phase of cell division (Wirtz, et al. 2011). Another mechanism by which Tregs suppress anti-tumour immunity is via direct cytolysis of T effector cells. Although granzyme-mediated cytolysis has been widely considered the forte of NK cells and $CD8^+$ T effector cells (Lieberman 2003), activated human Tregs are shown to express granzyme A and perforin which directly kill the target cells via the adhesion of CD18 (Grossman, et al. 2004). Furthermore, not only human Tregs, mouse Tregs are also found to mediate their functional suppression via upregulation of granzyme B expression (McHugh, et al. 2002, Herman, et al. 2004). Tregs derived from granzyme B-deficient mouse are found to be less suppressive *in vitro*, confirming the important role of granzyme B in the suppressive potential of Tregs. Tregs also shown to mediate cytolysis in a granzyme B-dependent and perforin-dependent and independent manners (Herman, et al. 2004, Cao, et al. 2007). Tregs can also mediate their cytolytic function to kill $CD8^+$ T effector cells by inducing apoptosis via the TRAIL pathway, especifically TRAIL-death receptor 5 (TRAIL-DR5) (Ren, et al. 2007).

Treg can also induce apoptosis on effector T cells through the expression of galetin-1 which is highly upregulated in human and mouse Tregs. Galectin-1-defecient Tregs have shown a significant reduction in their suppressive activity in vitro, confirming the importance of galectin-1 signalling in mediating regulatory activity of Tregs (Garin, et al. 2007). Tregs can also attenuate proliferation of $CD8^+$ T effector cells metabolically. It has been suggested that Tregs can inhibit proliferation of effector T cells by depleting local IL-2 via the expression of CD25. Depleting IL-2 induces cellular starvation and therefore apoptosis (Pandiyan, et al. 2007). However, IL-2 depletion alone is found not required for Tregs to inhibit effector T cells (Oberle, et al. 2007). Coordinated expression of CD39 and CD73 is shown to maintain suppressive activity of Tregs. These ectoenzymes allow Tregs to hydrolyse ATP (adenosine triphosphate) to adenosine which suppresses functional activity of $CD8^+$

T effector cells (Deaglio, et al. 2007). Exogenous ATP is hydrolysed to ADP (adenosine diphosphate) and 5'AMP (adenosine monophosphate), while CD73 hydrolyses 5'AMP to adenosine (Bynoe and Viret 2008). Adenosine is shown to suppress T effector cells through stimulation A2A (adenosine 2A) receptor. Upon stimulation, A2A receptor signalling could inhibit production of proinflammatory cytokine IL-6 and enhance production of TGF-β which in turn promote the induction of induced Tregs (Zarek, et al. 2008).

Tregs are also shown to inhibit T effector cells indirectly by modulating the maturation and function of DCs which act as robust activators of effector T cells (Whiteside 2012). It has been reported that Tregs can control the maturation and function of DCs through interaction between the inhibitory receptor "cytotoxic T-lymphocyte antigen-4 (CTLA-4) and the co-stimulatory receptors "CD80/CD86" which is constitutively expressed by Tregs and CDs, respectively. This interaction could inhibit the expression of co-stimulatory receptors of DCs, thereby reducing their ability to activate T effector cells (Tadokoro, et al. 2006, Read, Malmstrom and Powrie 2000, Oderup, et al. 2006). Moreover, the interaction between CTLA-4 and CD80/CD86 could induce DCs to produce IDO (indoleamine 2,3-dioxygenase) which is a potent regulatory molecule that inhibit proliferation of T cells (Fallarino, et al. 2003).

IDO is one of three different heme-containing enzymes that catalyse the first rate-limiting step of tryptophan breakdown into immune suppressive kynurenines (Prendergast 2008). IDO can activate the GCN2 stress-response kinase, which inhibits the proliferation of T cells and mediates the differentiation of naïve CD4+ T cells into Tregs cells. Also, IDO produces soluble factors, such as kynurenines, which bind the aryl hydrocarbon receptor (AhR) to induce Treg differentiation (Nguyen, et al. 2014).

Tregs can also affect the function of DCs through lymphocyte-activation gene-3 (LAG-3) which is found expressed on Tregs, activated T effector cells and NK cells (Liang, et al. 2008). LAG-3 is homologous to CD4 and has high affinity to bind MHC class II complexes. LAG-3 is required to maintain immunosuppressive activity of Tregs. Upon binding to MHC class II expressed on immature DCs, LAG-3 initiates an immunoreceptor tyrosine-based activation motif (ITAM) which mediates an

inhibitory pathway to constrain DC maturation and their immunostimulatory ability (Huang, et al. 2004, Liang, et al. 2008). Consistence with this, it has been found that human MHC class II-expressing Tregs are more suppressive than MHC class II-negative Tregs, supporting the notion that MHC class II-positive Tregs might inhibit activated T cells by ligating LAG-3 expressed on them (Baecher-Allan, Wolf and Hafler 2006). It has also been reported that Tregs could modulate the function of DCs via neuropilin-1 which is found differentially expressed on Tregs (Sarris, et al. 2008).

1.4.2.5. T helper 17 (Th17) cells

Th17 cells are highly plastic cells differentiated independently from a $CD4^+$ T cell lineage distinct from Th1, Th2 and Tregs (Harrington, et al. 2005). Development of Th17 from naïve $CD4^+$ T cells is mediated by several cytokines including IL-1β, IL-6 and TGF-β. In addition to IL-21 and IL-23 cytokines that maintain the developed Th17 cells for long-term (Stritesky, Yeh and Kaplan 2008). Th17 cells are identified by their ability to secrete IL-17A, IL-17F, IL-21, IL-22 and CCL20 (Liang, et al. 2006, Chang and Dong 2007, Dong 2008). Th17 cells have shown the capability to transdifferentiate mainly into Th1 or Tregs, and occasionally into Th2, T follicular helper (TFH) cells and type 1 regulatory (Tr1) cells based on the presence of cytokines and chemokines secreted in the local microenvironment (Guéry and Hugues 2015). For instance, the enrichment of proinflammatory cytokines including IL-1β, IL-23, IL-6 and IL-12 could mediate the conversion of Th17 into Th1 which promotes activation of $CD8^+$ T cells in the local milieu, whereas Th17 cells could be converted into Tregs in the presence of TGF-β, but in the absence of IL-6 (Bailey, et al. 2014).

In cancer, Th17 are found to play a paradoxical role in tumour immunity as they can boost tumour progression by acquiring and initiating immunosuppressive activities and induce anti-tumour immune responses by secreting proinflammatory cytokines which can enhance the function of $CD8^+$ T cells and NK cells. Th17 cells are found to have no direct cytotoxic effect on tumour cells as they do not produce perforin and granzymes (Kryczek, et al. 2009, Yen, et al. 2009). Instead, Th17 cells can mediate anti-tumour responses indirectly by mediating recruitment of other immune cells into the tumour site. For instance, by secreting IL-17, Th17 cells can induce tumour

cells to upregulate CXCL9 and CXCL10 chemokines which in turn recruit effector T cells into the tumour microenvironment (Kryczek, et al. 2009). IL-17 can also stimulate macrophages to produce IL-12 which mediates the expension of $CD8^+$ cytotoxic T cells and boots their functional activities (Markiewicz, et al. 2009). However, Th17 cells can promote tumour progression by inhibiting anti-tumour responses. IL-17 secreted by TH17 is found to induce angiogenesis in tumour context by stimulating stromal cells (myeloid cells and fibroblasts) and epithelial cells to produce VEGF and angiogenin-2 (Chung, et al. 2013). IL-17 is found also to provoke tumour-related fibroblasts to generate G-CSF which triggers myeloid cells to generate MMP-9, prokineticin-2/Bv8 and proinflammatory S100A8/9 molecules to mediate angiogenesis (Chung, et al. 2013). The interaction between IL-17 and stromal cells-related cytokines is found to be correlated with the increased levels of TGF-β (Jeon, et al. 2007). It has been reported that Th17 cells can upregulate CD39 and CD73 on their surface in the presence of TGF-β, and therefore acquire the ability to hydrolyse ATP to adenosine which suppresses $CD8^+$ T cells proliferation (Bailey, et al. 2014), thereby exerting its regulatory activity. IL-17 is also found to stimulate upregulation of CCL17 and CCL22 by DCs which in turn recruits Tregs to the tumour site (Halim, et al. 2017). IL-17 mediates the migration of MDSCs into the tumour milieu. Injection of IL-17 has significantly increased in the number and suppressive activity of MDSCs within the tumour microenvironment in $IL\text{-}17R^{-/-}$ tumour -bearing mice (He, et al. 2010). IL-17 has shown to trigger tumour cells to activate the COX-2/PGE2 pathway which mediate the differentiation of M2 macrophages that promotes tumour progression (Li, et al. 2014). IL-17 has also been found to trigger hepatocellular tumour cells to upregulate expression of MMP-2 and MMP-9 via NF-kB pathway which promotes tumour invasiveness and metastasis (Li, et al. 2011).

1.4.2.6. Mast cells

Mast cells are generated in the bone marrow as precursor cells that migrate to peripheral tissues where they become fully mature. Mast cells belong to granulated innate immune cells and express various receptors such as Toll-like receptors (TLRs), Nod-like receptors (NLRs), Fc receptors and complement receptors. These receptors allow mast cells to mediate several immunological responses including pro-inflammatory and anti-inflammatory immune responses (Galli, Nakae and Tsai 2005, Gurish and Austen 2001). Mast cells have been found dramatically enriched

within tumour microenvironment and correlated with worse prognosis and reduced survival in many types of human cancer including B-cell chronic lymphocytic leukaemia (Molica, et al. 2003), Hodgkin lymphoma (Glimelius, et al. 2005), melanoma (Ribatti, et al. 2003), prostate (Nonomura, et al. 2007), pancreatic adenocarcinoma (Strouch, et al. 2010) and oesophageal-squamous cell carcinoma (Elpek, et al. 2001). It has been reported that tumour-derived stem cell factor (SCF) mediates migration and activation of mast cells into tumour bed by binding to CD117 (KIT receptor) which is highly expressed on mast cells. Activation SCF/Kit pathway in mast cells mediates tumour progression (Huang, et al. 2008). Mast cells are found to mediate angiogenesis by releasing several angiogenic molecules including TGF-β, VEGF, fibroblast growth factor-2 (FGF-2), angiopoietin-1, IL-8, tryptase and chymase (Ribatti and Crivellato 2012). Mast cells are also found to produce IL-10 which can inhibit anti-tumour immune response (Galli, Grimbaldeston and Tsai 2008). Recently, it has been reported that tumour-derived TNF-α stimulates tumour-infiltrating mast cells to upregulate expression of PD-L1 via activation NF-kB pathway, and contribute to immune suppression and tumour progression in gastric cancer (Lv, et al. 2019).

1.4.2.7. B regulatory cells (Bregs)

Bregs are a unique subset of B cells and found to play a pivotal role in regulating and balancing immune responses in inflammatory diseases, autoimmune diseases and cancer (Berthelot, et al. 2013). Several subpopulations of Bregs have been identified in human and mouse models including IL-10-producting Bregs, TNF-α-secreting Bregs, Fas-ligand (Fas-L^+) Bregs, granzyme-B^+ Bregs, TGF-β-producing Bregs, $GITR^+$ Bregs and $STAT^+$ Bregs (Zhang, Gallastegui and Rosenblatt 2015). In pancreatic cancer, it is found that increased levels of IL-18 produced by tumour cells mediate the induction and immunosuppression of Bregs *in vitro* and *in vivo*. IL-18-induced Bregs have shown to upregulate expression of PD-L1 and produce IL-10 which both maintain their immunosuppressive activity (Zhao, et al. 2017). In gastric cancer, it is found that IL-35-secreting Bregs are significantly enriched in the peripheral blood of patients with advanced gastric cancer. Also, the accumulation of IL-35-secreting Bregs is found to be positively correlated with high frequencies of Tregs and MDSCs, confirming the potential role of Bregs in tumour progression (Wang, Liu and Li 2018). Bregs are also found to significantly enriched in the peripheral blood of

patients with lung and oesophageal cancer, compared to healthy individuals (Zhou, et al. 2014, Shi, et al. 2014). In 4T1 breast cancer mouse model, Bregs are detected significantly enriched in axillary and inguinal lymph nodes, spleen and blood. The accumulation of Bregs is found to promote lung metastasis in tumour-bearing mice. Also, Bregs are shown to produce high levels of TGF-β and therefore mediate the induction of Tregs from naïve $CD4^{+}CD25^{-}$ T cells (Olkhanud, et al. 2011). It has been reported that IL-21 can mediate the induction of Bregs in the microenvironment of several human tumour types including prostate, colorectal, breast and cervical carcinoma. IL-21-drived Bregs are found to express high levels of granzyme-B, along with IL-10, CD25 and IDO which inhibit T cell proliferation and function. These granzyme-B positive Bregs have also been found existed adjacent to IL-21-secreting Tregs, suggesting that Tregs may mediate the induction of Bregs through IL-21 expression (Lindner, et al. 2013).

According to what discussed above, it seems that TGF-β is a multifunctional master cytokine which contributes for the promotion of immunosuppressive tumour microenvironment through molecular interactions with different immunosuppressive T cells. The role of TGF-β in the biology of T cells and cancer will be discussed in the next section.

1.5. TGF-β cytokine

Transforming growth factor β (TGF-β) is a multifunctional cytokine that belongs to a family consisting of three distinct isoforms including TGF-β1, TGF-β2 and TGF-β3, while TGF-β1 is the prevalent isoform expressed in the immune system (Chin, et al. 2004). TGF-β is a pleiotropic cytokine that regulates a variety of essential cellular events including cell development, proliferation, differentiation, homeostasis and apoptosis. TGF-β is initially translated as a precursor form "pre-pro-TGF-β", which consists of pre-region signal peptide, pro-region latency-associated peptide (LAP) (N terminal peptide) and mature TGF-β peptide (C terminal peptide). The pre-pro-TGF-β resides within the endoplasmic reticulum where it undergoes post-translational modifications that result in the removal of signal peptide from LAP and the cleavage of the mature TGF-β peptide from LAP by a furin convertase (Annes, Munger and Rifkin 2003). To preserve TGF-β activity, the homodimeric LAP binds

and wraps the homodimeric mature TGF-β by noncovalent bonds, forming a soluble small latent complex (SLC) (Lawrence 2001) (Figure 1.5).

The SLC is then released from the cell to the extracellular matrix (ECM) where it can associate with milieu molecules such as LTBPs (latent TGF-β binding proteins) and LRRC32 (leucin rich repeat containing 32, which is also known as glycoprotein A repetitions predominant (GARP)). LTBP attaches the SLC by covalent bonds to form a large latent complex (LLC). LTBP facilitates the deposition of the LLC to the ECM (figure 1.6). The SLC can also bind the LRRC32/GARP to form a membrane-latent TGF-β (mLTGF-β). The LRRC32/GARP acts as a cognate receptor for the SLC (Unsold, et al. 2001, Wang, et al. 2012) . Therefore, TGF-β can be identified in three different forms; soluble SLC (small form), LLC (large form) and mLTGF-β (membrane form). These three forms represent the latent "inactive" state of TGF-β which is unable to bind its receptors and, therefore, requires activation to free the TGF-β as an active form to initiate TGF-β downstream signalling pathways (Tran 2011). Activation of TGF-β is mediated by several mechanisms that involve acidification, matrix metalloproteases, thrombospondin-1, plasmin, αv-integrins and proteases (Henderson, et al. 2013, Murphy-Ullrich and Poczatek 2000, Jenkins 2008, Hyytiäinen, Penttinen and Keski-Oja 2004). Once being exposed to these activator, the latent form of TGF-β undergoes proteolytic cleavage and structural modifications that result in conformational changes to the LAP structure and liberation of the bioactive TGF-β (Kubiczkova, et al. 2012).

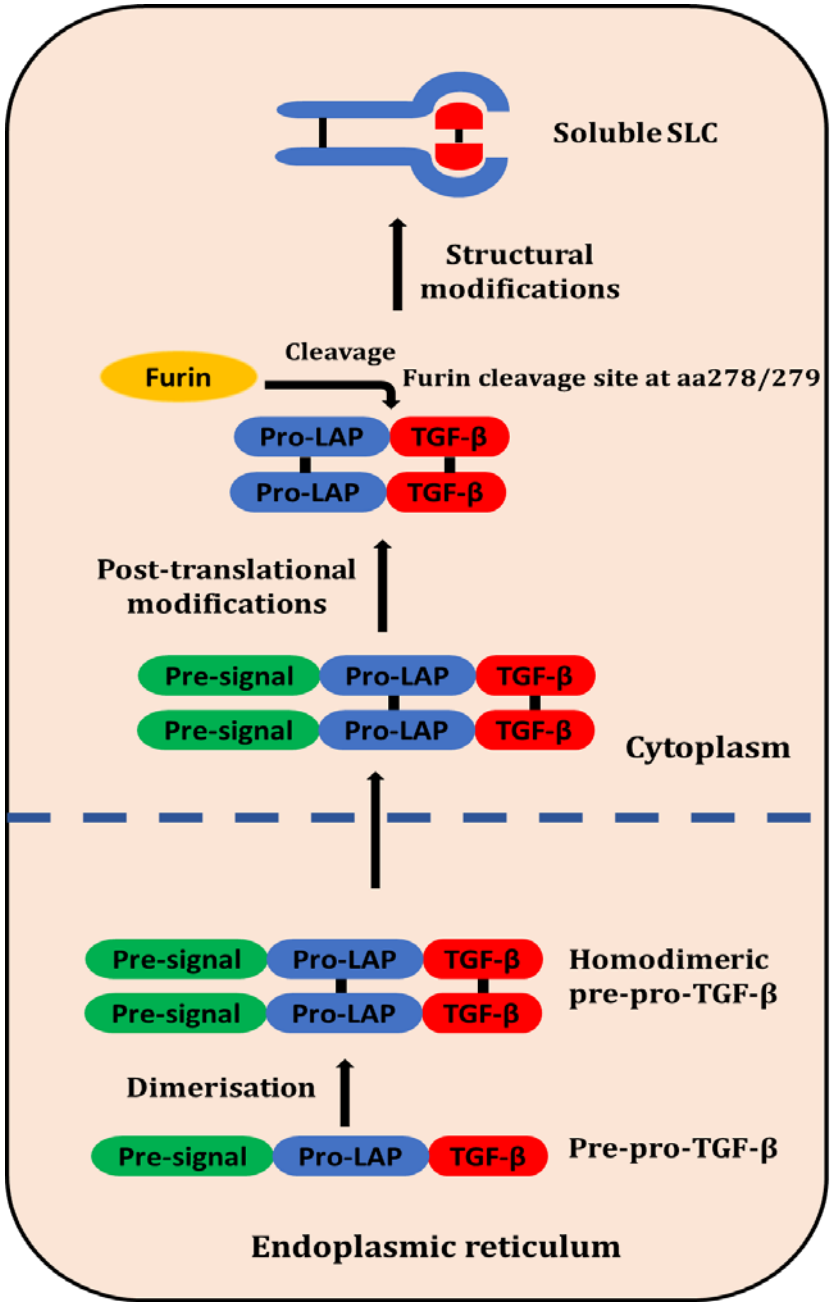


Figure 1.5. Mechanism of production of TGF-β. Initially, TGF-β is translated as a precursor form "pre-pro-TGF-β" which consists pre-signal peptide, pro-region latency associated peptide (LAP) and mature TGF-β. After dimerization, homodimeric pre-pro-TGF-β is released to the endoplasmic reticulum where it undergoes post-translational modifications that leads to the removal of pre-signal peptide. Then TGF-β is cleaved from pro-LAP peptide by Furin at cleavage site. Then, pro-LAP attaches to the free TGF-β to prevent its activity in the cytoplasm. Attachment pro-LAP to TGF-β initiates structural modifications that results in the generation of soluble small latent complex (SLC) which is inactive form of TGF-β.

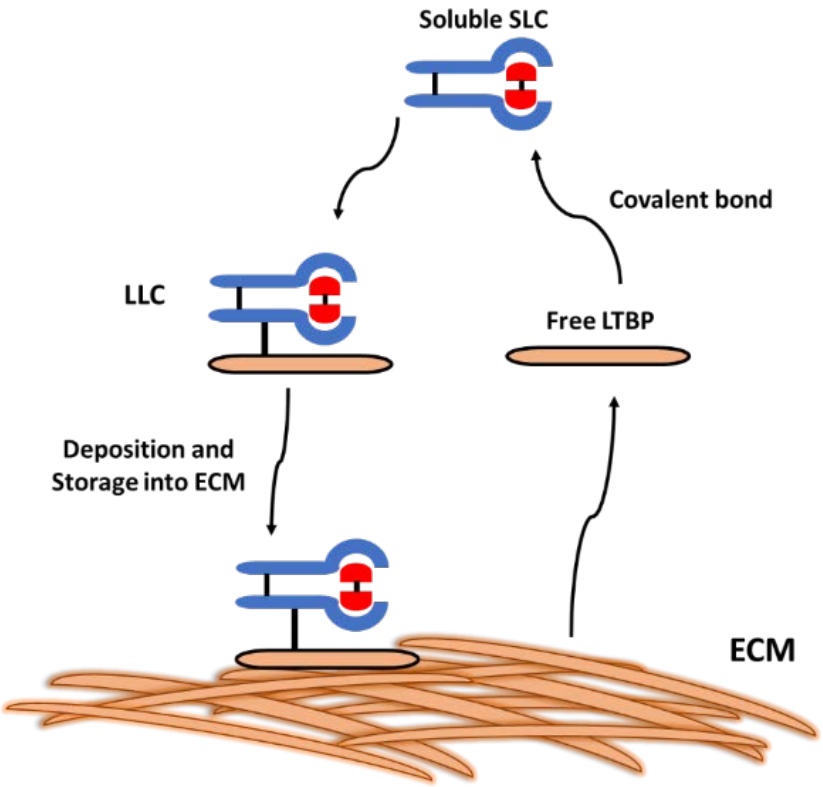


Figure 1.6. Mechanism of formation of large latent complex (LLC) from SLC. TGF-β is released from the cell as an inactive form "soluble SLC". Free soluble SLC binds to latent TGF-β binding protein (LTBP) which is released from the extracellular matrix (ECM), forming large latent complex (LLC). LLC is deposited and stored in the ECM as inactive form of TGF-β.

The bioactive form of TGF-β initiates its biological activities by ligation to TGF-β type I (TGF-βRI) and type II (TGF-βRII) receptors which are serine/threonine kinases. The bioactive TGF-β has a high affinity binding to TGF-βRII (Shi and Massagué 2003). Upon binding, TGF-βRII undergoes autophosphorylation and triggers TGF-βRI, forming a tetrameric receptor complex which comprises of two TGF-βRII and two TGF-βRI molecules (figure 1.7). This activated tetrameric complex induces downstream signalling pathways. Two key mechanisms have been identified to mediate TGF-β-downstream signalling, including Smad-dependent and Smad-independent signalling pathways (Derynck and Zhang 2003). Smad-dependent pathway is triggered through phosphorylation and activation of Smad2/3 complex

by the activated TGF-β ligand-receptor tetrameric complex. Smad-2 and Smad-3 are receptor-regulated effector proteins which are also known as “R-Smads” (Shi and Massagué 2003). Transphosphorylation of TGF-βRI mediates its binding to other specialised scaffold proteins called SARA (Smad anchor for receptor activation) and Hrs (hepatic growth factor-regulated tyrosine kinase substrate). This binding induces phosphorylation of R-Smads at a C-terminal SSXS motif leading to activation of R-Smads (Chen 2009). The activated R-Smads can associate with Smad-4 (Co-Smad) in the cytoplasm forming a heterodimeric complex which is then transported to the nucleus where the complex attaches to the regulatory sequences of target genes in corporation with other transcriptional factors to regulate the expression of hundred genes (Massague, Seoane and Wotton 2005) (figure 1.7). The bioactive TGF-β can also induce several cell type-specific Smad-independent signalling pathways which are mediated by PI3K-Akt, mitogen-activated proteins kinase (MAKP), Rh0-like GTPases, protein phosphatase 2A (PP2A) and partition-defective 6 (Par6) signalling (Zhang 2017).

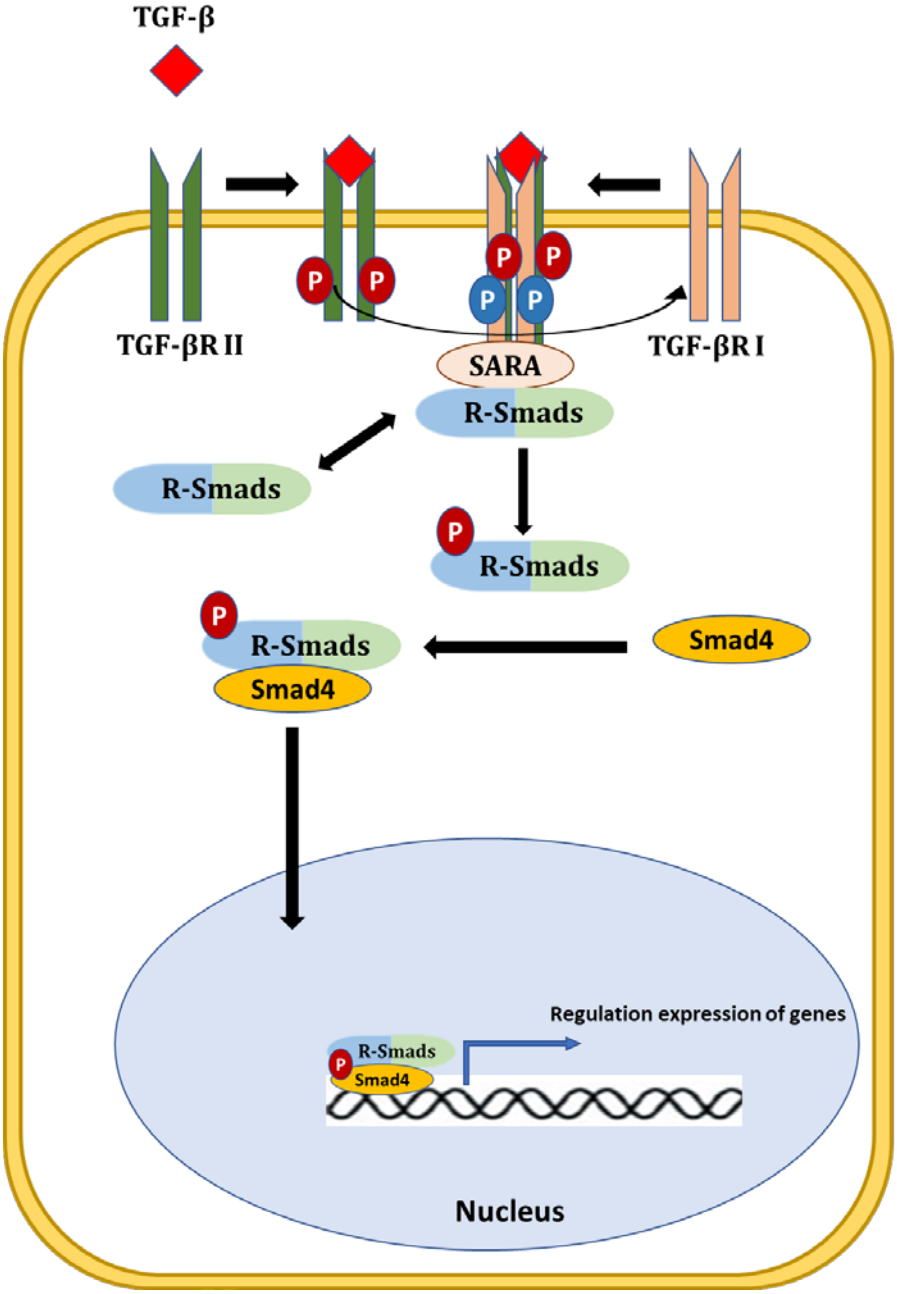


Figure 1.7. Schematic representation of Smad-dependent TGF-β pathway. Upon activation, free active TGF-β binds its receptor TGF-β type II (TGF-βRII). Upon binding, TGF-βRII undergoes autophosphorylation and stimulates TGF-βRI to form a tetrameric receptor complex consisting of two TGF-βRII and two TGF-βRI receptors. The activated tetrameric complex initiates downstream signalling cascade by mediating phosphorylation and activation of Smad2/3 complex (R-Smads) via SARA molecule which is a scaffold acting protein and functions to facilitate the phosphorylation process of R-Smads complex. The activated R-Smads complex phosphorylates and activates Smad4 to form a heterodimeric molecule which is directed to the nucleus where the heterodimeric binds to the regulatory sequences of targets genes.

1.5.1. The role of TGF-β in T cell biology

TGF-β plays a fundamental role in T cell development, differentiation, homoeostasis and tolerance. TGF-β has a paradoxical regulatory function that mediates both suppressive and inflammatory immune responses based on the status of the local microenvironment. This context-dependent function has enabled TGF-β to be the master of all T cell decisions (Li, et al. 2006).

1.5.1.1. TGF-β in the thymus

In the thymus, T precursor cells are subjected to a series of molecular and phenotypic modifications before they differentiate into mature T cells. The most critical stage in T cell differentiation is when immature $CD4^+CD8^+$ (double positive) T cells undergo thymic selection and lineage commitment. At this stage, TGF-β is found to play an important role in the differentiation of conventional $CD8^+$ T cells that requires TCR engagement and expression of IL-7 receptor that are the main requirements for $CD8^+$ T-cell lineage commitment (Park, et al. 2010). TGF-β is found to regulate the expression of IL-7 receptor α-chain (IL-7Rα) on the surface of developing $CD8^+$ T cells by repressing the expression of Gfi-1 (growth factor independent 1), a well-known transcription repressor, that inhibits the expression of *Il7Rα* in $CD8^+$ T cells (Ouyang, et al. 2013) (figure 1.8A).

TGF-β also controls the development of regulatory and innate-like T cells during thymic selection, including thymus-derived nTregs, invariant NK T (iNKT) cells and $CD8\alpha\alpha^+$ $TCR\alpha\beta^+$ intraepithelial lymphocytes (IELs) (figure 1.8A). The development of these specific subsets is based on their interactions with MHC-presenting self-ligands (agonist ligands) in a high-affinity manner. TGF-β is found to mediate and promote these interactions (Stritesky, Jameson and Hogquist 2012). The development of intrathymic nTregs is based on a combination of rigorous TCR interactions, co-stimulation and cytokines signals. Early studies have reported that TGF-β signalling is not required for nTregs development in mice with T cell-specific *Tgfbr2* (TGF-β receptor II) deletion as these mice have been shown to possess normal populations of $Foxp3^+$ thymocytes (Li, Sanjabi and Flavell 2006, Marie, Liggitt and Rudensky 2006). However, Liu et al. (2008) have confirmed the contribution of TGF-β signalling to the early development of intrathymic nTregs based on the finding that 3 to 5 day-old mice with T cell-specific *Tgfbr1* (TGF-β

receptor I) deletion have shown a significant reduction in the frequency of Foxp3$^+$ thymocytes (Liu, et al. 2008). Tone et al. have identified a conserved Smad3-binding sequence in the Foxp3 enhancer which maintains Foxp3 expression in intrathymic nTregs. Smad3 signalling is found to be required for activating the enhancer and thereby inducing Foxp3 expression, confirming the role of TGF-β in enhancer-related activation of Foxp3 expression in intrathymic nTregs (Tone, et al. 2008). However, Schlenner et al. have confirmed that the development of intrathymic nTregs is not regulated by TGF-β based on the finding that conserved Smad3-binding sequence-deficient mice have shown normal intrathymic Tregs development identical to wild-type mice (Schlenner, et al. 2012). Zheng et al. have also supported this finding and demonstrated that conserved noncoding sequence-1 (CNS1), which includes the conserved Smad3-binding sequence, at the Foxp3 locus is dispensable for nTregs development within the thymus (Zheng, Josefowicz, Chaudhry, Peng, Forbush and Rudensky 2010a). However, although TGF-β signalling is dispensable for nTregs development and lineage commitment, it is found to be required for maintaining the nTregs survival and viability within the thymus. Intrathymic nTregs have been found to express high levels of proapoptotic molecules and undergo negative selection-induced apoptosis in TGF-βRII-deficient mice compared to wild-type mice, confirming the role of TGF-β for promoting T cell survival by antagonising thymic negative selection (Ouyang, et al. 2010).

iNKT cells are characterised by their ability to induce both innate and adaptive immune responses, and recognise lipids presented by the MHC Class I-like molecule CD1d. These cells develop from CD4$^+$CD8$^+$ thymocytes and, like intrathymic nTregs, are induced by high-affinity interactions with agonist ligands (figure 1.8A) (Brennan, Brigl and Brenner 2013). The finding that TGF-βRII-deficient mice have failed to develop thymic and peripheral iNKT cells has indicated the requirement of TGF-β signalling in iNKT cells development (Li, et al. 2006, Marie, et al. 2006, Doisne, et al. 2009). TGF-β is found to maintain iNKT cells development by suppressing the apoptosis of their precursor cells (Doisne, et al. 2009).

CD8αα$^+$ TCRαβ$^+$ IELs are innate-like T cells that have an essential role in intestinal homeostasis (Cheroutre, Lambolez and Mucida 2011) (figure 1.8A). These cells are thought to be developed and differentiated within and outside the thymus and

induced by high-affinity TCR interactions with their agonist ligands (Pobezinsky, et al. 2012). The importance of TGF-β signalling in the development of these cells has confirmed by the finding that deletion of TGF-βRI has led to a significant reduction in the frequency of CD8αα$^+$ TCRαβ$^+$ IELs population and their thymic precursors (Konkel, et al. 2011).

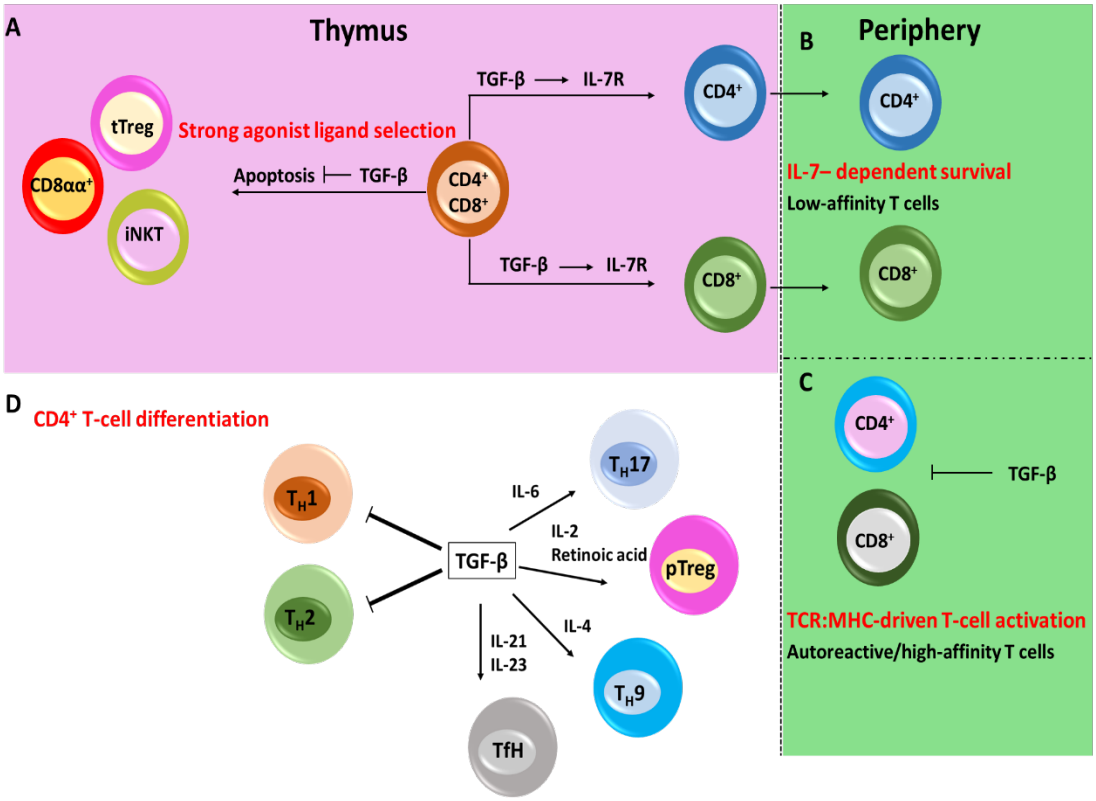


Figure 1.8. Regulation of T-cell biology by TGF-β. **(A)** In the thymus, TGF-β regulates the expression of IL-7R (IL-7 receptor) that is essential for the CD8$^+$ T cells lineage commitment. TGF-β also maintain the survival of tTregs, CD8αα$^+$ and iNKT cells which reacts with MHC-mounted self-ligand s (agonist ligands) with high-affinity by blocking apoptotic signals directed against tTregs, CD8αα$^+$ and iNKT cells. **(B)** In the periphery, TGF-β maintains homeostasis of the low-affinity CD4$^+$ and CD8$^+$ T cells by maintaining the expression of IL-7, whereas **(C)** TGF-β inhibits the proliferation of CD4$^+$/CD8$^+$ T cells that react with high-affinity to self-molecules. **(D)** TGF-β blocks the differentiation of Th1/Th2 cells, while it promotes the differentiation of other subpopulations of CD4$^+$ T cells in a cytokine-dependent context. In the presence of IL-2 and retinoic acid, TGF-β mediates the development of peripheral (induced) Tregs from naïve CD4$^+$ T cells. TGF-β also mediates the development of Th17 and Th9 in the presence of IL-6 and IL-4, respectively. The differentiation of TFH (T follicular helper) cells is also mediated by TGF-β in the presence of both IL-21 and IL-23 cytokines. This figure was adapted from (Soyoung and Ming 2013).

1.5.1.2. TGF-β in peripheral homeostasis

A competent immune system must produce and preserve a diverse naïve T cell pool within the confines of a relatively constant number of peripheral T cells. TGF-β is found to play a pivotal role in the maintenance of an efficient naïve T cell population by regulating T cell homeostasis, proliferation and repertoire diversity (McKarns and Schwartz 2005). During T-cell priming, TGF-β can suppress T cell proliferation by blocking transcription of the *Il-2* gene and thereby suppressing IL-2 production via Smad3-mdiated signalling (McKarns, Schwartz and Kaminski 2004). However, the capability of TGF-β in regulating T cell proliferation is based on the status of T cell proliferation and the presence of other co-stimulation signalling pathways involved in cell activation. For instance, TGF-β can suppress the proliferation of naïve T cells only, while it has no effect on activated T cells which reduce expression of TGF-βRII after activation (Cottrez and Groux 2001, Sanjabi, Mosaheb and Flavell 2009). In addition, CD28 co-stimulatory signalling is found to challenge and abrogate inhibitory effect of TGF-β on naïve T cells once being engaged with CD28. This confirms that TGF-β is unable to suppress the ability of activated APCs (antigen-presenting cells) to prime naïve T cells (Sung, Lin and Gorham 2003).

T cell proliferation must be properly controlled to maintain homeostasis under steady-state conditions and during immune challenges. Deletion or absence of TGF-β signalling affects homeostasis of both $CD4^+$ and $CD8^+$ T cells (Li, et al. 2006). For instance, TGF-βRII-deficient mice have shown a dramatic decrease in the peripheral T cell population. In addition to its role in $CD8^+$ T cell lineage commitment, TGF-β is found to mediate the homeostasis of peripheral $CD4^+$ T cells by maintaining the expression of IL-7Rα during their development in the thymus, which enables naïve peripheral $CD4^+$ T cells to recognise IL-7 for their survival in an IL-7-dependent manner (Ouyang, et al. 2013) (figure 1.8 B). Interestingly, TGF-β is found to specifically preserve the homeostasis of naïve peripheral $CD4^+$ T cells expressing low-affinity TCR (low-affinity $CD4^+$ T cells) (figure 1.8 B, C). TGF-βRII-deficient mice in which T cells are engineered to express TCR with high affinity, TGF-βRII-deficient T cells with high-affinity TCR have been shown to express high levels of IL-7Rα and thereby maintaining their homeostatic survival, whereas T cells with low-affinity TCR show impaired homeostatic survival in the absence of TGF-β (Robinson and Gorham 2007). These findings confirm that the strength of both TGF-β and TCR

signalling contribute to IL-7Rα expression. Therefore, loss of TGF-β signalling can affect TCR diversity and repertoire changes of naïve peripheral T cells resulting from selective loss of low-affinity T cells (Robinson and Gorham 2007). TGF-β is also found to regulate the homeostasis of peripheral $CD8^+$ T cells by promoting IL-7Rα expression. Impaired TGF-β signalling has resulted in altered homeostasis and aberrant activation of $CD8^+$ T cells with limited TCR diversity (Johnson and Jameson 2012) (figure 1.8 C).

1.5.1.3. TGF-β in T cell differentiation

TGF-β generally suppresses T cell activation by interfering with TCR signalling. It also inhibits differentiation of specific T cells including T helper 1 (Th1) and Th2 cells (Chen, et al. 2003). TGF-β suppresses differentiation of Th1 cells by inhibiting the expression of transcription factors T-bet and Stat4 and thereby inhibiting production of IFN-γ via Smad2/3 signalling pathway (Gorelik, Constant and Flavell 2002, Lin, et al. 2005, Takimoto, et al. 2010) . TGF-β also inhibits Th2 differentiation by blocking the expression of GATA-3, a lineage-defining transcription factor that mediates Th2 differentiation (Gorelik, Fields and Flavell 2000). The inhibition of GATA-3 expression is mediated by Smad-signalling pathway which results in the induction of Sox-4, a transcription factor that binds and inhibits GATA-3 (Kuwahara, et al. 2012) (figure 1.8 D).

TGF-β can also induce IL-9- and IL-10-secreting Th9 cells in co-operation with the Notch pathway and IL-4 signalling. Although Th9 cells produce abundant levels of IL-10, they act mainly as effector cells not regulatory cells (Dardalhon, et al. 2008). TGF-β is also found to promote the differentiation of T-follicular helper (Tfh) cells. Here, TGF-β acts as a pivotal co-factor in the presence of IL-12 and IL-23 which in turn induce Tfh cells (Schmitt, et al. 2014) (figure 1.8 D).

TGF-β can also regulate the differentiation of Th17 cells that are characterised by the expression of retinoic acid receptor-related orphan receptor C (RORC) in humans and retinoic acid receptor-related orphan receptor-γ t (RORγt) in mice (Gutcher, et al. 2011). TGF-β, in the presence of IL-6, induces the expression of RORγt/RORC that promotes the differentiation of Th17 in a Smad2-dependent and -independent manner (Ivanov, et al. 2006, Takimoto, et al. 2010, Martinez, et al. 2010). In addition, TGF-β mediates Th17 differentiation by suppressing the

expression of STAT4, GATA-3, GFi-1 and Eomesodermin, thus curbing Th1- and Th2-cell differentiation (Das, et al. 2009, Ichiyama, et al. 2011, Zhu, et al. 2009) . Therefore, TGF-β enhances the differentiation of Th17 cells both directly and indirectly by blocking T-cell differentiation into other cell lineages. It is well-known that Th17 are highly plastic cells which function paradoxically as immunoregulatory or inflammatory cells based on cumulative signalling pathways mediated by different cytokines within the local microenvironment (Sharma, Kaveri and Bayry 2013). The presence of TGF-β within the local milieu affects the cellular decision of Th17 to act as immunoregulatory or inflammatory. The absence of TGF- β, but in the presence of IL-6, IL-1β and IL-23 induces differentiation of Th17 (Ghoreschi, et al. 2010), whereas the presence of TGF-β, IL-6 and IL-1β mediates the induction of regulatory Th17 (McGeachy, et al. 2007) (figure 1.8D), therefore suggesting its strong role in the developing regulatory T cells.

Differentiation of Th17 cells can be interfered with by the differentiation of peripheral or induced Tregs based on the concentration of TGF-β in the local microenvironment. At low concentrations, TGF-β can cooperate with IL-6 and IL-21 to induce the expression of IL-23 receptor, thus promoting the differentiation of regulatory Th17 cells. However, at high concentrations, TGF-β inhibits the expression of IL-23 receptor and induces iTregs or pTregs differentiation (Knochelmann, et al. 2018, Sanjabi, Oh and Li 2017). The divergent effects of TGF-β on pTregs versus Th17 cell-fate indicate the context-dependent function of this versatile cytokine.

Beside of nTregs that are generated in the thymus, naïve T cells can gain Foxp3 expression and differentiate into induced Tregs (iTregs) cells in peripheral tissue (Abul 2003). Peripherally induced Tregs constitute 20-40% of the fully-mature Treg-cell pool in naïve steady-state mice and act in a cooperation with nTregs to maintain immune tolerance (Petzold, et al. 2014). Activated $CD4^{+}CD25^{-}$ or $CD4^{+}Foxp3^{-}$ T cells can be converted into $CD4^{+}CD25^{+}Foxp3^{+}$ iTregs *in vitro* in the presence of TGF-β which maintain their immunosuppressive activities (Huber and Schramm 2006). The induction of Foxp3 expression in iTregs is mediated by Smad3 which can be recruited to bind to the conserved Smad3-binding sequence in the Foxp3 enhancer CNS1 region which activates the induction of Foxp3 expression

(Tone, et al. 2008). In addition to TGF-β, other factors are found to essential for the induction of iTregs, including TCR activation-derived signals and the presence of IL-2. IL-2 is found to stimulate the transcription factor STAT5 which binds to the Foxp3 promoter, thus inducing Foxp3 expression (Burchill, et al. 2007). Suboptimal TCR signals with decreased CD28 signalling can also promote Foxp3 induction and iTregs generation in the presence of IL-2 and TGF-β (Josefowicz, Lu and Rudensky 2012).

iTregs can also be developed in the gut-associated lymphoid tissue (GALT). GALT-related $CD103^+$ DCs are found enriched at the mucosal site and implicated in the generation of iTregs there (Kushwah and Hu 2011). $CD103^+$ DCs mediate the recruitment of T cells to the gut by inducing the expression of α4β integrin and CCR9 (homing receptors) on T cells (Del Rio, et al. 2010). In addition, $CD103^+$ DCs are found to expressed high levels of retinal dehydrogenase which is an enzyme that converts vitamin A into retinoic acid (RA). RA is redundantly required for $CD103^+$ DCs to upregulate gut-homing receptors on T cell (Iwata, et al. 2004). $CD103^+$ DCs are found to preferentially express αvβ8 integrin which support their ability to activate TGF-β, thus promoting induction of iTregs (Boucard-Jourdin, et al. 2016). The generation of iTregs by $DC103^+$ DCs is found to be mediated by RA and TGF-β-dependent manners (Coombes, et al. 2007). Moreover, $CD103^+$ DCs are also found to express TGF-βR1 which has a pivotal, cell-intrinsic role in the development of these cells, suggesting that 103^+ DCs may require TGF-β to maintain their maturation and in turn mediate the induction of iTregs (Bain, et al. 2017).

1.5.1.4. TGF-β in T cell tolerance

The immune system has two main mechanisms to prevent autoimmunity. Central tolerance is an essential process in which self-reactive T cells are carefully removed in the thymus. Peripheral tolerance is developed to monitor autoreactive T cells that have escaped central tolerance and released to the periphery (Mueller 2010, Hogquist, Baldwin and Jameson 2005). The fatal inflammatory disorders that have occurred in mice with deficiency or total loss of TGF-β signalling reflect the crucial role of TGF-β in immune tolerance (Marie, et al. 2006, Li, et al. 2006). Loss of tolerance that emerges in the absence of TGF-β is not only caused by impaired activity of Tregs as the adoptive transfer of Tregs to TGF-β-deficient mice fails to

completely manage inflammatory disorder (Marie, et al. 2006). These findings indicate that peripheral tolerance is regulated directly and indirectly by TGF-β signalling. TGF-β can directly regulate peripheral tolerance by maintaining homeostatic survival of low-affinity $CD4^+$ T cells which maintain a unique regulatory population of $CD4^+$ T cells, called "deletor" T cells. These deletor cells control TCR repertoire and diversity of T cells and restrict the frequency of antigen-specific T cells that are highly reactive to the same antigen in the peripheral T cell pool (Singh, Bando and Schwartz 2012). Loss of TGF-β signalling in the developing T cells may contribute to establishment a lymphopenic environment which lacks low-affinity T cells and is preferentially enriched with high-affinity autoreactive T cells in humans (Singh, et al. 2012). TGF-β is also found to curb tissue damage by converting pathogenic, autoreactive $CD4^+$ T cells to a non-pathogenic phenotype (Reis, et al. 2013). In a transgenic diabetes mouse model, the absence of TGF-β signalling leads to activation of pathogenic $CD4^+$ T cells that promote disease progression by destroying pancreatic islets by autoreactive $CD4^+$ T cells (Ishigame, et al. 2013). TGF-β can also control peripheral tolerance in an indirect way by promoting the development of peripheral iTregs that inhibit activation and proliferation of highly-autoreactive T cells thereby preventing the development of autoimmune diseases (Oh and Li 2013). The role of TGF-β in the generation of iTregs in peripheral tissues has been mentioned in the previous section (1.6.1.3).

1.5.2. TGF-β in cancer

TGF-β, as a pleiotropic cytokine, is found to have two conflicting roles in cancer development and progression. In initial stages of cancer, TGF-β serves as a tumour suppressor by promoting inhibition of cell proliferation and induction of apoptosis. However, in advanced stages of cancer, TGF-β acts as a tumour promoter which enhances cell motility, invasion, metastasis and maintenance of cancer stem cells. This opposite switch in the function of TGF-β is defined as "TGF-β paradox" (Wendt, Tian and Schiemann 2012).

TGF-β-mediated growth inhibition is mediated by the suppression of cyclin-dependent kinases (CDKs) and the down-regulation of MYC. TGF-β exerts its inhibitory activity during the G1 phase of cell cycle where TGF-β inhibits the activation of several CDKs including cyclin D-CDK4, cyclin D-CDK6, cyclin E-CDK2

and A-CDK2 via several mechanisms which results in dephosphorylation of retinoblastoma protein (pRb) and thereby cell cycle arrest (Zhang, Alexander and Wang 2017). Inhibition of the MYC gene by TGF-β is mediated via Smad3/Smad4 signalling pathway in which Smads complex is recruited to the nucleus to bind to a transcriptional inhibitory element on the MYC promoter and thus curb MYC expression. TGF-β can also induce cell apoptosis by promoting expression of several pro-apoptotic genes via Smad-dependent pathway. These genes include the TGF-β-inducible early-response gene (TIEG1), the death-associated protein kinase and the SH2 domain-containing inositol-5-phosphatase (Drabsch and Ten Dijke 2012).

This antimitogenic capacity of TGF-β is usually overwhelmed by the potent mitogenic cancer cells that can mutate both TGF-β receptors and Smads thereby switching the role of TGF-β to enhance cell growth, motility and invasion. This mis-regulation of TGF-β signalling contribute to the cancer progression and metastasis (Massagué 2008). Biallelic inactivation of TGF-βRII by mutations is found in many types of cancer including colon, gastric, biliary, ovarian, oesophageal and head and neck cancer (Levy and Hill 2006). TGF-βRII mutations are also highly associated with microsatellite instability and promote cancer progression through altered TGF-β signalling pathway (Levy and Hill 2006). TGF-βRI is also found mutated in ovarian and head and neck cancers (Chen, et al. 2001b, Chen, et al. 2001a) . In addition to mutation in the coding region of TGF-βRII and TGF-βRI, loss of expression of both receptors is also another type of TGF-β signalling alterations found frequently in cancer. Loss of expression of TGF-βRII is frequently detected and associated with poor prognosis in non-small cell lung carcinomas (NSCLCs), bladder cancers, head and neck squamous cell carcinomas (HNSCC), ovarian carcinomas, oesophageal carcinomas, prostate and breast cancers (Malkoski, et al. 2012, J. Bian, et al. 2013, Weber, et al. 2008, Roane, Arend and Birrer 2019, Fukai, et al. 2003, Tu, et al. 2003, Busch, et al. 2015). Although mutations in TGF-βRI are somewhat rare, loss of expression of TGF-βRI is more frequently found (Levy and Hill 2006). It is very commonly identified in HNSCC and bladder carcinomas (Fukai, et al. 2003, Tokunaga, et al. 1999).

Smad proteins, as key components of TGF-β signalling pathway, have been found mutated within certain cancers. Smad2 protein is found mutated at very low

frequency in 8% of cervical cancers, 8% in colorectal cancer, 3% in hepatocellular carcinoma (HCC) and 2% in NSCLC. Loss of Smad2 expression contribute to the delay in P15 upregulation and failure to inhibit c-MYC by TGF-β, thereby promoting tumour-cell growth (Samanta and Datta 2012). Loss of the expression of Smad3 is also found in some of gastric cancers and is associated with increased tumorigenesis. In choriocarcinoma, a subtype of uterine cancers, loss of Smad3 expression is found to be correlated with down-regulation of TIMP-1 (tissue inhibitor of metalloprotease-1 and thus promotes the activity of matrix metalloproteinases for supporting tumour invasion (Xu, Chakraborty and Lala 2003). Loss of Smad4 expression is also found and correlated with metastasis and poor prognosis in certain cancers including breast (Liu, et al. 2015), pancreas (Shugang, et al. 2016), colon(Yan, et al. 2016) and lung (Bian, et al. 2015).

TGF-β is found to mediate tumour progression by a process called epithelial-mesenchymal transition (EMT) in which tumour cells lose epithelial phenotype and acquire mesenchymal phenotype thereby promoting their migratory and invasive properties. This dynamic process is normally occurred during embryonic development, wound healing, tissue regeneration and fibrosis (Moustakas and de Herreros 2017). However, EMT is exploited by tumour cells to maintain cancer progression and metastasis. The essential features of EMT include interruption of cell-cell adhesion and cellular polarity, remodelling of the cytoskeleton and defects in cell-matrix adhesion (Guarino, Rubino and Ballabio 2007). TGF- β signalling is found to downregulate the expression of epithelial markers; E-cadherin, zona occludins-1 (ZO1), B4-integrin, α- and γ-catenin that maintain architecture of epithelial tissues. Whereas, upregulation of mesenchymal markers; vimentin, fibronectin, N-cadherin, smooth-muscle actin and Twist, is enhanced in response to TGF-β signalling to mediate tumour cell survival, migration and invasion (Drabsch and Ten Dijke 2012).

1.6. TGF-β enriched tumours

Overexpression of TGF-β and increased its plasma level has been correlated with advance tumour progression and metastasis in certain types of cancer.

In breast cancer, increased TGF-β1 serum levels in the blood of patients have been correlated with advanced stages of the disease (Sheen-Chen, et al. 2001). Also, TGF-β1 is found to be highly expressed in breast cancer tissues and is corelated with poor prognosis and shorter disease-free survival (Desruisseau, et al. 2006). In addition, high expression of TGF-β1 in breast tumours has been associated with increased lymph node metastasis (De Kruijf, et al. 2012). In triple negative breast cancer (TNBC), which is the most aggressive subtype of breast cancer, high expression of TGF-β1 is found to be significantly correlated with advanced tumour histological grade, axillary lymph node metastasis and shorter 5-year disease-free survival rate (Ding, et al. 2016).

In colorectal cancer, the expression of TGF-β1 is found to be higher in both tissue and plasma of patients with metastatic disease compared to those with non-metastatic disease. Even in patients with non-metastatic colorectal cancer, the expression of TGF-β1is found significantly correlated with lymph node involvement (Langenskiöld, et al. 2008). Xiong et al. have also found that patients with stage 3 and 4 of metastatic colorectal cancer show higher serum levels of TGF-β1 than those with stage 1 and 2 of non-metastatic tumours (Xiong, et al. 2002).

In prostate cancer, Reis et al. have found that TGF-β1 is highly overexpressed in 33% of prostate tumour tissues and is correlated with poor prognosis, confirming the importance role of TGF-β1 as a prognostic biomarker for prostate cancer (Reis, et al. 2011). Wu et al. have also found that elevated TGF-β1 expression is found highly correlated with aggressive advanced stages of prostate cancer and with accumulation of Tregs within tumour tissues (Wu, et al. 2015). The same author has also found that increased expression of TGF-β1 is significantly correlated with tumour-cell resistance to radiation therapy in patients with prostate cancer (Wu, et al. 2015).

In pancreatic cancer, Zhao et al. have demonstrated that high serum levels of TGF-β1 are dramatically correlated with reduced survival rate, advanced stages of disease and lymph node metastasis in patients with pancreatic ductal adenocarcinoma (PDAC) (Zhao, et al. 2016). Park et al. have found that serum level of TGF-β1 could be used as prognostic biomarkers at the diagnosis of unresectable pancreatic cancer patients. The authors have demonstrated that patients with high

serum levels of soluble TGF-β at the time of diagnosis have shown shorter overall survival rate compared to patients having low serum levels of TGF-β1 at the time of diagnosis (Park, et al. 2018). Javle et al. have also supported these findings. The authors have found that high expression of TGF-β1 in PDAC tumour tissues and high serum levels of TGF-β1 are significantly correlated with reduced overall survival rate in patients with advanced PDAC (Javle, et al. 2014).

In lung cancer, Li et al. have confirmed that high expression of TGF-β correlates positively with worse prognosis in patients with lung cancer (Li, et al. 2019). Luo et al. have found that lung cancer patients who have high serum levels of TGF-β1 following radiotherapy show unresponsive progressive disease than those with low levels of TGF-β1. The authors have also found that high serum levels of TGF-β1 are significantly correlated with $CD4^+$ T cells and $CD8^+$ T cells percentages, suggesting the impact of TGF-β1 on circulating T cells for promoting disease progression (Luo, et al. 2018). Kim et al. have also found that high serum levels of VEGF, IL-1β and TGF-β1 in patients with advanced non-small cell lung carcinoma are significantly correlated with shorter overall survival rate. The authors also found that high serum levels of these cytokines are significantly correlated with increased number of leukocytes (Kim, et al. 2013). Moreover, Xie et al. have demonstrated that tumours with positive expression of TGF-β1 have been correlated with late stages, lymph node involvement and reduced overall survival rate in patients with lung adenocarcinoma (Xie, He and Wei 2015).

In hepatocellular carcinoma (HCC), Kohla et al. have found that increased serum levels of TGF-β1 are associated with disease aggressiveness and shorter overall survival rate in patients with HCC (Kohla, et al. 2016). Lin et al. have also found that elevated serum levels of TGF-β1 are significantly correlated with advanced stage of disease and reduced progression-free survival and overall survival rates in patients with HCC. The high levels of TGF-β1 are also found associated with increased resistance to sorafenib, a multikinase inhibitor (Lin, et al. 2015). An et al. have also confirmed that high serum levels of TGF-β are significantly correlated with high percentage of TGF-β-expressing Tregs in the blood of patients with HCC, and both TGF-β serum levels and TGF-β-expressing Tregs percentage are correlate with reduced 5-year overall survival rate (An, et al. 2018).

In oesophageal cancer, Von Rahden et al. have found that TGF-β1 is found significantly overexpressed in tissues derived from patients with primary oesophageal adenocarcinoma and its overexpression is significantly associated with advanced stages of disease, poor prognosis and low cumulative survival rate (von Rahden, et al. 2006). Sun et al. have also noticed that elevated serum levels of TGF-β1 in patients with oesophageal carcinoma following radiotherapy are significantly correlated with advanced unresponsive disease and very shorter overall survival, compared to patients who have low serum levels of TGF-β1 (Sun, et al. 2007).

According to the results of studies discussed above, it seems that the expression of TGF-β1 could be used as a prognostic and predictive biomarker for cancers in which TGF-β1 is significantly enriched within tumour tissue or in the peripheral blood. In addition, some of studies have shown that the expression of TGF-β1 is significantly correlated with increased percentage of Tregs in the blood of patients, confirming the clinical importance of this correlation as prognostic biomarkers for TGF-β1-enriched cancers. The clinical importance of Tregs in cancer will be discussed in the next section.

1.7. The clinical importance of Tregs in cancer

In cancer, immunosuppressive T regulatory cells (Tregs) are found to inhibit immune responses directed against tumour cells and promote immune evasion and cancer progression (Elkord, et al. 2010). Many studies have confirmed that tumour-infiltrating Tregs are associated with a worse prognosis in various types of cancer, including breast cancer, ovarian cancer, lung cancer, hepatocellular carcinoma, glioblastoma, pancreatic ductal adenocarcinoma, renal carcinoma, non-Hodgkin's lymphomas, melanoma and other malignancies (Zhang, et al. 2017, Curiel, et al. 2004, Tao, et al. 2012, Sun, et al. 2017, Sayour, et al. 2015, Tang, et al. 2014, Yang, et al. 2014, Yang, et al. 2006, Li, et al. 2010, Gerber, et al. 2014, Shang, et al. 2015a).

In breast cancer, Merlo et al. have found that increased Foxp3 expression within tumour microenvironment is significantly associated with poor overall survival rate. The authors have also found that Foxp3-positive tumours is significantly associated with lymph node metastasis, distant relapse and shorter distant-metastasis (DM) free survival rate (Merlo, et al. 2009). Sun et al. have confirmed that increased frequency of $Foxp3^+$ Tregs and $PD\text{-}1^+$ immune cells in breast cancer patients is

significantly correlated with reduced survival rate (Sun, et al. 2014). Kim et al. have also found that increased frequency of $Foxp3^+$ Tregs within the tumour stroma derived from breast cancer patients is significantly correlated with decreased infiltration of $CD8^+$ T cells and reduced 5-year disease free survival rate (Kim, et al. 2014). Zhou et al. have also found that high $Foxp3^+$ Tregs infiltration with the tumour microenvironment is correlated with high histological grade, $HER2^+$ tumours and reduced recurrence-free survival rate in breast cancer patients (Zhou, et al. 2017). Li et al. have also found that high expression of PD-L1 is significantly correlated tumour-infiltrating $Foxp3^+$ Tregs and both are correlated with reduced overall and recurrence-free survival rates in patients with breast cancer (Li, et al. 2010).

In lung cancer, Liu et al. have demonstrated that Tregs are significantly enriched in the blood of patients with NSCLC at first relapse following radiotherapy. The authors have also found that the accumulation of Tregs in the peripheral blood is significantly correlated with reduced progression-free survival rate (Liu, et al. 2017b). Tao et al. have also demonstrated that increased tumour-infiltrating Tregs within the tumour tissue derived from patients with NSCLC is significantly associated with shorter overall and relapse-free survival rate (Tao, et al. 2012). Kotsakis et al. have also confirmed that the frequency of circulating Tregs is significantly elevated in the peripheral blood derived from untreated NSCLC patients compared to healthy individual. Increased circulating Tregs is found significantly correlated with reduced progression-free and overall survival rate (Kotsakis, et al. 2016). Hanagiri et al. have also found that increased proportions of Tregs in the regional lymph node lymphocytes in patients with NSCLC following surgery is significantly correlated with lymph node metastasis and shorter overall survival rate (Hanagiri, et al. 2013).

In ovarian cancer, Sato et al. have demonstrated that high Tregs frequency relative to total T cells or $CD8^+$ T cells is significantly associated with reduced overall survival rate in patients with ovarian compared to low Tregs frequency group of patients (Sato, et al. 2005). Curiel et al. have also found that enrichment of tumour-infiltrating Tregs is correlated with advanced stages of ovarian cancer. Increased infiltration of Tregs within the tumour tissue is also significantly correlated with

reduced overall survival rate in patients with ovarian cancer from stage 1 to 4. In addition, patients with ovarian cancer stage 4 have shown more significant reduction in the overall survival rate than those at stages 1 to 3, confirming the role of Tregs in the progression of ovarian cancer (Curiel, et al. 2004).

In pancreatic cancer, Hwang et al. have found that increased frequency of tumour-infiltrating $Foxp3^+$ Tregs within the microenvironment is significantly correlated with low frequency of granzyme B^+ CD8 T cells and reduced overall survival rate in patients with PDAC (Pancreatic ductal adenocarcinoma) following distal pancreatectomy (Hwang, et al. 2016). Liu et al. have also confirmed that elevated percentages of circulating Tregs in the peripheral blood of patients with unresectable PDAC prior to chemotherapy is significantly correlated with reduced overall survival rate. The author have also found that high percentage of circulating Tregs following chemotherapy treatment is significantly correlated with decreased ratio of circulating $CD8^+$ T cells and reduced overall survival rate in patients with unresectable PDAC (Liu, et al. 2017a). Tang et al. have also demonstrated that $Foxp3^+$ Tregs are significantly enriched in the juxtatumoural stroma immediately beside tumour epithelial cells in patients with PDAC and their enrichment is significantly correlated with reduced frequency of $CD8^+$ T cells and reduced overall rate. The enrichment of $Foxp3^+$ Tregs in the tumour stroma is also found highly correlated with undifferentiated tumour cells and high histological grades (Tang, et al. 2014).

In hepatocellular carcinoma (HCC), Sun et al. have performed a comprehensive meta-analysis on 3854 patients with HCC collected from 27 cohort studies. The authors have found that high frequency of $Foxp3^+$ Tregs in the tumour tissues and peripheral blood derived from HCC patients is significantly correlated with advanced disease stages, high histological grades with poorly undifferentiated tumour cells and reduced overall and disease-free survival rate (Sun, et al. 2017).

In gastric cancer, similar findings have been found by Lee et al. The authors have confirmed that elevated expression of $Foxp3^+$ Tregs is significantly correlated with poor prognosis and reduced overall survival rate in 2941 patients with gastric cancer (Zheng, et al. 2017).

In brain cancer, Yue et al. have demonstrated that elevated infiltration of Foxp3$^+$ Tregs within the tumour tissue derived from patients with glioblastoma (GBM) is significantly corelated with reduced overall and progression-free survival rat, compared to patients with low Tregs infiltration (Yue, et al. 2014).

Taken together, the presence and infiltration of Foxp3$^+$ Tregs within the tumour microenvironment contributes to the tumour progression and poor prognosis in the most common types of cancer as mentioned above. Therefore, inhibiting their role within tumour milieu will improve the treatment outcomes of cancer therapy, especially cancer immunotherapy (Shitara and Nishikawa 2018).

Several approaches have been followed for targeting tumour-infiltrating Tregs which represent a major obstacle for effective anti-tumour immunotherapy. These approaches including: **(1)** nonspecific depletion or modulation of Tregs using conventional anticancer drugs such as cyclophosphamide, sunitinib, sorafenib and imatinib (Galluzzi, et al. 2012, Ge, et al. 2012, Adotevi, et al. 2010, Desar, et al. 2011), **(2)** specific depletion of Tregs using Daclizumab (anti-CD25 monoclonal antibody) and Denileukin Diftitox which is an engineered protein combining IL-2 and Diphtheria toxin (Rech, et al. 2012, Attia, et al. 2005), **(3)** targeting mechanisms mediating Tregs recruitment into tumours using Mogamulizumab (anti-CCR4 monoclonal antibody) (Kurose, et al. 2015), (4) depletion of Tregs using checkpoint inhibitors such as Ipilimumab (anti-CTLA-4 monoclonal antibody) and Nivolumab (anti-PD-1 monoclonal antibody) (Hellmann, et al. 2018, Larkin, et al. 2015). However, these therapies are designed to deplete Tregs systemically and have found to induce severe autoimmune problems (Bakacs, Mehrishi and Moss 2012, Kurose, et al. 2015, Larkin, et al. 2015). Therefore, therapeutic agents that can selectively deplete tumour-infiltrating Tregs are still required. However, developing such agents requires an in-depth understanding of the difference between tumour-infiltrating Tregs and other circulating Tregs at normal tissue sites, in addition to identification additional markers that can selectively distinguish tumour-infiltrating Tregs from other circulating Tregs (Liu, Workman and Vignali 2016a, Togashi, Shitara and Nishikawa 2019).

The origin of tumour-infiltrating Tregs is still controversial. Studies have found that most of Tregs within tumour are induced Tregs (iTregs) developed from naïve CD4$^+$

T cells in the presence of tumour-drived TGF-β and other co-stimulatory cytokines secreted by DCs and tumour-infiltrating macrophages (Valzasina, et al. 2006, Liu, et al. 2007, Ghiringhelli, et al. 2005). Other studies have reported that natural Tregs (nTregs) are also enriched within the tumour nest and contribute to tumour immunosuppressive. Moreover, several studies have assessed the crucial role of chemokines receptors in the recruitment of both naïve $CD4^+$ T cells and circulating Tregs into the tumour microenvironment (Gobert, et al. 2009, Mizukami, et al. 2008, Facciabene, et al. 2011). However, it is difficult to identify the origin of tumour-infiltrating Tregs without finding a distinct biomarker that can discriminate iTregs from nTregs since both subsets rise from $CD4^+$ T cells and share many features.

1.8. Aims and objectives

The main aim of this study is to identify cell surface biomarkers that can differentiate phenotypic features of peripherally-induced Tregs (iTregs) from thymic-drived natural Tregs (nTregs) in the blood and tissue of cancer patients. The identification of such biomarkers will enable us to further delineate the origin of tumour-infiltrating Tregs and circulating Tregs in the tissue and blood of cancer patients, respectively by understanding their biology in detail. This will also facilitate development of therapeutic agents that can selectively target iTregs without inhibiting the function of nTregs, thereby enhancing the efficacy of immunotherapy in combination with other conventional anti-cancer therapies for better treatment outcomes and minimising the chance of development autoimmune diseases. Identification of such biomarkers will also help to predict the prognosis of cancer patients at the time of diagnosis based on their immunological profiling.

To achieve the main aim, objectives are as follows:

1- Establishment of mice model for generating $CD4^+CD25^+Foxp3^+$ induced Tregs from purely sorted naïve $CD4^+CD25^-$ T cells in the presence and absence of tumour cells *in vitro*.

2- Isolation of a pure population of naïve natural Tregs, in vitro-induced Tregs, naïve $CD4^+CD25^-$ T cells using cell sorting.

3- Optimisation of a quantitative proteomics approach to profile iTregs and nTregs populations

4- Verification and validation of any discovered biomarker/s.

The reasons of establishing a mice model instead of human model are that; **first,** establishing a successful mice model will help in validating the discovered biomarkers pre-clinically and will be required as a routine step of any clinical trial for biomarker validation. **Second,** based on the main aim of this project, finding healthy individuals living in a pathogen-free environment and who have not exposed to any infection during their lives is impossible. Therefore, isolation and sorting Tregs from healthy donors' blood based on the expression of CD25 and Foxp3 will not give a clue whether the sorted Tregs are natural or induced Tregs. Thus, studying phenotypic features and discrepancies between nTregs and iTregs will be very difficult.

Chapter 2.

Development of an *in vitro* murine model to study induced T regulatory cells

2.1. Introduction

The most critical trait of immune system is the ability of maintaining homeostasis in the periphery. Based on this feature, immune system prevents the development of autoimmunity. Regulatory T cells (Tregs) are specialised subsets of $CD4^+$ T cells and function to maintain immune homeostasis and tolerance in the periphery. Tregs can be divided into two main subpopulations: thymic-derived natural Tregs (nTregs) and peripherally-derived Tregs or induced Tregs (p/i Tregs). nTregs are developed within the thymus during medullary thymic selection and function to maintain immune tolerance against auto-antigens, thus preventing autoimmunity. Induced or peripheral Tregs are developed in the peripheral tissues, mainly the intestine, to maintain immune tolerance against innocuous antigens expressed by commensal microflora. The early process of the development of both nTregs and p/i Tregs that occurs in the thymus where all T-cell populations develop and mature will be discussed in the introduction of this chapter.

2.1.1. Thymopoiesis

Thymopoiesis is defined as the process by which thymocytes differentiate into mature T-lymphocytes in the thymus. It is a highly regulated process which involves migration of the earliest T-lineage progenitors (ETPs) from bone marrow niches to the thymus where further development and differentiation of mature T cells occurs (Schwarz and Bhandoola 2006). The thymus is a specialised lymphoepithelial organ and consists of two main compartments: the cortex and the medulla, each of which contains particular populations of thymic epithelial cells, dendritic cells and endothelial cells, thus furnishing an exceptional microenvironment for the optimal production of a diverse T cell repertoire (Gordon and Manley 2011).

The origin of the ETPs settled in the thymus is thought to be unknown. However, to be qualified as a T-lineage progenitor, a bone marrow progenitor must achieve two

key requirements. First, it must show the ability to effectively generate T cells in the thymus. Second, it must have the potential to migrate from the bone marrow to the thymus (Witt and Robey 2005). Several bone marrow progenitors are found to have the potential of T-lineage progenitors, including pluripotent self-renewing haematopoietic stem cells (HSCs), committed myeloid-lymphoid progenitors (MLPs) that lack the potential for erythroid and megakaryocyte lineage, and common committed lymphoid progenitors (CLPs) (Wu 2006). Although the definite identity of the bone marrow cell precursors that generate T cells is still debated, the origin of the ETPs has been proposed to be derived from HSCs that can self-differentiate into LSK progenitors that show a phenotype $Lineage^{negative}$ $Sca\text{-}1^{hi}$ $c\text{-}Kit^{hi}$ (LSK) (figure 2.1). The LSK progenitors have the potential to differentiate into two main subpopulations; $CD62L^{+}$ IL-7Rα^{-} LSK progenitors and $CD62L^{-}$ IL-7Rα^{-} LSK progenitors that are also called early lymphoid progenitors (ELPs) (Wu 2006). The $CD62L^{+}$ LSK and ELP progenitors are thought to be released from the bone marrow into the bloodstream where they may migrate to thymus via the expression of CD62L (L-selectin) which is function as a cell adhesion molecule (Wu 2006). The migration of ELPs to the thymus is through to be via the expression of P-selectin, however this is still enigmatic. In the thymus, $CD62L^{+}$ LSK progenitors undergo further differentiation to become ETPs which is mediated by Notch-1 signalling (Wu 2006) (figure 2.1).

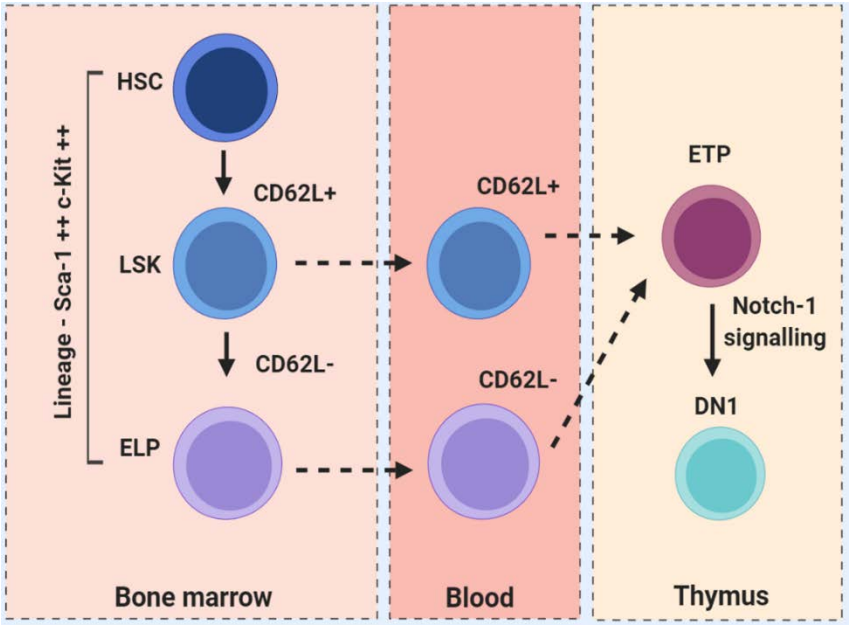
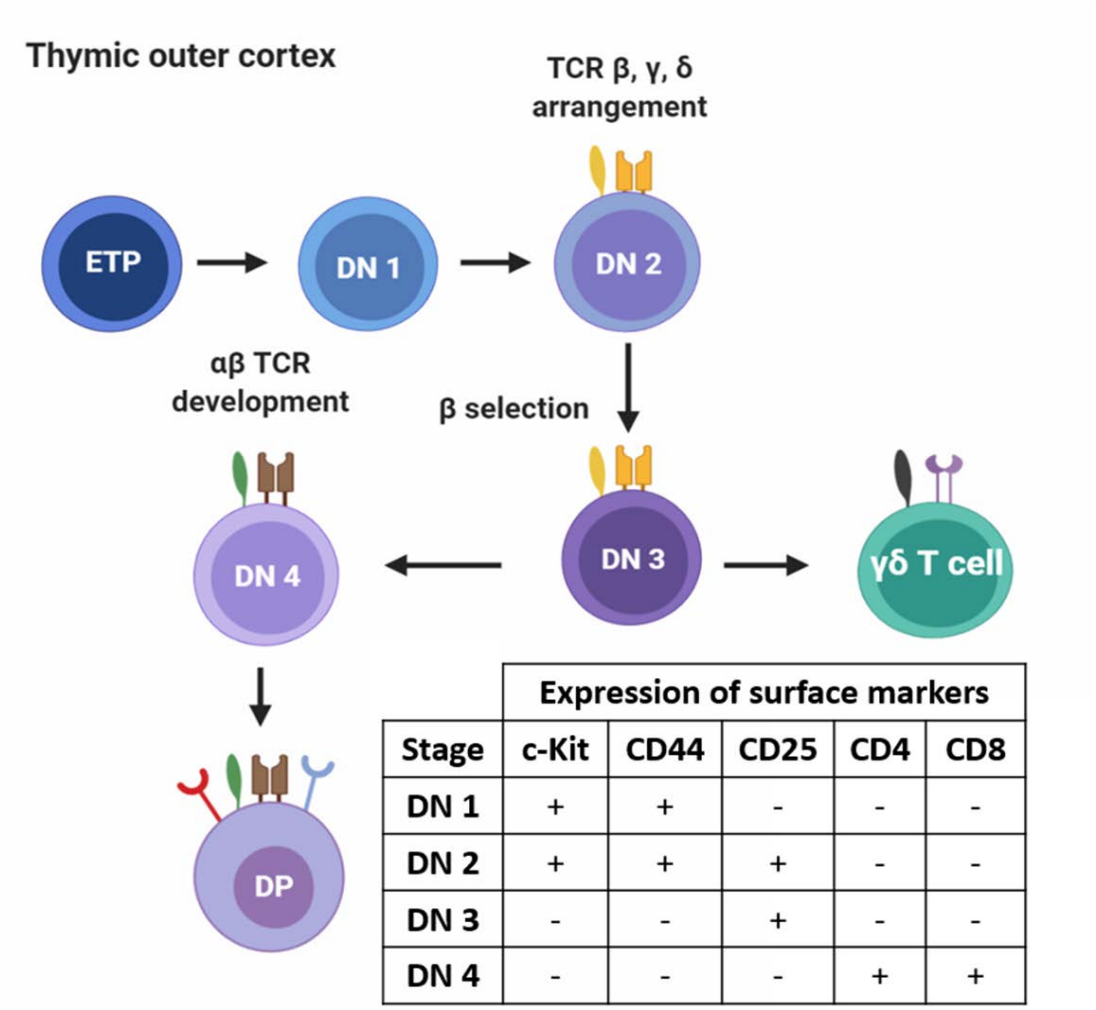


Figure 2.1. A proposed representation for initial stages of early T cell progenitors (ETPs) development from the bone marrow to the thymus. HSC: haematopoietic stem cell, LSK: Lineage Sca-1 c-Kit phenotype, ELP: early lymphoid progenitor, ETP: early T-lineage progenitor, DN: double negative. This figure is modified from (Wu 2006).

2.1.2. Ontogeny of T cells

After arriving the thymus, T-cell precursors (ETPs) face Notch-1 ligands that are amply expressed by the thymic epithelium. Notch-1 signalling is considered as a key mediator of T cell-lineage commitment and regulates the early decision between whether to become a T or B lymphocytes (Deftos and Bevan 2000). After interaction with with Notch-1 ligands, T-cell precursors migrate to the outer thymic cortex where they start to develop their T-cell receptors. The primitive T cell precursors lack the expression of CD4 and CD8 and therefore recognised as double negative (DN) cells. These DN cells can be further classified into four subsets, DN1 – DN4, based on certain cell surface antigens including c-Kit (CD117), CD44 and CD25 (Germain 2002). T-cell precursors at the stage DN1 of development show phenotypic features of $CD117^{++}CD44^{+}CD25^{-}$ on their surface, and then after being cross-linked with Notch-1 ligands for T-cell lineage commitment, they travel to the cortex where they acquire the expression of CD25, becoming DN2 thymocytes ($CD117^{++}CD44^{+}CD25^{+}$). At this critical stage of development, the rearrangement for genes of the T cell receptors β, γ, and δ occur. At the end of DN2 stage, thymocytes or T cell precursors are entirely committed to the T-cell lineage (Wu 2006). During DN3 stage, thymocytes begin to reduce their expression of CD117 and CD44 and can

be identified as ($CD117^{+/-}CD44^{-}CD25^{+}$) thymocytes. DN3 cells then progress to develop their TCR-γ, -δ, and –β chains, and have the potential to make the first critical decision in T-cell development to become TCR- γδ or TCR-αβ T-cell lineage via a process called β-selection (Germain 2002). This process includes a glycoprotein called the pre-Tα chain that is specifically expressed at this stage. The pre-Tα plays as a surrogate for the actual TCR-α which has not been rearranged yet and binds to the successfully rearranged β chains and CD3 complex. This Immature TCR/CD3 complex is called the pre-T cell receptor (pre-TCR) (Spits 2002). At DN4 stage, DN4 thymocytes that have successfully developed their pre-TCR begin to lose the expression of CD25 and CD117 to become $CD117^{-/low}CD44^{-}CD25^{-}$ thymocytes and mature directly to $CD4^{+}CD8^{+}$ double positive (DP) thymocytes after completing TCR-α chain locus rearrangement. $CD4^{+}CD8^{+}$ DP thymocytes express mature TCR-αβ on their cell surface (Harrington 2019).



	Expression of surface markers				
Stage	c-Kit	CD44	CD25	CD4	CD8
DN 1	+	+	-	-	-
DN 2	+	+	+	-	-
DN 3	-	-	+	-	-
DN 4	-	-	-	+	+

Figure 2.2. Schematic representation showing developmental stages of T cells in the murine thymic outer cortex. Upon arrival the thymus, ETPs migrate to the outer cortex where they differentiate into DN1 thymocytes (c-Kit$^+$CD44$^+$) that lack the expression of CD4 and CD8. DN1 thymocytes acquire CD25 expression during their development to become DN2 (c-Kit$^+$CD44$^+$CD25$^+$). During DN2 developmental stage, the rearrangement for genes of the T cell receptors β, γ, and δ occur. At DN3 stage, DN3 thymocytes lose the expression of c-Kit and CD44 and develop their TCR-β, -γ, and -δ. DN3 thymocytes undergo β-selection process by which they give rise to either γδ T cell or αβ T cell lineage. At DN4 stage, DN4 thymocytes develop expression of pre-αβ TCR and mature to CD4$^+$CD8$^+$ DP thymocytes after losing the expression of CD25. ETP: early T-lineage progenitor, DN: double negative, DP: double positive. This figure is modified from (Germain 2002).

2.1.3. Thymic selection of the T-cell repertoire

The random rearrangement of TCRs genomic loci during thymocytes development leads to generate a diverse TCR repertoire with different recognition specificities in the newly developed CD4$^+$CD8$^+$ DP thymocytes. At this stage, unlike mature CD4$^+$ and CD8$^+$, the DP thymocytes lack a well-programmed affinity for foreign antigen

plus a self-MHC (major histocompatibility complex) molecule; in theory they should have the potential to recognise soluble antigens either foreign or self, self-MHC molecules or antigen bound with a non-self MHC molecule (Singer, Adoro and Park 2008). However, to be qualified as a mature functional T cell, the T cell must recognise only foreign antigen attached with self-MHC molecules before leaving the thymus to peripheral tissues. To reach this level, DP thymocytes undergo two selection processes: positive selection that ensures MHC restriction, and negative selection that confirms self-tolerance (Starr, Jameson and Hogquist 2003) (figure 2.3). Positive selection is for DP thymocytes expressing receptors that can bind self-MHC molecules either class I or class II which are expressed by cortical thymic epithelial cells (cTECs). Cells that do not interact with MHC molecules will die by apoptosis (Takahama 2006). Then, positively selected thymocytes travel to the thymic medulla where they encounter with a variety of antigen-presenting cells including haematopoietic dendritic cells (DCs) and medullary thymic epithelial cells (mTECs). mTECs are found to produce a wide range of self-proteins, tissue specific self-molecules and display the majority of self-antigens, providing the thymocytes a test environment for negative selection (figure 2.3). Negative selection is a process of elimination thymocytes expressing receptors with high-affinity for self-MHC molecules or self-antigen presented by self-MHC (Klein, et al. 2014b, Sprent and Kishimoto 2002). The successful thymocytes that pass the negative selection will then leave the medulla as self-tolerant, mature naïve $CD4^+$ and $CD8^+$ T cells to the periphery (figure 2.3). A small subpopulation of thymocytes with high-affinity for self-MHC can survive the elimination of negative selection and are destined to become T regulatory cells (Tregs) (Sakaguchi, et al. 2008).

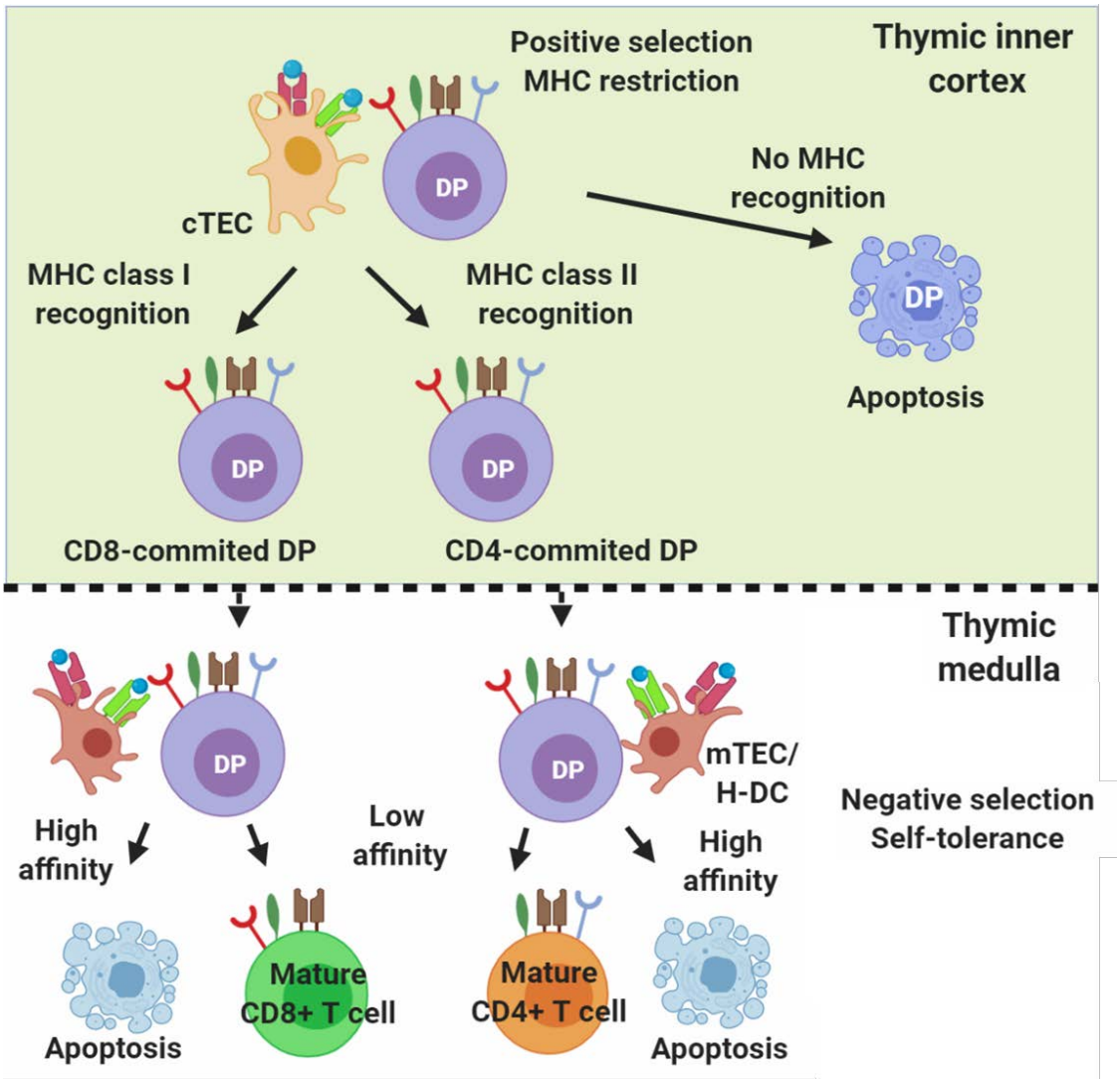


Figure 2.3. Schematic representation showing positive and negative thymic selection of the T-cell repertoire. After maturation, DP thymocytes migrate to the thymic inner cortex where they undergo positive selection that ensures MHC restriction. During this selection, DP thymocytes are exposed to self-MHC molecules expressed by cTECs. DP thymocytes that recognise self-MHC class II pass the selection to become CD4-commited DP thymocytes, whereas those recongise self-MHC class I become CD8-committed DP thymocytes. DP thymocytes that do not recognise self-MHC molecules are eliminated by apoptosis. Positively-selected CD4-commited and CD8-commited DP thymocytes migrate to the thymic medulla where they undergo negative selection that ensures self-tolerance. During this selection, both CD4-commited and CD8-commited DP thymocytes are exposed to self-MHC molecules attached with self-antigens that are expressed by mTEC and H-DC. CD4-commited DP thymocytes that recognise self-MHC class II with high affinity are eliminated by apoptosis, whereas those recognise self-MHC class II with low affinity are selected to become mature $CD4^+$ T cells. CD8-commited DP thymocytes that interact with self-MHC class I with low affinity are selected to become mature $CD8^+$ T cells, while those interact with high affinity are eliminated by apoptosis. Both mature $CD4^+$ and $CD8^+$ T cells are then released into the periphery. cTEC: cortical thymic epithelial cell, DP: double positive, mTEC: medullary thymic epithelial cell, H-DC: Haematopoietic dendritic cell, MHC: major histocompatibility complex. This graph is modified from (Germain 2002).

2.1.4. The role of thymic medulla microenvironment in Thymic tolerance

The negative selection imposed by the thymic medulla ensures that αβ-T cells leaving the thymus can recognise the self-MHC proteins and are tolerant to self - antigens simultaneously. The specialised mTECs and DCs, by expressing most of self-antigens enforce $CD4^+$ and $CD8^+$ single-positive (SP) thymocytes to undergo negative selection if they show high-affinity recognition for self-antigens. Thus, this clonal deletion is programmed to eliminate most of the self-reactive thymocytes (Klein, et al. 2014b). However, thymic tolerance machineries do not ensure the elimination of all autoreactive thymocytes, while efficiently prune the frequency of self-reactive cells leaving the thymus to the periphery through a process that may bypass the presence of flaws in the αβTCR repertoire. To compensate for the chaotic inefficiencies in negative selection, the medulla also distorts a subpopulation of $CD4^+$ SP thymocytes into the $Foxp3^+$ Tregs lineage (Kakugawa, et al. 2017). When migrated to the periphery, these cells are highly efficient in inhibiting autoreactive responses introduced by self-reactive T cells that evade thymic tolerance (Fan, et al. 2018).

Although, the precise mechanisms by which the mTECs select T cell repertoire is not fully elucidated, it is believed that the mTECs and thymic DCs are thought to act as key regulators for the development of $Foxp3^+$ Tregs, thereby confirming the important role of the thymic medulla in that process (Cowan, et al. 2013, Perry, et al. 2014). The complex heterogeneity of mTECs and DCs allows the medulla to manage the maturation of diverse T cells populations including conventional SP T cells and $CD4^+Foxp3^+$ Tregs (Takahama, et al. 2017).

In the murine thymus, mTECs are generally stratified into two distinct $mTECs^{low}$ ($MHC\text{-}II^{low}CD80^{low}$) and $mTECs^{high}$ ($MHC\text{-}II^{high}CD80^{high}$) subpopulations. The $mTECs^{low}$ have been proposed to function as precursors for mature $mTECs^{high}$ as they both share a linear developmental progress (Rossi, et al. 2007). Of note, the conversion of $mTECs^{low}$ to $mTECs^{high}$ can be regulated by certain TNF receptor superfamily (TNFRSF) signals such as RANK and CD40, the upregulation of the autoimmune regulator (Aire), and the expression of different peripheral tissue

antigens (PTAs) including insulin-2 and salivary protein-1 (Rossi, et al. 2007), (Anderson, et al. 2002), (Irla, et al. 2008, Hikosaka, et al. 2008, Akiyama, et al. 2008).

The murine thymic DCs contain many subpopulations that are different in their developmental origins, and perhaps in their functional contributions to thymic tolerance. The thymus is thought to involve thymic conventional DCs (cDCs) that can be classified into cDC1 cells ($CD8\alpha^+$ $Sirp\alpha^-$ $CD11b^-$) and cDC2 ($CD8\alpha^-$ $Sirp\alpha^+$ $CD11b^+$) cells, and plasmacytoid DCs (pDCs) that are characterised by the expression of PDCA-1 (Liu and Nussenzweig 2010). cDCs arise from pre-cDC precursors that evolve within the bone marrow and egress into the periphery before developing into cDC1 and cDC2 cells (Liu, et al. 2009). Whereas pDCs are generated directly from the common DC progenitor within the bone marrow. Both cDCs and pDCs have the expression of CD11c on their surfaces. After maturation in the peripheral tissues, cDC2 cells and pDCs are attracted to the thymus via the expression of CCR9 and CCR2 that bind to the chemokine ligands CCL25 and CCL8, respectively, which are produced by thymic stroma (Baba, Nakamoto and Mukaida 2009, Hadeiba, et al. 2012). By contrast, unlike cDC2, cDC1 cells are developed within the thymus from migrant immature pre-cDC progenitors that express CCR7 and are recruited to the thymus via the secretion of CCL21 from the mTECs. Also, these intrathymic DCs can be regulated by mTECs via the chemokine XCL-1 that is produced in an Aire-dependent manner (Schlitzer, et al. 2015, Lei, et al. 2011). Migration of cDC2 cells, pDCs and pre-cDC progenitors to the thymus must be tightly balanced as it may be essential for efficient thymic tolerance. It has been proposed that the exact location of DCs within the thymic medulla plays a crucial role in tolerance mechanisms. A study has confirmed that in both Aire-deficient and XCL-1 deficient mice intrathymic cDC1 cells were displaced and unsettled in the thymic medulla that lack the expression of Aire and XCL-1. This DC mispositioning was also followed by deficient self-tolerance and insufficient induction of Tregs. These findings suggest that mTECs via the expression of Aire and XCL-1 can direct cDC1 cells to migrate to the relevant medullary areas for optimal thymic tolerance and sufficient development of Tregs (Lei, et al. 2011).

2.1.5. The development of intra-thymic natural T regulatory cells (nTregs)

Despite the crucial role of the thymic medulla in the development of $Foxp3^+$ Tregs, the initial mechanisms that control commitment to the thymic Treg lineage remain controversial. However, two main subpopulations of $CD4^+$ SP thymocytes have been proposed to act as direct precursors of mouse $CD4^+CD25^+Foxp3^+$ Tregs, including $CD25^+Foxp3^-$ and $CD25^-Foxp3^+CD4^+$ SP thymocytes (Lio and Hsieh 2008, Tai, et al. 2013) (figure 2.4). $CD25^+Foxp3^-CD4^+$ thymocytes were found to induce $CD25^+Foxp3^+$ Tregs after being transferred intrathymically (Lio and Hsieh 2008). Whereas $CD25^-Foxp3^+CD4^+$ thymocytes were found to give rise to $CD25^+Foxp3^+CD4^+$ Tregs following intrathymic adoptive transfer (Tai, et al. 2013). These findings suggest that both subsets may act as Tregs precursors intrathymically. However how both subsets contribute to the newly developed intrathymic Tregs is still enigmatic, especially when two separate development pathways are involved in the process. However, a recent mouse model (Nr4a3-Tocky) has been developed by Bending et al. to study the dynamics of Tregs development and differentiation via TCR signalling within the thymus using a fluorescent timer protein. In this model, TCR signalling generates unstable chromophore that spontaneously shifts its emission from blue fluorescence to red fluorescence when it becomes maturely stable. Thus, upon receiving a TCR trigger cells emit only blue fluorescence and become $blue^+$ red^-. Then, with persistent TCR engagement cells shift the emission to become $blue^+$ red^+, whereas they express red fluorescence only ($blue^-$ red^+) when TCR signals are disrupted. This model has shown that $CD25^+Foxp3^-CD4^+$ SP thymocytes were mainly $blue^+$ red^-, while both $CD25^+Foxp3^+$ Tregs and $CD25^-Foxp3^+$ thymocytes were mostly $blue^+$ red^+. These findings suggested that $CD25^+Foxp3^-CD4^+$ SP thymocyte could act as main precursors for thymic Treg lineage as they received persistent TCR signals to generate $CD25^+Foxp3^+$ Tregs, compared to $CD25^-Foxp3^+$ thymocytes that found to be enriched with mature $Foxp3^+$ cells that have received less TCR signals. These findings confirm the important role of persistent TCR signals for generating intrathymic Tregs (Bending, et al. 2018). However, another mouse model (Rag2GFP) has reported that $CD25^-Foxp3^+CD4^+$ SP thymocytes might have the potential to play as precursors for intrathymic Tregs development (Cowan, McCarthy and Anderson

2016). Both mouse models support the hypothesis that intrathymic Tregs development may rise from two distinct precursor subsets.

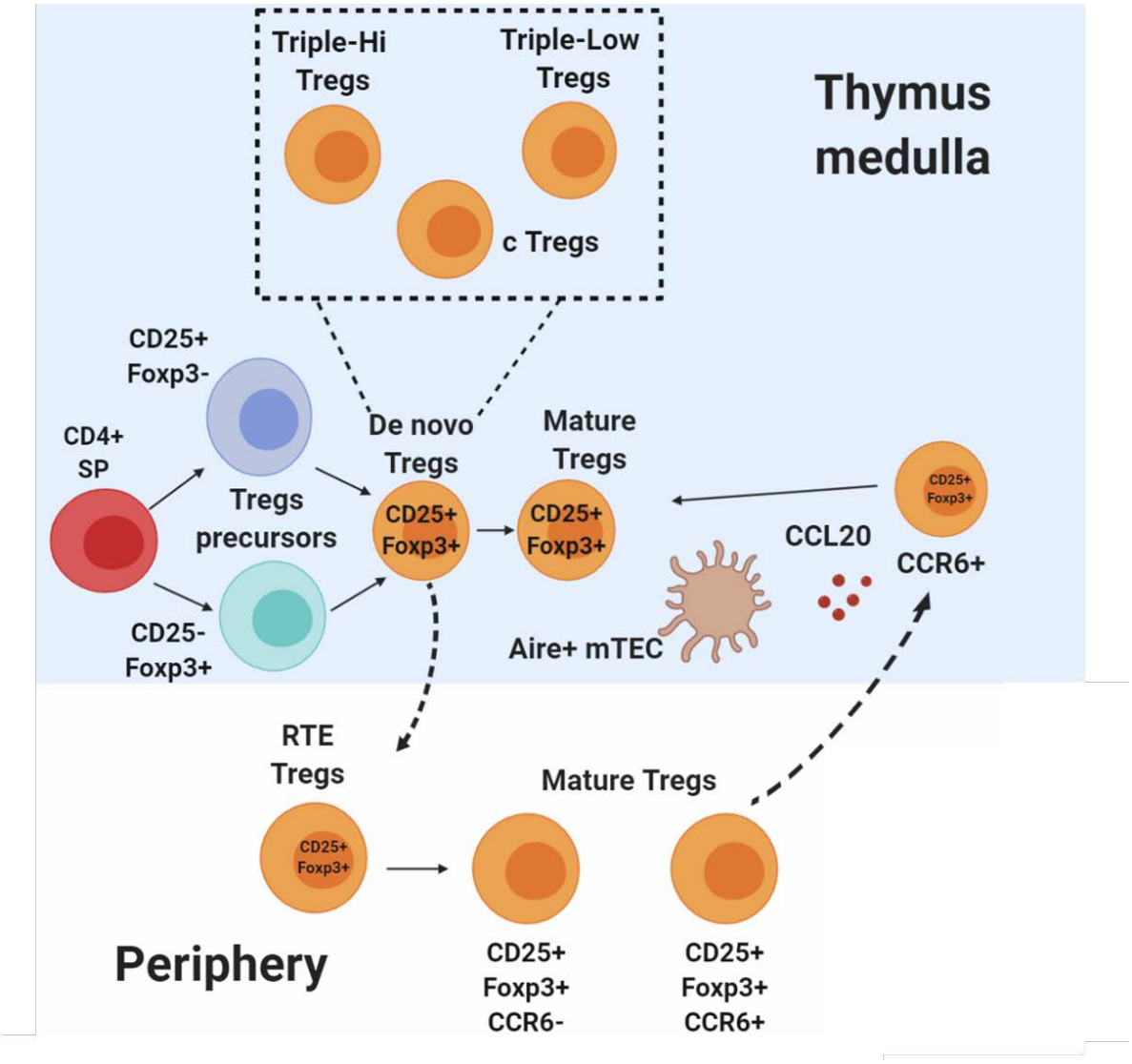


Figure 2.4. The development and recirculation of Foxp3$^+$ nTregs in the murine thymus. The thymic medulla provides a suitable microenvironment for the commitment and development of CD4$^+$CD25$^+$Foxp3$^+$ intrathymic Tregs from newly selected CD4$^+$ SP thymocytes. These CD4$^+$ SP thymocytes are found to give rise to CD25$^+$Foxp3$^-$ and CD25$^-$Foxp3$^+$ thymocytes which function as precursors for the development of *de novo* CD4$^+$CD25$^+$Foxp3$^+$ intrathymic Tregs. *De novo* Tregs have been reported to be a heterogeneous population containing Triple^{hi} (CD25^{hi} PD1^{hi} GITR^{hi}), Triple^{lo} (CD25^{lo} PD1^{lo} GITR^{lo}) and central Tregs (CD62L$^+$ CD44$^-$). *De novo* Tregs depart the thymus as RTE Tregs to the periphery where they join the mature peripheral Treg pool. Mature peripheral Tregs are found to acquire the expression of CCR6$^+$ during their maturation in the periphery, which enables them to re-enter the thymus again in response to the production of CCL20 from Aire$^+$ mTEC. SP: single positive, RTE: recent thymic emigrants, mTEC: medulla thymic epithelial cell. This figure is modified from (Inglesfield, et al. 2019).

The presence of two different precursors for intrathymic Treg-lineage commitment may highlight the ability of the thymic medulla of producing heterogeneous Treg subsets that are phenotypically and functionally distinct. Indeed, mature Tregs that are exported to peripheral tissues show distinct function with diverse phenotypic features especially in the expression of chemokine receptors and adhesion molecules. This phenotypic heterogeneity may boost their ability to maintain various immune responses across different body organs (Smigiel, et al. 2014, Wyss, et al. 2016). However, whether this heterogeneity is formed intrathymically during Tregs development or after leaving the thymus is obscure. Accordingly, two $Foxp3^+$ Treg populations have been identified in mice including central Tregs (cTregs) and effector Tregs (eTregs) that are $CD44^-CD62L^+$ and $CD44^+CD62L^-$, respectively. By using the Rag2GFP mouse model, cTregs have been found to be phenotypically similar to the newly developed intrathymic Tregs, proposing that cTregs might be initially developed in the thymus before travelling to peripheral tissues where they mature to eTregs (figure 2.4) (Wyss, et al. 2016). Similar results have been shown using adoptive transfer of sorted cTregs which displayed $CD44^+CD62L^-$, the phenotype of eTregs, following transfer in mice (Smigiel, et al. 2014). By contrast, another study has found two distinct subsets of intrathymic Tregs including triple high cells ($CD25^{hi}$ PD-1^{hi} $GITR^{hi}$) and triple low cells ($CD25^{lo}$ PD-1^{lo} $GITR^{lo}$), and both subsets have shown to involve newly developed intrathymic Tregs in Rag2GFP mice, supporting the hypothesis of the heterogeneity of Tregs development within the thymus (figure 2.4). Moreover, this study has shown that both triple high and low cells express different sequenced TCR repertoire and have different affinity for self-antigens, based on the strength of TCR signals using Nur77-GFP reporter expression. Triple high cells have shown higher expression of Nur77-GFP than triple low cells, confirming that they might receive stronger TCR signals than triple low cells. Also, triple cells have shown higher affinity for self-antigens than triple low, suggesting that the heterogeneity of Tregs development in the thymus might be based on the affinity of their TCR for self-antigens (Wyss, et al. 2016).

It has also been proposed that the heterogeneity of the *de novo* Tregs in the thymus may due to the presence of mature recirculated Tregs that have been found alongside their newly developed Tregs in the thymic medulla (figure 2.4). This has been widely demonstrated using Rag2GFP mice where thymocytes undergoing

intrathymic development can be characterised as GFP^+, whereas fully mature T cells are GFP^-. In these mice, the newly developed intrathymic Tregs are found to be densely contaminated by GFP^- Tregs, suggesting that GFP^- Tregs might recirculate back to the thymus or be recruited to the thymus from the periphery (Cowan, et al. 2016, E. Yang, et al. 2014, Thiault, et al. 2015). Studies have been carried out to understand the mechanisms that regulate the process of thymic recirculation and its possible effect on thymus function. A study by Thiault et al. has shown that the migration of mature peripheral Tregs to the thymus might be mediated by the expression of CXCR4 that is expressed by GFP^- Tregs in the thymus. Also, the number of Tregs homing to the thymus is reduced about 50% following injection of Tregs into mice treated with a CXCR4 inhibitor (Thiault, et al. 2015). However, another study has disproved this finding by using $Cd4^{cre}/Cxcr4^{floxed}$ mice where the number of intrathymic Tregs is not affected (Lucas, et al. 2017). A recent study has shown that the intrathymic recirculating Tregs are decreased in $Ccr6^{-/-}$Rag2GFP mice compared to wild-type controls. Also, it is found that the newly developed intrathymic Tregs are $CCR6^-CCR7^+GFP^+$, whereas mature Tregs show $CCR6^+CCR7^-GFP^-$, proposing that mature thymic Tregs acquire the expression of CCR6 in the periphery, which is a key regulator for Tregs recirculation to thymus. These findings highlight the possibility of involvement of the thymic medulla in Tregs recirculation mechanism as $CCR6^+$ Tregs may be attracted to the thymus via the secretion of CCL20 by mTECs. Relevant to this, it has been found that mTECs from $Aire^{-/-}$ mice show a significant reduction in the expression of CCL20 compared to the wild-type controls, and then a reduction in the GFP^- recirculating Tregs, proposing the role of AIRE-derived mTECs in the process of thymic Tregs recirculation via CCR6-CCL20 signalling (Cowan, et al. 2018) (figure 2.4). This finding also supports the role of AIRE-derived mTECs in controlling and directing the thymic DCs within the medulla via XCL-1 signalling as mentioned above (Lei, et al. 2011).

Recirculating Tregs have been found to inhibit the development of intrathymic Tregs, particularly when mice become old; as older mice show a reduction in thymic Treg pool and an increase in the number of mature recirculating Tregs within the thymus compared to younger (Thiault, et al. 2015). However, other experiments on osteoprotegerin (OPG)-deficient mice where recirculating Tregs are increased has not shown any reduction in the development of intrathymic Tregs compared to

wild-types controls (McCarthy, et al. 2015). Taken together, there are still many unanswered questions such as to what extent does the thymus need recirculating Tregs? Do thymus recirculating Tregs need to travel back to the periphery again? How do recirculating Tregs mediate their function in the thymus? Which peripheral tissues do recirculating Tregs come from?.

2.1.6. The development of peripherally-induced T regulatory cells

In addition to natural or intrathymically derived Tregs, another population of regulatory cells has been identified to maintain immune homeostasis in the periphery. Unlike natural Tregs, peripherally-induced T regulatory cells (iTregs) are immunosuppressive cells that differentiate from naïve $CD4^+$ T cells in peripheral tissues following exposing to antigen stimulation in the presence of certain cytokines including interleukin-2 (IL-2), transforming growth factor (TGF-β1) and interleukin-10 (IL-10) (Abul 2003). Different abbreviations have been used to describe induced Tregs including "iTreg" when cells are generated *in vitro,* or "pTregs" or adaptive Tregs when generated *in vivo*. Despite pTregs constituting only a small percentage of Tregs as a whole, these populations are found highly enriched in particular organs including maternal placenta and the gut (Mizrahi and Ilan 2009). Therefore, pTregs are thought to sustain immune tolerance against commensal bacteria, foods, allergens and the fetus in the uterine (Samstein, et al. 2012, Luu, Steinhoff and Visekruna 2017).

In humans, the development of pTregs is found to occur mainly in the intestine where mucosal DCs and commensal bacteria are enriched. Mucosal DCs, particularly $CD103^+$ lamina propria DCs, can secrete TGF-β1 that is known as a master regulator of pTregs, and retinoic acid (RA) that shows a potential of induction of pTregs from naïve $CD4^+$ T cells (Darrasse-Jeze, et al. 2009, Yamazaki, et al. 2007). Also, commensal bacteria including *Bacteroides fragilis* and *Clostridia* are found to promote the production of pTregs in the gut in a process to protect the host against inflammatory bowel diseases such as ulcerative colitis. *B. Fragilis* have shown to directly induce pTregs by triggering Toll-like receptor 2 (TLR2) that is expressed on activated T cells via the expression of capsular polysaccharide A of *B. Fragilis*. *B. Fragilis* – induced Tregs are found to secrete IL-10, an anti-inflammatory cytokine, and protect the host against colitis (Round and Mazmanian 2010). A specific

pathogen-free (SPF) mice model has been proposed by Atarashi et al. to study the impact of indigenous *Clostridium* species on intestinal $CD4^+$ T cells. *Clostridium* bacteria, especially cluster XIVa and IV, are found predominantly enriched in the colonic mucosa and work to maintain immune homeostasis by secreting antigens which attract intraepithelial lymphocytes and IgA^+ cells. These antigens are directly presented by $CD103^+$ DCs which then activate naïve $CD4^+$ T cells to become either pTregs or effector $CD4^+$ T cells based on the physiological condition (Atarashi, et al. 2011) . $CD103^+$ DCs play an essential role in manipulating the immune responses, for examples during steady state, DCs produce TGF-β1 and RA regulatory molecules to induce pTregs, whereas during inflammatory conditions DCs secrete inflammatory molecules such as IL-23 to initiate the inflammation response by promoting differentiation of effector $CD4^+$ T cells (Lathrop, et al. 2011).

The development of pTregs/iTregs is mainly mediated by TGF-β1 signalling that promotes the induction of Foxp3 expression. The main signalling pathway initiated by TGF-β1 is the activation and phosphorylation of its downstream transcription factors "Smads"; mainly Smad2 and Smad3. Smad2/3 are key mediators for the induction of Foxp3. Once activated, Smad2/3 are recruited into the conserved non-coding sequence 1 (CNS1) region in the Foxp3 gene locus (Xu, Kitani and Strober 2010). IL-2 is also found to be an essential factor for the induction of Foxp3 expression, which stimulates transcription factor STAT5 to bind the Foxp3 promoter and CNS2 region which is the major Treg-specific demethylated region (TSDR) (Kanamori, et al. 2016).

pTregs are thought to have a different TCR repertoire compared to nTregs as they develop in various conditions in the periphery. pTregs, unlike nTregs, are found to be more specific for foreign or non-self antigens, given that they are differentiated from naïve conventional T cells (Haribhai, et al. 2011). Although both pTregs and nTregs are biased for different TCR-antigen affinities, it has been found that both subsets share minimal overlap between their TCR repertoire, suggesting that this shared overlap may be required for both subsets in the resolution of autoimmune diseases as the combination shows a more diverse TCR repertoire (Schmitt and Williams 2013). Moreover, another study has shown that pTregs derived from the intestinal mucosa had a distinct TCR repertoire compared with Tregs from other

location, confirming that pTregs are originally developed outside the thymus (Lathrop, et al. 2011). In the periphery, various types of antigen-presenting cells could generate different peptide-MHC complexes by processing and displaying tissue-derived proteins/peptides, and then trigger TCR on T cells. These interactions have been found to induce Foxp3$^+$ T cells with a broader TCR repertoire that could recognise a diverse pool of antigens (Yadav, Bluestone and Stephan 2013). By contrast, other studies on TCR repertoire have demonstrated that the generation of pTregs in non-mucosal tissues such as the central nervous system and the pancreas might be restricted as both T conventional cells and Tregs found to share a minimal overlap in TCR repertoire, suggesting that Tregs could be recruited to rather than induced at these locations (Liu, et al. 2009, Wong, Mathis and Benoist 2007). These findings imply that tumours developed in mucosal cells may be more enriched with peripheral or induced Tregs.

2.1.7. Aim of this study

The aim of this chapter is to establish a mice model for generating CD4$^+$CD25$^+$Foxp3$^+$ induced Tregs from purely sorted naïve CD4$^+$CD25$^-$ Foxp3$^-$ T cells in the presence and absence of tumour cells *in vitro*. The objectives of this chapter are:

1- Selection of an appropriate model to generate and study iTregs.

2- Purification of naïve CD4$^+$CD25$^-$Foxp3$^-$ T cells and CD4$^+$CD25$^+$Foxp3$^+$ nTregs using beads and cell sorting based technology.

3- Optimisation of Tregs induction parameters including CD3/28 TCR activation signal, the concentration of TGF-β1 and the duration of induction.

4- Validation of the induction of Tregs from naïve CD4$^+$CD25$^-$ T cells using flow cytometry, Foxp3-TSDR methylation status and functional assay.

5- Scaling up the model for obtaining sufficient cell number for MS-based profiling.

2.2. Materials and Methods

2.2.1. Mice and Cell lines

Female naïve BALB/c and C57 mice, 8-12 week of age, were brought from Charles Rivers laboratories (UK) and maintained in pathogen-free environment at the animal unit of the John van Geest Cancer Research Centre in accordance with established guidelines of the Animal Care and Use of Nottingham Trent University. 4T1 cells (ATCC® CRL-2539™), a highly metastatic epithelial breast cancer cell line derived from BALB/c mice, were cultured in complete 4T1 cell medium (RPMI 1640 medium supplemented with 10% FCS, 10 mM HEPES, 2 mM L-glutamine, 1 mM sodium pyruvate and 2250 mg/500 mL glucose).

2.2.2. Isolation of $CD4^+CD25^+$ and $CD4^+CD25^-$ T cells and Cell sorting

$CD4^+CD25^+$ T cells were purified from mouse spleens using a mouse $CD4^+CD25^+$ T Regulatory Cell Isolation Kit (Miltenyi Biotec, Cat no: 130-091-041) according to the manufacturer's instructions.

$CD4^+CD25^-$ T cells were purified from mouse spleens using two different ways. **First,** by using the mouse $CD4^+CD25^+$ T Regulatory Cell Isolation Kit (mentioned above), $CD4^+CD25^-$ T cells were obtained from the *flow-through* tube during positive selection of $CD4^+CD25^+$ cells by MS columns. **Second,** naïve $CD4^+$ T cells were purified using Dynabeads™ Untouched™ Mouse CD4 Cells Kit (Invitrogen, Thermo Fisher Scientific, Cat no: 11415D) according to manufacturer's instructions. The isolated $CD4^+$ T cells were stained with mouse anti-CD4 monoclonal antibody (eFluor 450, Clone GK1.5, eBioscience™, Thermo Fisher Scientific, Cat no: 48-0041-82) and mouse anti-CD25 monoclonal antibody (PE, Clone PC61.5, eBioscience™, Thermo Fisher Scientific, Cat no: 12-0251-82). The stained cells were then sorted using a Beckman Coulter MoFlo XDP cell sorter (sorting mode: purify, the purity of sorting was ≥ 98%). The strategy for sorting a pure population of naïve $CD4^+CD25^-Foxp3^-$ T cells and naïve $CD4^+CD25^+Foxp3^+$ nTregs is discussed in detail in the chapter 3 (section 3.2.2.1).

2.2.3. Flow Cytometry

Single-cell suspensions of splenocytes, purified, sorted and cultured T-cells were prepared using 40µm cell strainers (Greiner Bio-one). Red blood cells were lysed using RBC lysis buffer 1X (BD Pharm Lyse™, BD Biosciences, Cat no: 555899). Cell counting and viability were assessed following staining with solution 18 (Chemometec, Cat no: 910-3018), which contains Acridine Orange (AO, 80 µg/mL) for staining total living and dead cells, and 4′,6-diamidino-2-phenylindole (DAPI, 40 µg/mL) for counterstaining only dead cells, using NucleoCounter cell counter (NucleoCounter® NC-250™, Chemometec). For flow cytometry, cells were washed twice with DPBS buffer, centrifuged at 350 x g for 5 min at room temperature and stained with Fluorochrome-conjugated monoclonal antibodies (mAbs) (table 2.1) and analysed using a flow cytometer (Gallios, 10 channels, Beckman Coulter). Details about the Gallios flow cytometer are summarised in table 2.2. To prevent non-specific binding, mouse FCR blocking reagent (purified anti-mouse CD16/32 antibody, BioLegend, Cat no: 101302) was used before staining with Fluorochrome-conjugated antibodies. LIVE/DEAD™ Fixable Yellow Dead Cell Stain Kit (405 nm excitation, Thermo Fisher Scientific) was used to detect the viability of cells prior to intracellular staining.

Table 2.1: List of antibodies that were used for flow cytometry analysis.

Channel	mAbs	Fluorochrome	Clone	Company	Cat no.
FL1	CD73	FITC	TY/11.8	BioLegend	127219
FL2	CD25	PE	PC61.5	ThermoFisher	12-0251-82
FL3	Tim-3	PE/Dazzle	B8.2C12	BioLegend	134013
FL3	CTLA-4	PE/Dazzle	UC10-4B9	BioLegend	106317
FL5	CD39	PE/Cy7	Duha59	BioLegend	143806
FL5	LAG-3	PE/Cy7	C9B7W	BioLegend	125225
FL5	GITR	PE/Cy7	DTA-1	BioLegend	126317
FL6	FOXP3	eFluor 660	FJK-16s	ThermoFisher	50-5773-82
FL8	CD127	APC-eFluor 780	A7R34	ThermoFisher	47-1271-80
FL9	CD4	eFluor 450	GK1.5	ThermoFisher	48-0041-82
FL10	DEAD/LIVE	Yellow 405		ThermoFisher	L34959

Table 2.2: List of lasers and filter that are available on the Gallios flow cytometer.

Laser	Channel	Filters	Fluorochromes
Blue Laser 488nm	FL-1	525/40	FITC
	FL-2	575/30	PE
	FL-3	620/30	ECD, PE Tx Red, PE/Dazzle
	FL-4	695/30	PerCP-Cy5.5
	FL-5	755LP	PE-Cy7
Red Laser 638nm	FL-6	660/20	EFluor 660, APC
	FL-7	725/20	Alexa-Fluor 700
	FL-8	755LP	APC-eFluor 780, APC-Cy7
Violet Laser 405nm	FL-9	450/40	Pacific Blue
	FL-10	550/40	Krome Orange

2.2.4. Generation of induced T regulatory cells *in vitro*

Naïve $CD4^+CD25^-$ T cells were purified and sorted (as mentioned in section 2.2.2) and rested overnight in complete T cell medium (RPMI 1640 supplemented with 10% FCS, 1% penicillin/streptomycin, 20 mM HEPES buffer, 50 µM β-mercaptoethanol, 2 mM L-glutamine) at 37^0C, 5% CO_2. Then, a 24-well culture plate was prepared by adding 900 µL of complete T cell medium in each well. Hanging cell culture inserts for 24 well plate, PET (polyethylene terephthalate) membrane bottom, transparent, pore size 0.4 µm (SARSTEDT, Cat no: 83.3932.041) were placed in each well of the 24-well plate. After that, $2.5x10^5$ of sorted naïve $CD4^+CD25^-$ T cells resuspended in 200 µL of complete T cell medium were transferred into each of hanging cell culture inserts. Then, cells were activated using Dynabeads™ Mouse T-Activator CD3/CD28 for T-Cell Expansion and Activation Kit (Invitrogen, Thermo Fisher Scientific, Cat no: 11452D) according to the manufacturer's instruction at ratio 1:1 (naïve $CD4^+CD25^-$ T cells: CD3/38 beads). The cells were grown in the presence of IL-2 (30 U/mL) (Recombinant Murine IL-2, PeproTech, Cat no: 212-12), and TGF-β1 (5 ng/mL) (Mouse TGF beta 1 Recombinant

Protein, Thermo Fisher Scientific, Cat no: 14-8342-80) for 5 days at 37^{0}C, 5% CO_2. The overall steps of the iTregs genertion *in vitro* are shown in figure 2.5.

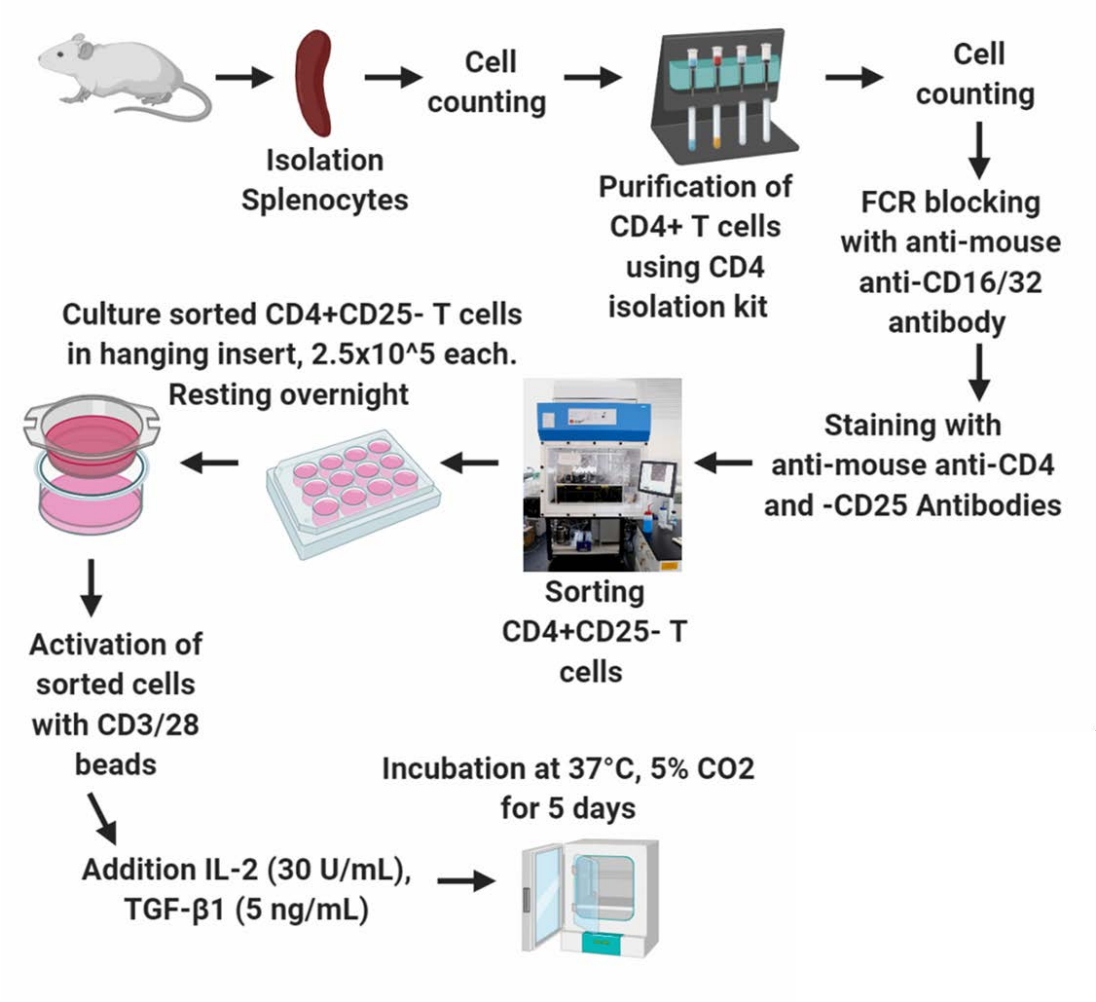


Figure 2.5. A diagrammatic representation showing the process of iTregs generation *in vitro*.

2.2.5. Generation of induced T regulatory cells *in vitro* in the presence of tumour cells (contactless co-culture)

Murine breast cancer tumour cells (4T1) resuspended in 900 µL of the complete 4T1 cell medium (mentioned in section 2.2.1) were seeded in a 24-well plate and incubated overnight at 37^{0}C, 5% CO_2. Hanging cell culture inserts were placed over

already seeded tumour cells. Subsequently, $2.5x10^5$ of sorted naïve $CD4^+CD25^-$ T cells resuspended in 200 µL of complete T cell medium were transferred into hanging inserts to be separately co-cultured with tumour cells at ratio 1:20 (tumour cells: $CD4^+CD25^-$ T cells). The sorted naïve $CD4^+CD25^-$ T cells were activated using CD3/28 beads and cultured in the presence of IL-2 (30 U/mL) and TGF-β1 (5 ng/mL) for 5 days at 37^0C, 5% CO_2.

2.2.6. Intracellular (nuclear) staining of Foxp3

For intracellular nuclear and cytoplasmic staining, eBioscience™ Foxp3 / Transcription Factor Staining Buffer Set (Thermo Fisher Scientific, Cat no: 00-5523-00) was used. For nuclear staining, Fixation/Permeabilisation (F/P) solution was prepared by mixing one part of F/P concentrate with three parts of F/P diluent. Then, 1x solution of Permeabilisation buffer (PB) was prepared by mixing one part of 10x PB with nine parts of nano-pure distilled water. After staining with cell surface antibodies and live/dead stain, cells were fixed by F/P solution (1 mL for $1x10^6$ cells) in the dark at room temperature for 40 min. After the incubation, the cells were washed by DPBS (2 mL for $1x10^6$) and centrifuged at 350 x g for 5 min. The cells were resuspended in 200 µL of PB buffer and stained with Foxp3 mAb for 30 min in the dark at room temperature. For intracellular cytoplasmic staining, the same protocol of nuclear staining was followed. However, cells were fixed by F/P concentrate only (500 µL/$1x10^6$ cells) without mixing with F/P diluent.

2.2.7. Establishment of functional assay

Induced T regulatory cells were generated as shown in figure 2.5. At day 5 of iTregs generation, naïve $CD8^+$ T cells were purified using Dynabeads™ Untouched™ Mouse CD8 Cells Kit (Invitrogen™, Thermo Fisher Scientific, Cat no: 11417D) according to the manufacturer's instructions from the spleen. The purified $CD8^+$ T cells were washed twice with DPBS and stained with anti-mouse CD8 mAb (Alexa Fluor® 700, clone: 53-6.7, BioLegend, Cat no: 100730) for cell sorting using purify mode as mentioned above. Sorted $CD8^+$ T cells were rested for 4 h and stained again with CD8 mAb to assess the purity of cell sorting which was ≥ 98% using the flow cytometer. Then induced $CD4^+CD25^-$ T cells were harvested, washed and stained with CD4, CD25 and Foxp3 mAbs to assess the percentage of induction of Foxp3 after 5 days of culture using flow cytometer. After that, sorted $CD8^+$ T cells were

resuspended in DPBS supplemented with 0.1% BSA (1 mL/1x10^6 $CD8^+$ T cells). While vortexing, 1 µL of 5 µM CFSE (carboxyfluorescein succinimidyl ester) dye was added into the cell suspension at (1:1000) dilution using CellTrace™ CFSE Cell Proliferation Kit, for flow cytometry (Invitrogen™, Thermo Fisher Scientific, Cat no: C34554). Then cells were incubated for 10 min in the dark at 37°C. after that, while vortexing, cells were washed by adding 10 mL of pre-warmed complete T cell medium and incubated for 5 min in the dark at room temperature. After incubation cells were centrifuged at 350 x g for 8 min at room temperature. Washing step was repeated twice to ensure removing extra CFSE dye in the solution. Then cells were centrifuged, counted and seeded into a 96-well plate, rounded bottom. After that, CFSE-stained $CD8^+$ T cells were activated using Dynabeads™ Mouse T-Activator CD3/CD28 for T-Cell Expansion and Activation Kit in the presence of IL-2 (30 U/mL). Then, induced $CD4^+CD25^+Foxp3^+$ T cells (iTregs) were co-cultured with CFSE-stained $CD8^+$ T cells at six different ratios including 1:2, 1:1, 2:1, 4:1, 8:1, 16:1 ($CD8^+$ T cells : iTregs), in addition to controls that contain $CD8^+$ T cells cultured without iTregs, for 4 days at 37°C, 5% CO_2. Cells were co-cultured in 96-well plate in 200 µL of complete T cells medium as a total volume. At day 4, co-cultured cells were harvested, washed twice. CD3/28 beads were magnetically removed. Then cells were stained with Dead/Live stain and CD8 mAb, and then analysed on flow cytometer for assessing the proliferation of $CD8^+$ T cells in the presence and absence of iTregs. The steps of establishment the functional assay are shown in figure 2.6.

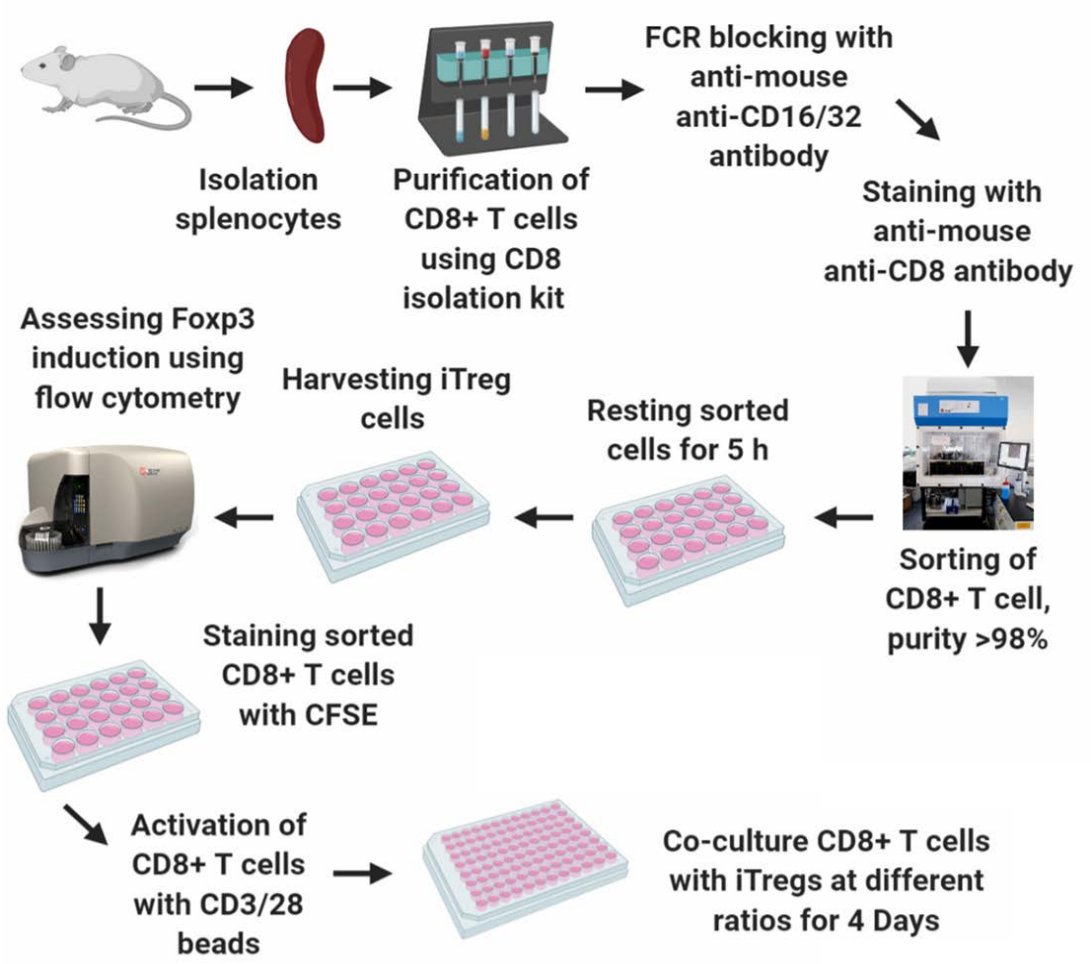


Figure 2.6. A diagramatic representation showing the steps of establishment of functional assay for assessing the suppressive activity of iTregs.

2.2.8. The stability of Foxp3 expression of induced Tregs

Induced Tregs were generated as shown in figure 2.5 for 5 days. At day 5, total $CD4^+$ Cells, which contain both activated $CD4^+CD25^-$ T cells and $CD4^+CD25^+Foxp3^+$ iTregs, were harvested and washed twice with DPBS. CD3/28 beads were magnetically removed. The Cells were centrifuged at 350 x g for 5 min and then resuspended in a fresh complete T cell medium. The cells were counted and $1x10^6$ of cells were stained with Live/Dead dye, anti-CD4, -CD25 and -Foxp3 mAbs to assess the percentage of induction of Foxp3 expression. After that, the remaining cells were cultured again in a fresh T cell medium in the presence of IL-2 (30 U/mL) for another 5 days at 37°C, 5% CO_2. The expression of Foxp3 of iTregs was assessed at day 3 (day 8 of Fox3 induction) and day 5 (day 10 of Foxp3 induction) of the second culture period.

2.2.9. Foxp3-TSDR methylation assay

The methylation status of Foxp3-TSDR region was assessed in 6 different subpopulation of $CD4^+$ T cells including naïve $CD4^+CD25^-$ T cells, activated $CD4^+CD25^-$ T cells, TGF-β1-drived $CD4^+CD25^+Foxp3^+$ iTregs, TGF-β1/4T1-drived $CD4^+CD25^+Foxp3^+$ iTregs, naïve $CD4^+CD25^+Foxp3^+$ nTregs and activated $CD4^+CD25^+Foxp3^+$ nTregs. Naïve $CD4^+CD25^-$ T cells were purely sorted as mentioned in section 2.2.2. The cells were activated using CD3/CD28 beads at ratio 1:1 (naïve $CD4^+CD25^-$ T cells: CD3/28 beads) in the presence of IL-2 (30 U/mL) for 5 days at 37°C, 5% CO_2. TGF-β1-drived $CD4^+CD25^+Foxp3^+$ iTregs were generated as mentioned in section 2.2.4. TGF-β1/4T1-drived $CD4^+CD25^+Foxp3^+$ iTregs were generated as mentioned in section 2.2.5. naïve $CD4^+CD25^+Foxp3^+$ nTregs were purely sorted using the cell sorter as mentioned in section 2.2.2. The sorted naïve nTregs were then activated using CD3/28 beads as mentioned above. The generated iTregs were purely sorted following the sorting strategy that is mentioned in the chapter 3 (section 3.2.2.1). $1x10^6$ cells from each of the 6 subpopulations were carefully washed twice with DPBS, centrifuged at 350 x g for 5 min. a cell pellet from each of the subpopulations was freezed at -80°C and shipped to EpigenDX company (MA, USA) for assessing the methylation status of Foxp3-TSDR.

The mouse Foxp3 methylation assay (Assay ID: ADS568-FS2) was carried out by EpigenDX, MA, USA. This assay was optimised to assess the methylation status of 4 different CpGs within TSDR region, intron 1, location (-2238 to -2207) from ATG, and +4442 to +4473) from TSS. Table 2.3 shows the list of CpG regions that were assessed during the assay.

Table 2.3: List of CpG regions that were examined to assess the methylation status of mouse Foxp3 interon1-TSDR region.

Assay ID	ADS568-FS2 (Intron 1-TSDR Region, Foxp3, Mouse)			
CpG#	CpG# -22	CpG# -21	CpG# -20	CpG# -19
From ATG	-2238	-2219	-2215	-2205
From TSS	+4442	+4461	+4465	+4473

2.3. Results

2.3.1. Establishment of a murine model to generate induced T regulatory cells from naïve CD4⁺CD25⁻Foxp3⁻ T cells

To develop a representative murine model for studying the difference between nTregs and iTregs, two most appropriate mouse strains were selected based on the availability of commercial mouse tumour cell lines that are compatible with the mouse strain. The percentage of nTregs in the spleen of BALB/c and C57BL/6 mice was assessed. The results showed that the percentage of naïve CD4⁺CD25⁺Foxp3⁺ nTregs from total naïve CD4⁺ T cells was significantly higher in the spleen of BALB/c (mean=12%) than C57BL/6 mice (mean=2.7%) (Figure 2.7).

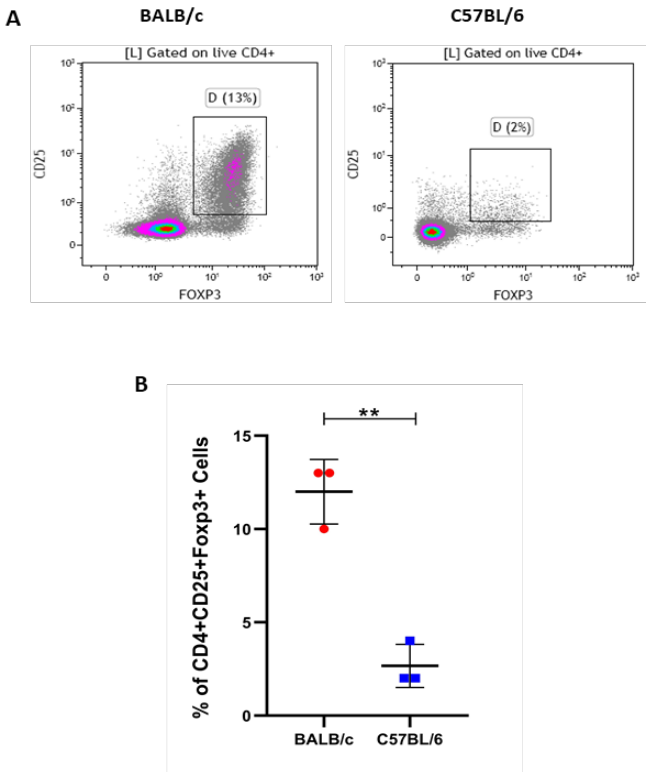


Figure 2.7. The percentage of naive CD4⁺CD25⁺Foxp3⁺ nTregs from total naive CD4⁺ T cells in the spleens of BALB/c and C57BL/6 mice. **(A)** Flow cytometric density plots showing the expression of CD25⁺Foxp3⁺ (gate D) in naïve nTregs of both BALB/c and C57BL/6. These plots were gated on live CD4⁺ T cells. **(B)** Scatter plots showing the mean of the percentage of naïve CD4⁺CD25⁺Foxp3⁺ T cells in the spleens of BALB/c (12%) and C57BL/6 (2.7%). The results were statistically significant (p>0.0015). Three independent experiments were carried out, a spleen from each of mouse strains was analysed in each experiment, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Unpaired t test was used using GraphPad Prism 8.0.1 software. Kalusa software (version 1.3) was used for analysing flow cytometric data.

According to the results shown in figure 2.7, the BALB/c mice strain was selected to develop the model.

Then it was decided to deplete naïve $CD25^+Foxp3^+$ Tregs within naïve $CD4^+$ T cells population to obtain pure naïve $CD4^+CD25^-Foxp3^-$ T cells as a control population for the model. To achieve this, it was assessed whether all naïve $CD25^+$ T cells express positively Foxp3; since depletion of nTregs based on the expression of Foxp3 was not suitable in this model; as Foxp3 is mainly expressed in the nucleus. Therefore, the depletion of $CD4^+CD25^+Foxp3^+$ nTregs was based on the co-expression of CD4 and CD25 as both are surface markers. The mouse $CD4^+CD25^+$ T Regulatory Cell Isolation Kit (Miltenyi Biotec) was used as it could isolate both $CD4^+CD25^+$ T cells and $CD4^+CD25^-$ T cells simultaneously. The results of isolation confirmed that 92% of $CD25^+$ T cells were positive for Foxp3 expression, confirming that CD25 expression could be useful for depleting $CD25^+Foxp3^+$ T cells in this model (figure 2.8). The results also confirmed the efficiency of that kit for isolating naïve nTregs from purified $CD4^+CD127^-$ T cells.

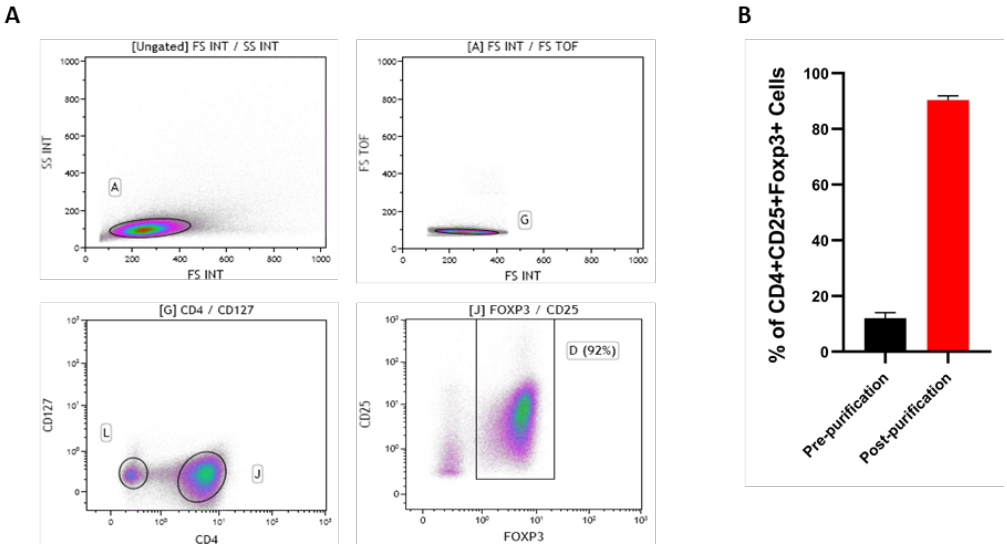


Figure 2.8. The percentage of $Foxp3^+$ cells within the naïve $CD4^+CD25^+$ T cells population after isolating naïve $CD4^+CD25^+$ T cells from purified naïve $CD4^+$ T cells using the Miltenyi kit. **(A)** Flow cytometric density plots showing the percentage of $CD25^+$ T cells that positively express Foxp3. The plots also show the gating strategy for analysing flow cytometric data; $Foxp3^+CD25^+$ cells (gate D) were gated on $CD4^+CD127^-$ cells (gate J) that were gated on singlets cells (gate G) of total stained cells. **(B)** Column bar graph displaying the effectiveness of the Miltenyi kit in isolating nTregs from purified $CD4^+$ T cells; the isolation purity was ≥ 90%. Three independent experiments were carried out, a spleen from each of mouse strains was analysed in each experiment, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. GraphPad Prism was used for drawing bar graph. Kalusa software (version 1.3) was used for analysing flow cytometric data.

The same kit was used to deplete $CD4^+CD25^+$ T cells nTregs for obtaining pure $CD4^+CD25^-Foxp3^-$ T cells as a control population to generate iTregs. Accordingly, naïve nTregs were depleted, and $CD4^+CD25^-$ T cells were stained with CD4, CD25 and Foxp3 mAbs as well as live/dead staining to assess the purity of depletion. Then $CD4^+CD25^-$ T cells were cultured for 3 days in the presence and absence of IL-2 to confirm the purity of the control population in this model. Surprisingly, the results showed that $CD4^+CD25^-$ T cells (control population) were positive for CD25 expression (38%) after 3 days of culture, whereas they showed a negativity for CD25 expression (0%) on day 0 of culture after depletion in the absence of IL-2, confirming that the $CD25^+$ T cells were not originally depleted. To confirm these results, the same experiment was repeated and a different conjugated anti-CD25 mAbs (PerCP-Cy5.5-conjugated) was used to assess the expression of CD25 after depletion; because PE-conjugated anti-CD25 antibody and anti-PE microbeads were included in the kit. The results of staining showed that $CD25^+$ T cells were completely depleted following depletion using both CD25 mAbs, however after 3 days of culture 30-42% of CD25 depleted cells were positive for CD25 expression (figure 2.9).

As the purity of depletion of $CD25^+$ T cells using the Miltenyi kit was ~ 62%, it was suggested to use cell sorting for obtaining pure $CD4^+CD25^-$ T cells. The results of cell sorting showed that $CD4^+CD25^-$ T cells were 100% negative for the expression of CD25 following cell sorting at day 0 and after 5 days of culture (figure 2.10). As a result, cell sorting using purity mode was selected to be used in this model.

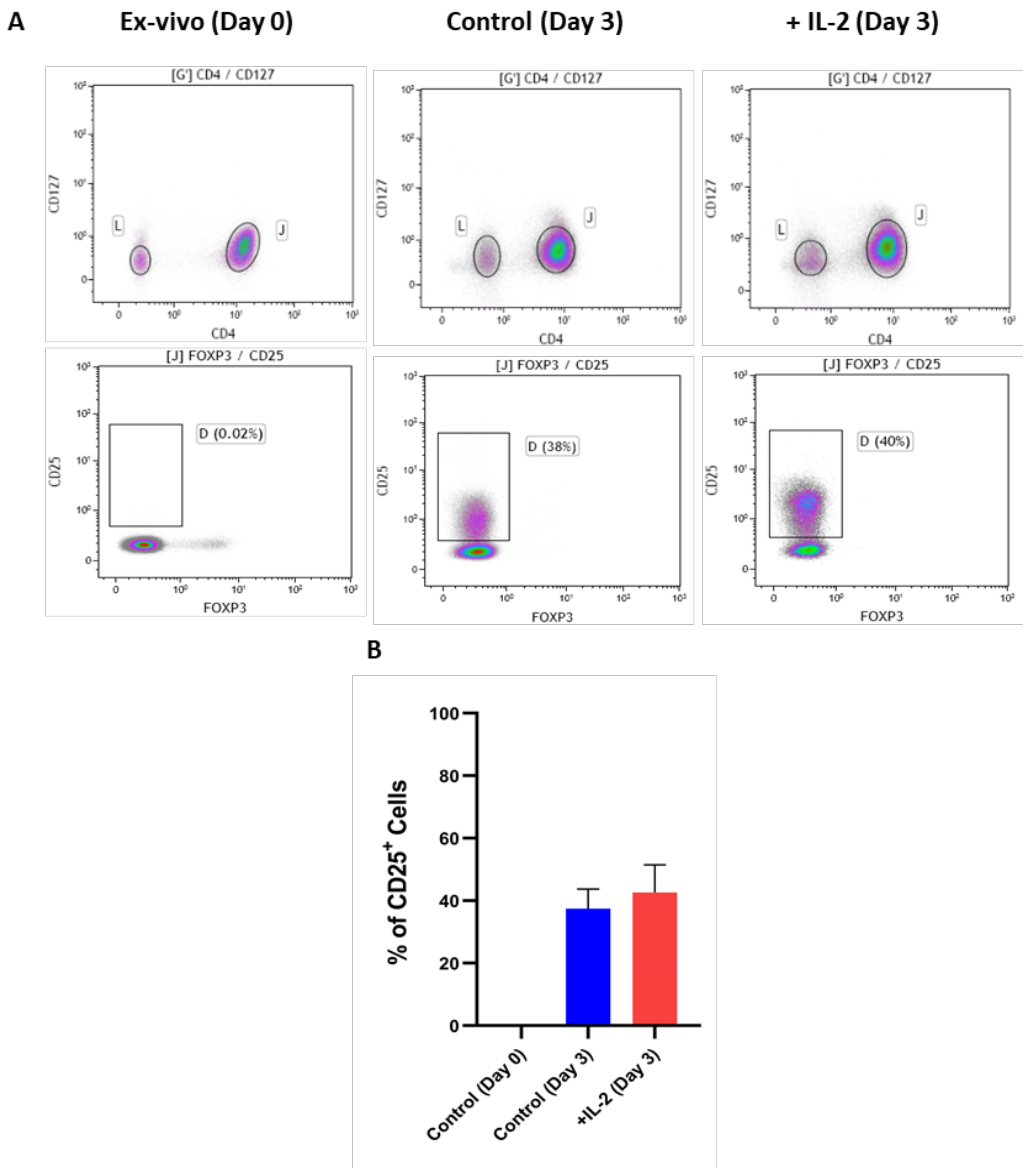


Figure 2.9. The percentage of CD25⁺ T cells following depletion of CD25 expressing cells using the Miltenyi kit. **(A)** Flow cytometric density plots showing the staining of CD25 (gate D) following depletion of CD25⁺ T cells *ex-vivo* at day 0 and after 3 days of culture in the presence and absence of IL-2. Gate D was gated on CD4⁺CD127⁻ T cells (gate J) which was gated on live CD4⁺ T cells. **(B)** Column bar graph showing the percentage of CD25⁺ T cells after their depletion ex-vivo on day 0 and day 3 of culture in the presence and absence of IL-2 (30 U/mL). Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve CD4⁺CD25⁻ T cells were cultured in the complete T cell medium in the absence of IL-2.

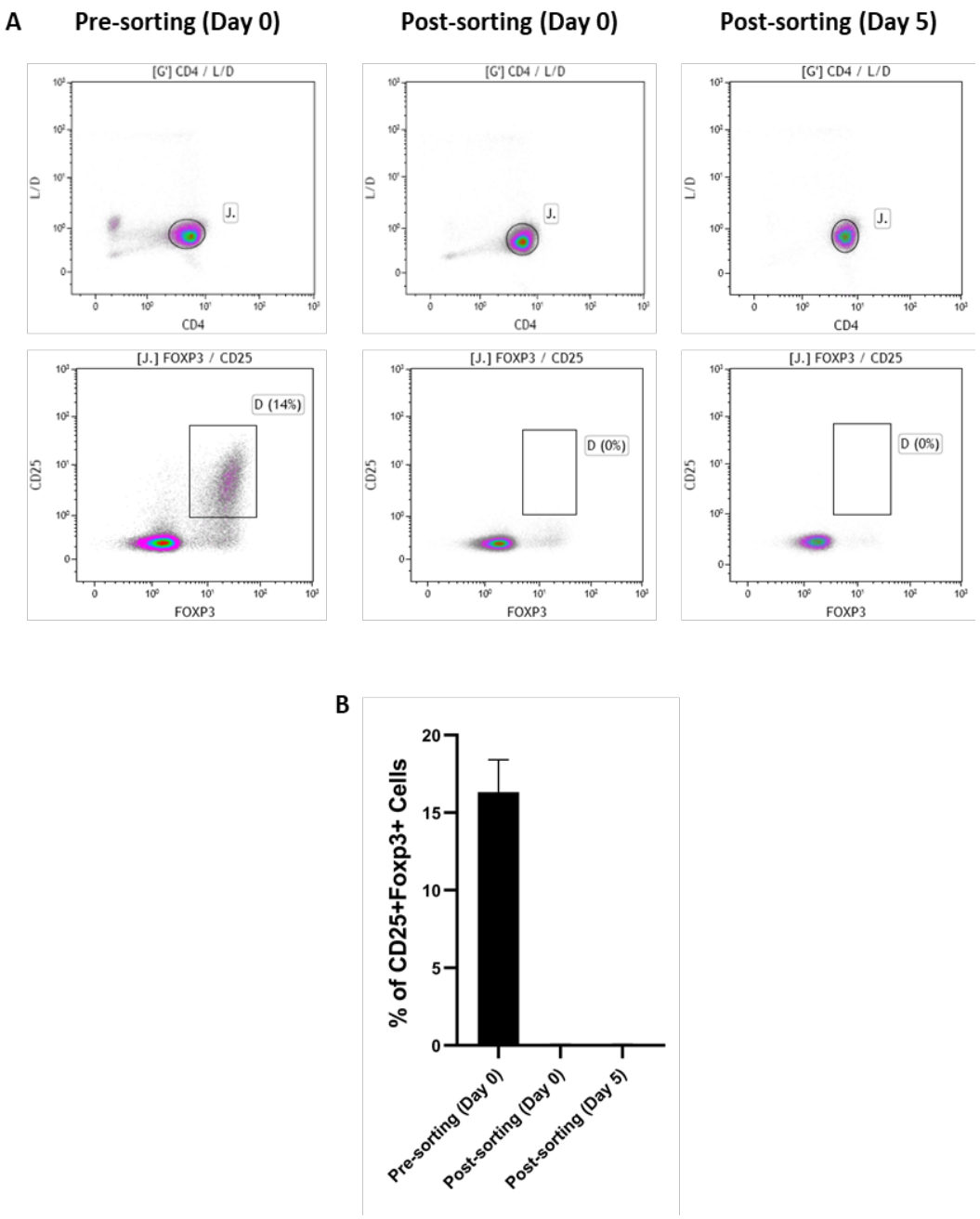


Figure 2.10. The percentage of $CD25^+$ T cells before and after sorting of $CD4^+CD25^-$ T cells from total purified $CD4^+$ T cells. **(A)** Flow cytometric density plots showing the staining of CD25 expression (gate D) on total $CD4^+$ T cells (pre-sorting) and sorted $CD4^+CD25^-$ T cells (post-sorting) at day 0 and day 5 of culture. Gate D was gated on live $CD4^+$ T cells (gate J.). **(B)** Column bar graph illustrating the percentage of $CD25^+Foxp3^+$ T cells within total $CD4^+$ T cells (pre-sorting) and sorted $CD4^+CD25^-$ T cells (post-sorting) at day 0 and day 5 of culture. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

2.3.2. Conversion of naïve $CD4^+CD25^-$ T cells to $CD4^+CD25^+Foxp3^+$ iTregs *in vitro*

2.3.2.1. Generation of iTregs with and without CD3/CD28 TCR activation signals

It is known that TCR activation signals are essential for generating iTregs from naïve $CD4^+CD25^-$ T cells. To assess whether 4T1 tumour cells could induce naïve $CD4^+CD25^-$ T cells to iTregs without CD3/28 TCR activation in the presence and absence of TGF-β1 cytokine. The results showed that only 1.3% of naïve $CD4^+CD25^-$ T cells were induced to iTregs in the presence of tumour cells, whereas 2.2% of iTregs were generated in the presence of both 4T1 cells and TGF-β1 (figure 2.11).The results confirmed that 4T1 cells were able to generate iTregs even in a small percentage in contactless culture, suggesting that 4T1 cells might secrete molecules for mediating that induction independently of TCR activation despite the difference in the induction between both culture conditions was not statistically significant. Although iTregs could be generated without TCR activation in the presence of tumour cells and TGF-β1, the induction was in a very small percentage.

To assess the effect of CD3/CD28 TCR activation signals on the differentiation of iTregs from naïve $CD4^+CD25^-$ T cells, we repeated the same experiment. Naïve $CD4^+CD25^-$ T cells were co-cultured with 4T1 cells in the presence of CD3/28 T cell activator beads and IL-2, and in the presence and absence of TGF-β1 for 5 days. The co-culture was contactless using hanging culture inserts. The results (figure 2.12) showed that only 2%-3% (mean = 2.6 %) of naïve $CD4^+CD25^-$ T cells were induced to iTregs in the presence of 4T1 cells, IL-2 and CD3/28 beads, whereas the induction of iTregs was significantly increased in range from 58% to 67% (mean = 61.3%) in the presence of 4T1, IL-2, CD3/28 beads and TGF-β1 (figure 2.12). The results also confirmed that the induction of iTregs was mainly due to the presence of TGF-β1 and CD3/28 beads as 4T1 cells alone could yield that much induction. However, comparing these results with the results shown in figure 2.11, it is clear that the induction of iTregs results from CD3/28 TCR activation signals and TGF-β1 as the induction was 2.6 % in the presence of 4T1 cells and TGF-β1 without CD3/28 beads, while it was significantly increased to 61.3 % in the presence of 4T1, TGF-β1 and CD3/28 beads.

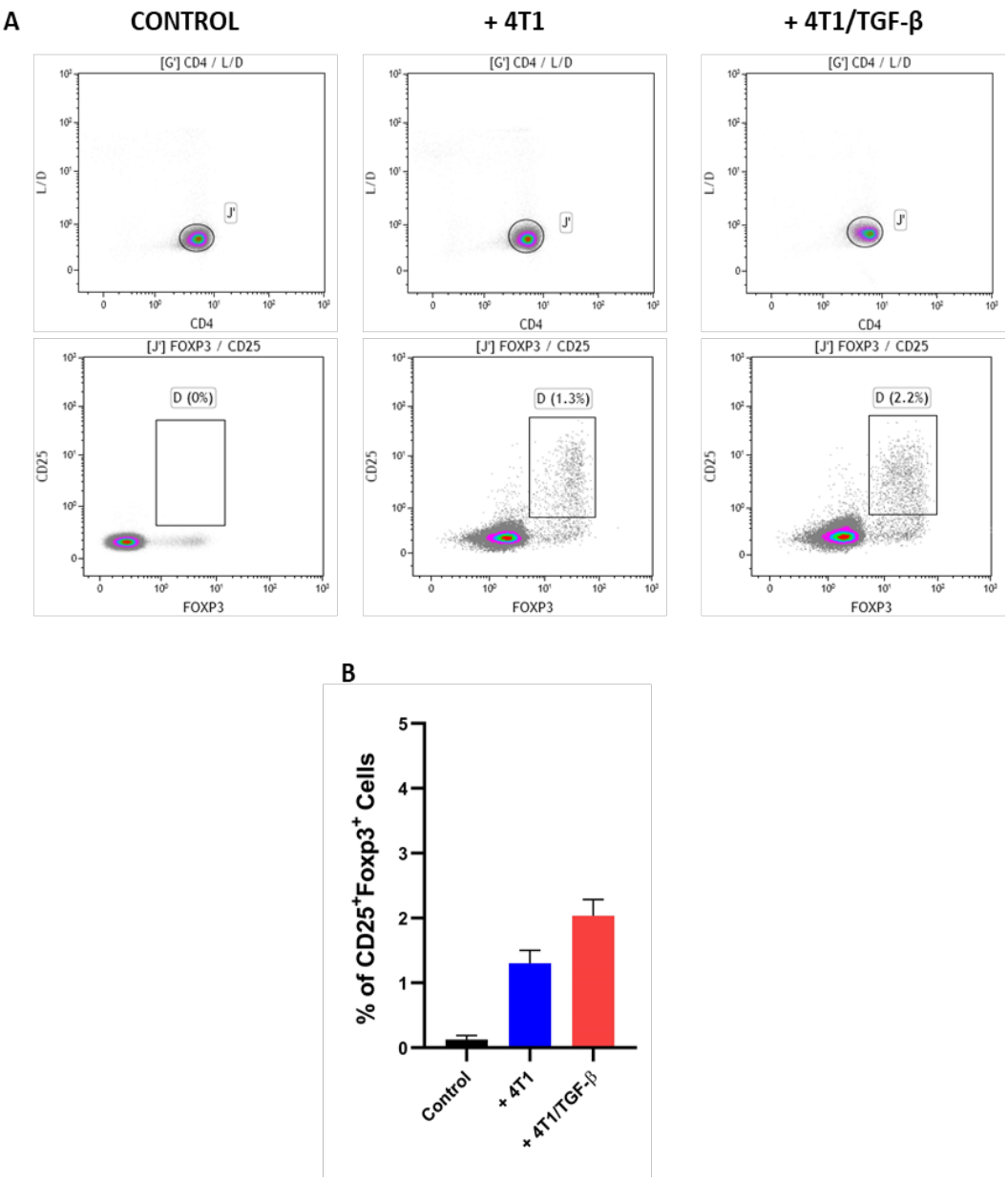


Figure 2.11. The percentage of induction of iTregs from naïve CD4⁺CD25⁻ T cells that were co-cultured with 4T1 tumour in a **contactless** system using hanging culture inserts in the presence and absence TGF-β1 (5 ng/mL) without TCR CD3/CD28 activation. IL-2 (30 U/mL) was added into all culture conditions. Culture ratio of 4T1 cells to naïve CD4⁺CD25⁻ T cells was 1 : 5, respectively. Cells were cultured for 5 days. Control cells were cultured alone in the presence of IL-2 only. **(A)** Flow cytometric density plots displaying the percentage of CD25⁺Foxp3⁺ iTregs (gate D) differentiated from naïve CD4⁺CD25⁻ T cells (gate J'). **(B)** Column bar graph illustrating the percentage of iTregs raised from naïve CD4⁺CD25⁻ T cells cultured in different conditions as mentioned above. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve CD4⁺CD25⁻Foxp3⁻ T cells were cultured in the complete T cell medium in the presence of IL-2 only.

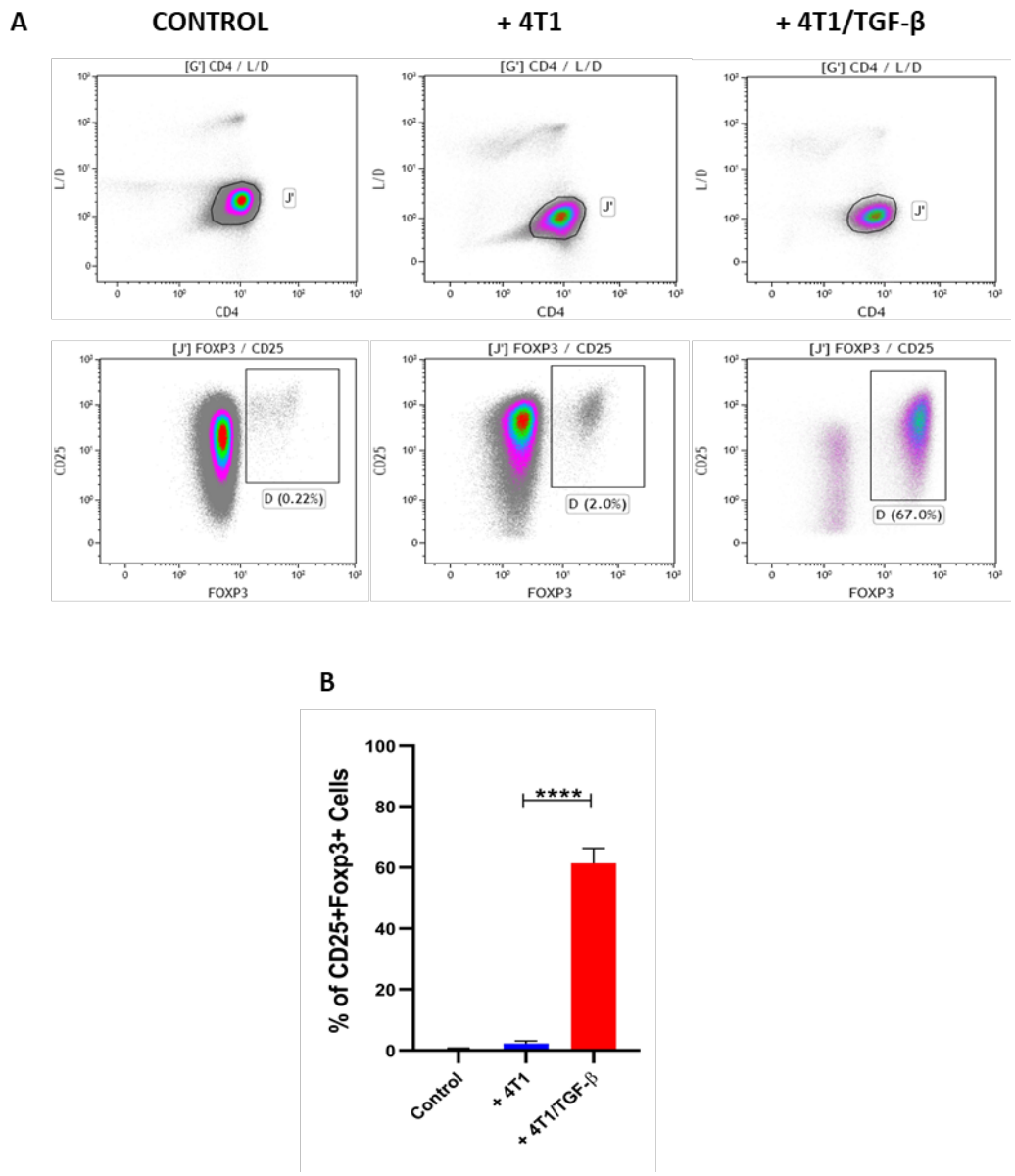


Figure 2.12. The percentage of induction of iTregs from naïve $CD4^{+}CD25^{-}$ T cells that were co-cultured with 4T1 cells (**contactless** co-culture) and CD3/CD28 beads in the presence and absence of TGF-β1 (5 ng/mL) for 5 days. Culture ratio of 4T1 cells to naive$CD4^{+}CD25^{-}$ T cells was 1 : 5, respectively. IL-2 (30 U/mL) was added into all cultures. Control cells (naïve $CD4^{+}CD25^{-}$) were cultured alone in the presence of IL-2 and CD3/28 beads. **(A)** Flow cytometric density plots illustrating the percentage of induction of iTregs (gate D) from live $CD4^{+}CD25^{-}$ T cells (gate J'). **(B)** Column bar graph showing the percentage of $CD25^{+}Foxp3^{+}$ iTregs generated from naïve $CD4^{+}CD25^{-}$ T cells cultured with 4T1 tumour cells in the presence and absence of TGF-β. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells were cultured in the complete T cell medium in the presence of IL-2 only. ($p \leq 0.0001$). Ordinary one-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

2.3.2.2. Generation of iTregs using contact and contactless co-culture models

To define the contribution of 4T1 cells in the induction of iTregs. Then, naïve $CD4^+CD25^-$ T cells were cultured in the presence of IL-2, CD3/28 beads and TGF-β1 in the presence and absence of 4T1 cells using hanging culture inserts (contactless co-culture), in addition to the control where naïve $CD4^+CD25^-$ T cells were cultured in the presence of IL-2 and CD3/28 beads only. The results (figure 2.13) showed that the percentage of induction of iTregs by adding TGF-β1 was 50.7%, whereas it was significantly increased to 61.3 % by adding TGF-β1 in the presence of 4T1 tumour cells, although 4T1 cells contributed to only 2.5 % of induction in the absence of TGF-β1 (figure 2.13). These results suggested that TGF-β1 might direct 4T1 cells to contribute to the induction of iTregs.

To assess whether the presence of 4T1 cells in a cell-cell contact with naïve $CD4^+CD25^-$ T cells could affect the percentage of iTregs induction. Naïve $CD4^+CD25^-$ T cells were co-cultured with 4T1 cells in a contact co-culture without using hanging inserts in the presence and absence of TGF-β1. The results of this experiment showed (figure 2.14) that cell-cell contact co-culture between 4T1 tumour cells and naïve $CD4^+CD25^-$ T cells yielded 62.7 % of iTregs induction, while contactless co-culture showed 61.3 % (figure 2.13), confirming that 4T1 secretome (4T1-conditioned media) was enough for mediating the induction of iTregs in the presence of TGF-β1. This shows that naïve $CD4^+CD25^-$ T cells do not require cell-cell contact with tumour cells to be differentiated to iTregs. Therefore, the contactless co-culture system was used in this model.

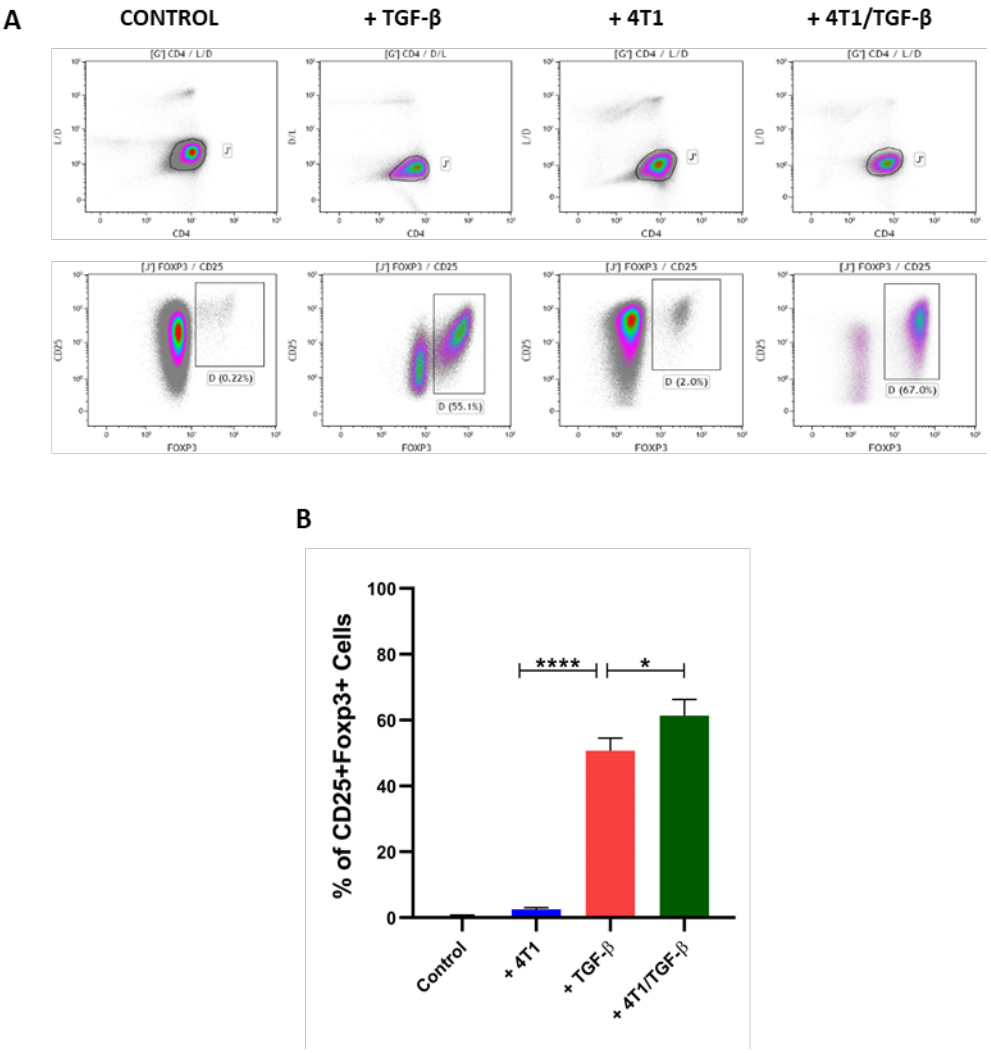


Figure 2.13. The percentage of induction of iTregs from naïve $CD4^+CD25^-$ T cells that were cultured in different conditions for 5 days in vitro. Here the co-culture of naïve $CD4^+CD25^-$ T cells and 4T1 cells was **contactless**. IL-2 (30 U/mL) was added into all cultures. TGF-β1 (5 ng/mL) was added into the relevant cultures (+TGF-β). Culture ratio of 4T1 cells to naive$CD4^+CD25^-$ T cells was 1 : 5, respectively. Control cells (naïve $CD4^+CD25^-$) were cultured alone in the presence of IL-2 and CD3/28 beads. **(A)** Flow cytometric density plots illustrating the percentage of induction of iTregs (gate D) from live $CD4^+CD25^-$ T cells (gate J'). **(B)** Column bar graph showing the percentage of $CD25^+Foxp3^+$ iTregs generated from naïve $CD4^+CD25^-$ T cells as mentioned above. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve $CD4^+CD25^-Foxp3^-$ T cells were cultured in the complete T cell medium in the presence of IL-2 only. ($p \leq 0.0001$ (****), $p \leq 0.01$ (*)). Ordinary one-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

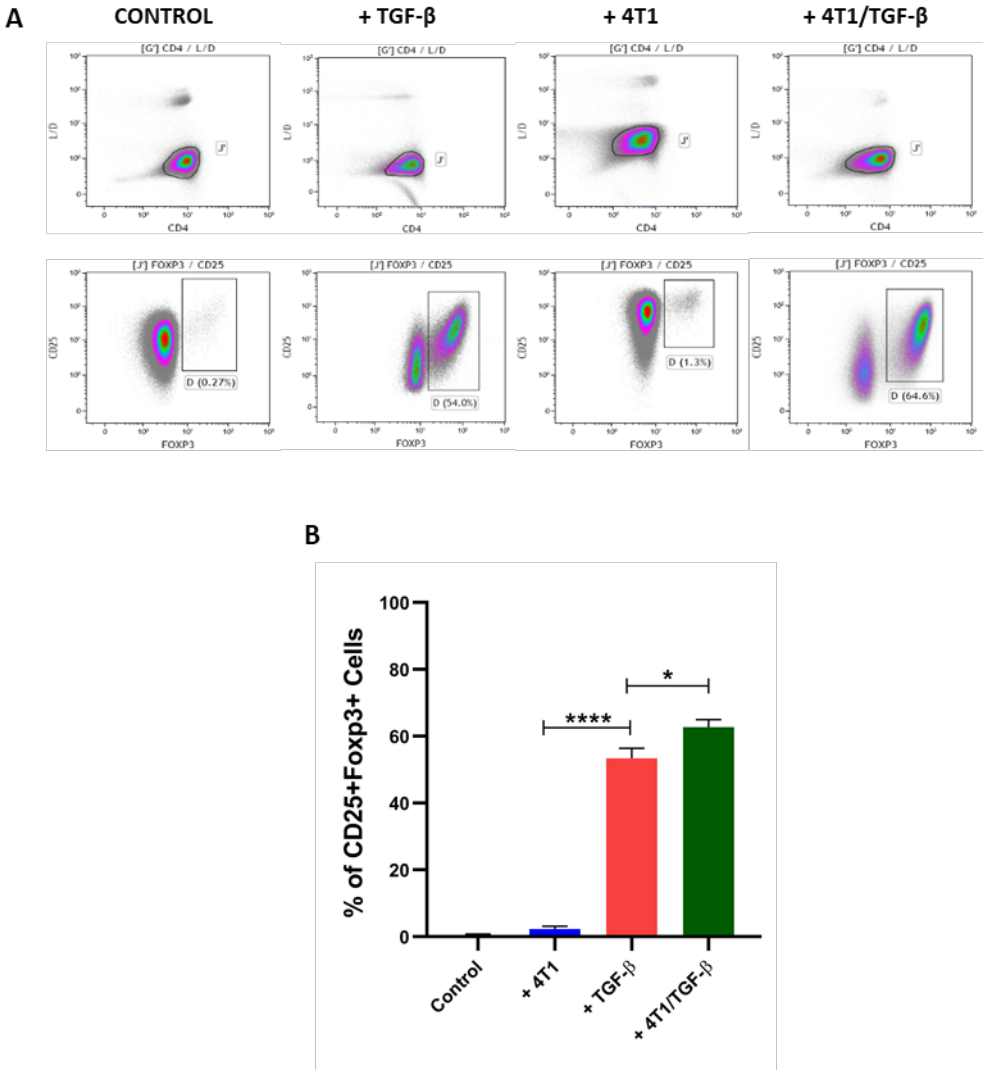


Figure 2.14. The percentage of induction of iTregs from naïve $CD4^+CD25^-$ T cells that were cultured in different conditions for 5 days in vitro. Here the co-culture of naïve $CD4^+CD25^-$ T cells and 4T1 cells was **cell-cell contact**. IL-2 (30 U/mL) was added into all cultures. TGF-β1 (5 ng/mL) was added into the relevant cultures (+TGF-β). Culture ratio of 4T1 cells to naive$CD4^+CD25^-$ T cells was 1 : 5, respectively. Control cells (naïve $CD4^+CD25^-$) were cultured alone in the presence of IL-2 and CD3/28 beads. **(A)** Flow cytometric density plots illustrating the percentage of induction of iTregs (gate D) from live $CD4^+CD25^-$ T cells (gate J'). **(B)** Column bar graph showing the percentage of $CD25^+Foxp3^+$ iTregs generated from naïve $CD4^+CD25^-$ T cells as mentioned above. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve $CD4^+CD25^-Foxp3^-$ T cells were cultured in the complete T cell medium in the presence of IL-2 only. ($p \leq 0.0001$ (****), $p \leq 0.01$ (*)). Ordinary one-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

2.3.2.3. The presence of TGF-β1 is essential during CD3/28 TCR activation signals for optimal induction of iTregs

As both CD3/28 TCR activation and TGF-β1 were required for inducing iTregs from naïve $CD4^+CD25^-$ T cells *in vitro*. Next, it was assessed whether the initiation of CD3/28 TCR activation signals simultaneously in the presence of TGF-β1 is required for iTregs differentiation from naïve $CD4^+CD25^-$ T cells. To achieve that, naïve $CD4^+CD25^-$ T cells were activated with CD3/28 beads in the presence of IL-2, then TGF-β1 was added into cultures at different time points: immediately following adding CD3/28 beads (0 h), 2 hours after adding CD3/28 beads (2 h), 5 hours after adding CD3/28 beads (5 h) and 24 hours after adding CD3/28 beads (24 h). This experiment was done in the presence and absence of 4T1 tumour cells. Cultures were incubated for 5 days at 37^oC, 5% CO_2. The results showed that the presence of TGF-β1 immediately during CD3/28 TCR activation signals was substantial for generating iTregs (figure 2.15). The mean of the percentage of TGF-β1-derived iTregs induction was 54.3%, 43.6%, 38.8% and 11.6% at 0 h, 2hrs, 5 h and 24 h, respectively. These results showed that the induction of iTregs was significantly declined by increasing the period between CD3/28 TCR activation of naïve $CD4^+CD25^-$ T cells and the addition of TGF-β1. Similarly, the mean of the percentage of TGF-β1/4T1-derived iTregs was significantly decreased from 66% at 0 h to 15.06 % at 24 h of adding TGF-β1 into CD3/28 activated $CD4^+CD25^-$ T cells *in vitro* (figure 2.15). Taken together, these results confirmed that the induction of iTregs is tightly dependent on the time when naïve $CD4^+CD25^-$ T cells received TCR-activation signals and the availability of secreted TGF-β1 in the induction milieu. Moreover, these results suggest that naïve $CD4^+CD25^-$ T cells might resist TGF-β1 signalling after prolonged CD3/28 TCR activation and become differentiated to effector $CD4^+$ T cells.

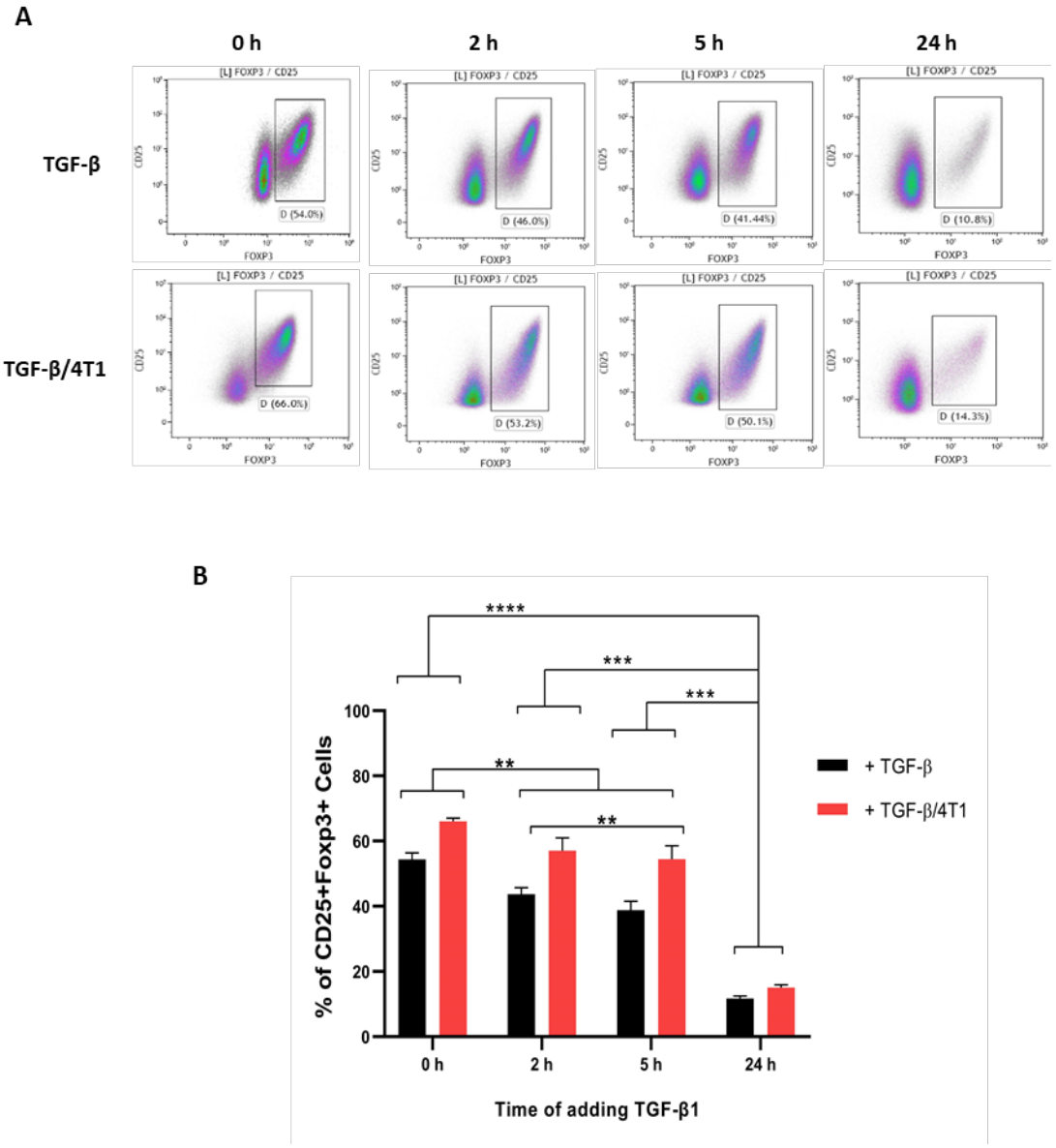


Figure 2.15. The percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs from CD3/28-activated $CD4^+CD25^-$ T cells following adding TGF-β1 (5 ng/mL) at different time points. Cells were cultured for 5 days in vitro. IL-2 (30 U/mL) was added into all cultures. Culture ratio of 4T1 cells to naïve $CD4^+CD25^-$ T cells was 1 : 5, respectively. **(A)** Flow cytometric density plots illustrating the percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs (gate D) that was gated on live $CD4^+CD25^-$ T cells (gate L (not shown)). **(B)** Column bar graph showing the percentage of $CD25^+Foxp3^+$ TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs generated from CD3/28-activated $CD4^+CD25^-$ T cells. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. ($p \leq 0.0001$ (****), $p \leq 0.0007$ (***), $p \leq 0.0020$ (**)). Two-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

2.3.2.4. Generation of iTregs using different concentrations of TGF-β1

Due to the crucial role of TGF-β1 for mediating the induction of iTregs from naïve $CD4^+CD25^-$ T cells following CD3/28 TCR activation signals. Next, the induction of iTregs using different concentrations of TGF-β1 was assessed *in vitro*. Here, naïve $CD4^+CD25^-$ T cells were activated with CD3/28 beads and cultured in the presence and absence of 4T1 cells. Then TGF-β1 was added into cultures at different concentrations including 5 ng/mL, 3 ng/mL, 1 ng/mL, 0.5 ng/mL, 0.05 ng/mL and 0.005 ng/mL. The results (figure 2.16B) showed that the induction of iTregs was significantly decreased from 58.8% at concentration 5 ng/mL to 41.3% at concentration 0.5 ng/mL, 4.46% at concentration 0.05 ng/mL and 1% at 0.005 ng/mL (figure 2.16B). Also, the results confirmed that the induction of iTregs was increased with the increase of TGF-β concentration. Although there were no significant differences between concentrations 5 ng/mL, 3 ng/mL and 1 ng/mL, the concentration 5 ng/mL yielded the highest induction of iTregs. These results also suggested that the concentration 0.05 ng/mL of TGF-β could be the minimum concentration required for inducing iTregs *in vivo*. Moreover, the induction of iTregs was significantly increased in the presence of 4T1 tumour cells with all concentrations of TGF-β, however the highly significant difference in the induction of iTregs in the presence and absence of 4T1 cells was at concentration 0.05 ng/mL of TGF-β (20.6% and 4.46%, respectively) (figure 2.16C). Accordingly, the concentration of 5 ng/mL was selected to be used for the next experiments.

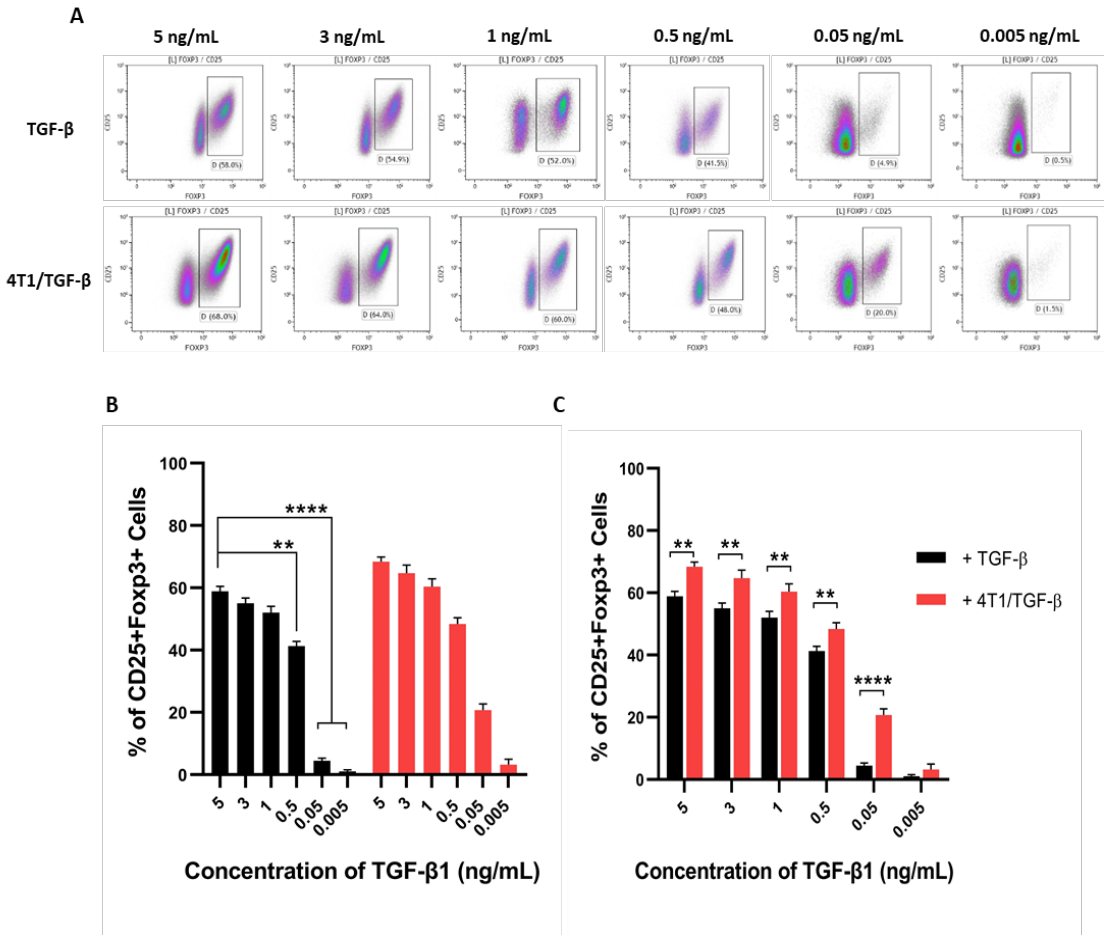


Figure 2.16. The percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs from CD3/28-activated $CD4^{+}CD25^{-}$ T cells following adding TGF-β1 at different concentrations. Cells were cultured for 5 days in vitro. IL-2 (30 U/mL) was added into all cultures. Culture ratio of 4T1 cells to naïve $CD4^{+}CD25^{-}$ T cells was 1 : 5, respectively. **(A)** Flow cytometric density plots illustrating the percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs (gate D) that was gated on live $CD4^{+}CD25^{-}$ T cells (gate L (not shown)). **(B)** Column bar graph showing the percentage of $CD25^{+}Foxp3^{+}$ TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs generated from CD3/28-activated $CD4^{+}CD25^{-}$ T cells using different concentrations of TGF-β1. **(C)** Column bar graph illustrating the percentage of iTregs induction in the presence and absence of 4T1 using different concentrations of TGF-β1. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. ($p \leq 0.0001$ (****), $p \leq 0.0065$ (**)). Two-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

2.3.2.5. Optimising the duration of induction of iTregs *in vitro*

As "5 days" period of culture was followed for all experiments mentioned above for generating iTregs *in vitro*. Next, the induction of iTregs was tracked over 6 days period of time to assess whether the induction of iTregs would increase with time in the presence of CD3/28 beads, IL-2 and TGF-β1. Here, naïve $CD4^+CD25^-$ T cells were activated with CD3/28 beads and cultured in the presence of TGF-β1 and IL-2 and in the presence and absence of 4T1 cells for 6 days at 37°C, 5% CO_2. Then, the induction of iTregs was observed in each day of culture. The results (figure 2.17) showed that the percentage of induction of iTregs was significantly increased from day 1 (11.3%; with TGF-β, and 18.3%; with TGF-β1 and 4T1 cells) until day 6 (51%; with TGF-β1, and 63%% with TGF-β1 and 4T1 cells). However, although there were no significant differences in the induction of iTregs between day 4, day 5 and day 6, day 5 showed the highest percentage of iTregs induction (figure 2.17). As a result, 5 days period of culture was selected to be used in the model. These results also justify why 5 days period of culture was followed in the previous experiment.

2.3.2.6. Optimising the co-culture ratio of naïve $CD4^+CD25^-$ T cells and 4T1 tumour cells for generating iTregs *in vitro*

As the induction of iTregs was significantly increased in the presence of 4T1 cells and TGF-β1, compared to cells cultured in the presence of TGF-β1 only. It was assessed whether iTregs induction could be affected by the number of tumour cells co-cultured with naïve $CD4^+CD25^-$ T cells. Next, 4T1 cells were co-cultured with naïve $CD4^+CD25^-$ T cells at different ratio in the presence of TGF-β1, IL-2 and CD3/28 TCR activation beads. The results (figure 2.18A,B) showed that the percentage of iTregs induction was significantly higher at ratio 1:2 (70%) and 1:3 (69%), compared with the control (56.1%) where naïve $CD4^+CD25^-$ T cells were cultured in the absence of 4T1 (figure 2.18A,B). Even in the presence of 4T1 cells, the induction of iTregs was significantly increased at ratio 1:4 (67%) and 1:5 (64.3%), compared with the control. There were no significant differences in the induction of iTregs between ratios 1:25 (61%), 1:50 (58.8%) and the control (figure 18A,B). These results suggest that the number of naïve $CD4^+CD25^-$ T cells infiltrating to the tumour microenvironment might contribute to increase the differentiation of

iTregs within tumour milieu, particularly if the infiltration ratio was 1:2 or 1:3 (tumour cells: naïve CD4 T cells).

However, although the induction of iTregs was increased when naïve $CD4^+CD25^-$ T cells were co-cultured with 4T1 cells in the presence of TGF-β1, it was noticed that the number of $CD4^+$ T cells recovered after 5 days of culture was significantly reduced, compared to the control population where naïve $CD4^+CD25^-$ T cells were cultured in the absence of tumour cells. The results (figure 2.18C) showed that an average of 6.4 x 10^6 total $CD4^+$ T cells (which include activated $CD4^+CD25^+Foxp3^-$ and $CD4^+CD25^+Foxp3^+$ iTregs) were proliferated from 1 x 10^6 $CD4^+CD25^-$ T cells in the presence of CD3/28 beads, IL-2 and TGF-β1 for 5 days of culture, whereas the average number of total $CD4^+$ T cells proliferated from 1 x 10^6 in the presence of 4T1 cells was in range between 1.52 x 10^6 cells (ratio 1:2) and 3.93 x 10^6 cells (ratio 1:50) (figure 2.18C). This significant reduction in the cell recovery after 5 days in the presence of 4T1 suggest that: (1) 4T1 cells might secrete stress signals that could affect the proliferation of naïve $CD4^+CD25^-$ T cells even in the presence of CD3/28 TCR activation beads, or (2) iTregs that were induced in the presence of 4T1 might be more suppressive cells that could inhibit the proliferation of activated $CD4^+CD25^+Foxp3^-$ T cells than TGF-β1-derived iTregs.

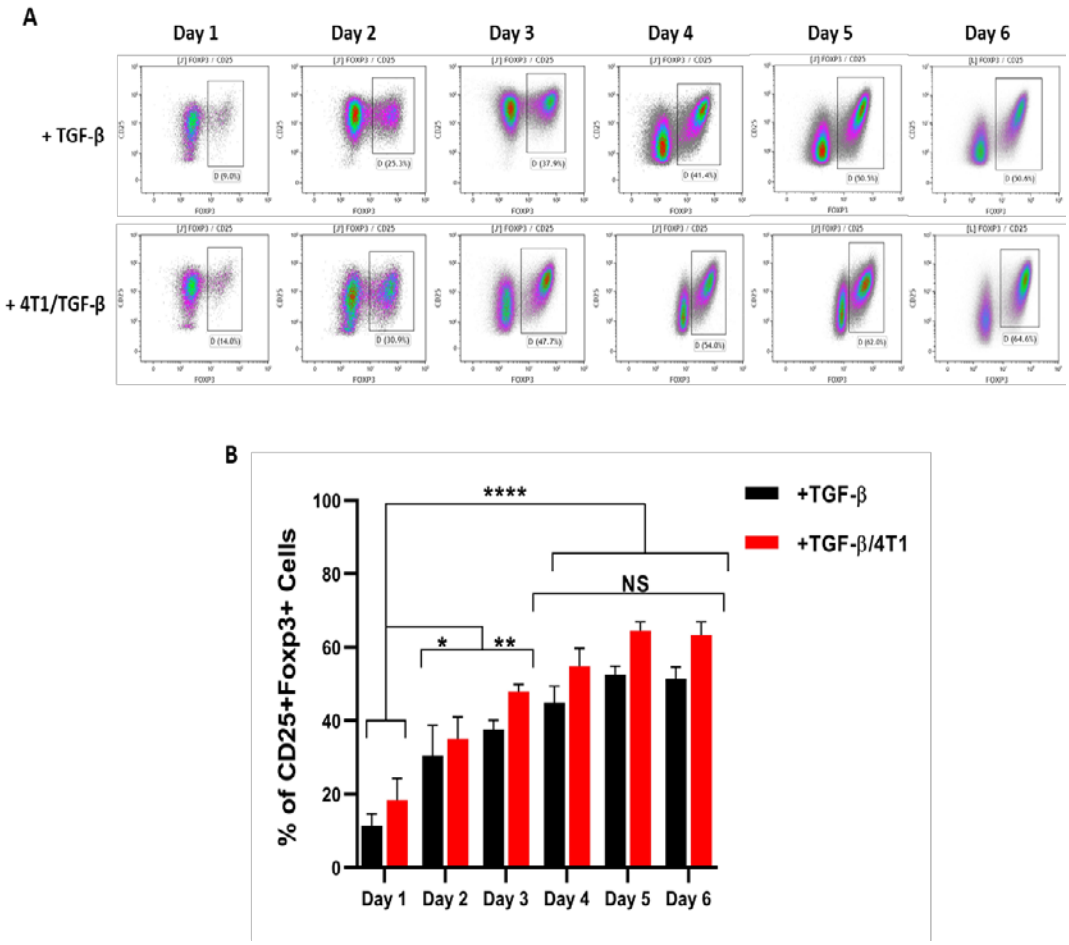


Figure 2.17. The percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs from CD3/28-activated $CD4^{+}CD25^{-}$ T cells over 6 days of culture in vitro. The induction of iTregs was assessed in each day. IL-2 (30 U/mL) was added into all cultures. Culture ratio of 4T1 cells to naive$CD4^{+}CD25^{-}$ T cells was 1 : 5, respectively. **(A)** Flow cytometric density plots illustrating the percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs (gate D) that was gated on live $CD4^{+}CD25^{-}$ T cells (gate L (not shown)). **(B)** Column bar graph showing the percentage of $CD25^{+}Foxp3^{+}$ TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs generated from CD3/28-activated $CD4^{+}CD25^{-}$ T cells in each day for 6 days period of culture. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. ($p \leq 0.0001$ (****), $p \leq 0.0003$ (***), $p \leq 0.0095$ (**), $p \leq 0.0118$ (*)). Two-way ANOVA test was used for analysis using GraphPad Prism 8.0.1.

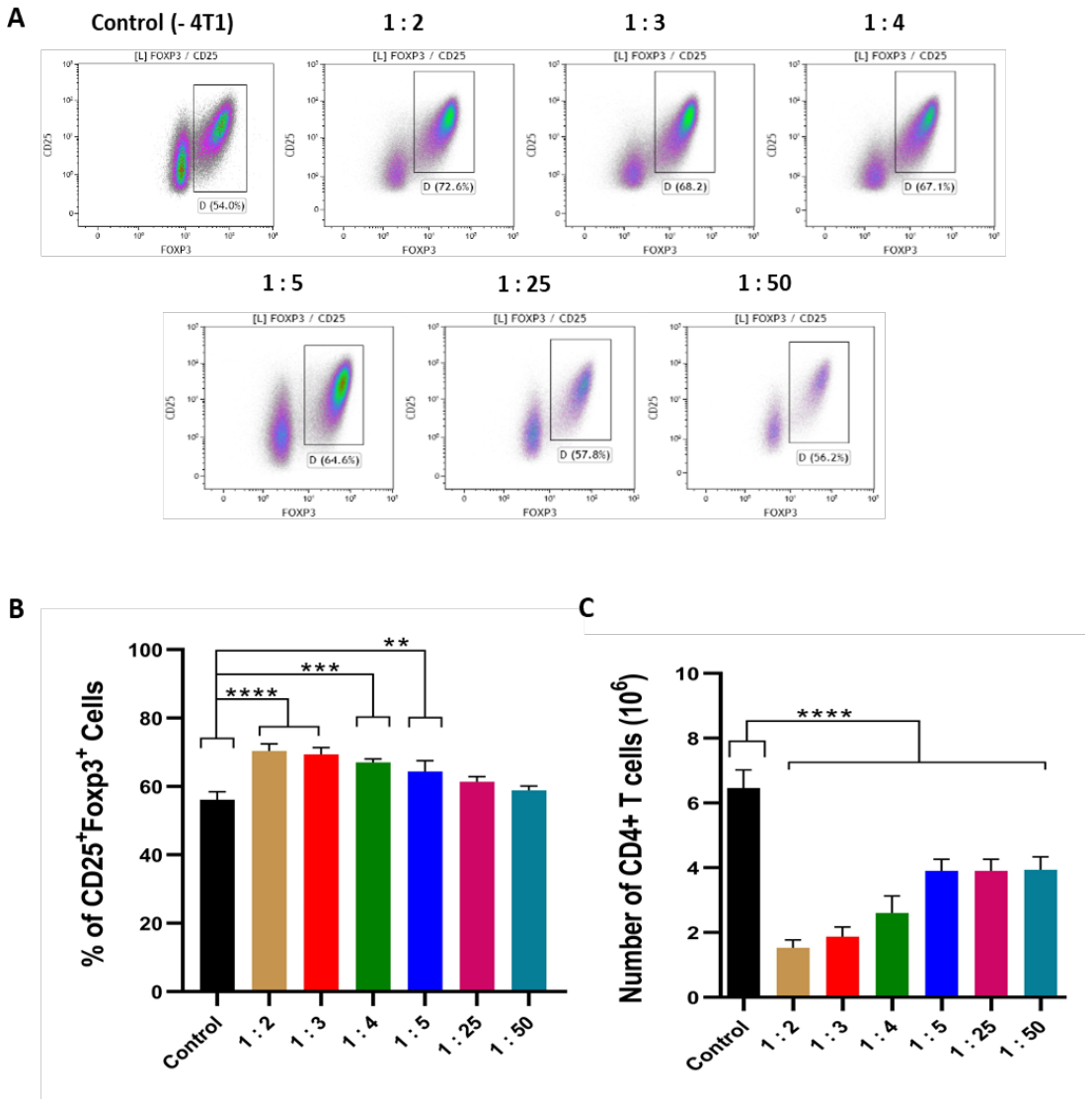


Figure 2.18. The percentage of induction of iTregs from CD3/28-activated CD4$^+$CD25$^-$ T cells co-cultured with 4T1 tumour cells at different culture ratios (4T1 cells : naïve CD4$^+$CD25$^-$ T cells) in vitro for 5 days. IL-2 (30 U/mL), TGF-β1 (5 ng/mL) were added into all cultures. (A) Flow cytometric density plots illustrating the percentage of induction of TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs (gate D) that was gated on live CD4$^+$CD25$^-$ T cells (gate L (not shown)). (B) Column bar graph showing the percentage of TGF-β1/4T1-derived iTregs generated from CD3/28-activated CD4$^+$CD25$^-$ T cells co-cultured with 4T1 cells in different ratios. Ordinary one-way ANOVA test was used for analysis using GraphPad Prism 8.0.1. (C) Column bar graph illustrating the average of total number of CD4$^+$ T cells that were proliferated from 1x10^6 of naïve CD4$^+$CD25$^-$ T cells co-cultured with 4T1 cells at different ratios. Control population includes naïve CD4$^+$CD25$^-$ T cells cultured without 4T1 cells. Ordinary one-way ANOVA test was used for analysis using GraphPad Prism 8.0.1. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. ($p \leq 0.0001$ (****), $p \leq 0.0003$ (***), $p \leq 0.0035$ (**)). Control: naïve CD4$^+$CD25$^-$Foxp3$^-$ T cells were cultured in the complete T cell medium in the presence of IL-2 and TGF-β1, in the absence of 4T1 tumour cells.

2.3.2.7. Establishment of functional assay

To assess whether iTregs generated in this model could function as immunosuppressive cells, a functional assay was developed by co-culturing the generated iTregs with purely sorted $CD8^+$ T cells that were stained with CFSE dye for 4 days. To achieve this, first, the purity of $CD8^+$ T cell sorting was assessed, and then the proliferation of naïve $CD8^+$ T cells following CFSE staining in the presence of IL-2 and CD3/28 TCR activator beads (1:1 ratio ($CD8^+$ cells:CD3/38 beads)) for 4 days was assessed *in vitro*. The results of the purity of cell sorting shown in figure 2.19. After pure sorting, naïve $CD8^+$ T cells were stained with CFSE (5µM), activated with CD3/28 beads and cultured in the presence of IL-2 (30 U/mL) for 4 days. The results of the proliferation of CFSE-stained $CD8^+$ T cells showed that 8 generations of activated $CD8^+$ T cells were proliferated from $CD8^+$ paternal T cells after 4 days of culture. The results also showed that the staining of CFSE was nicely shifted from paternal cells with higher CFSE intensity to the 8^{th} generation with lower CFSE intensity (figure 2.20).

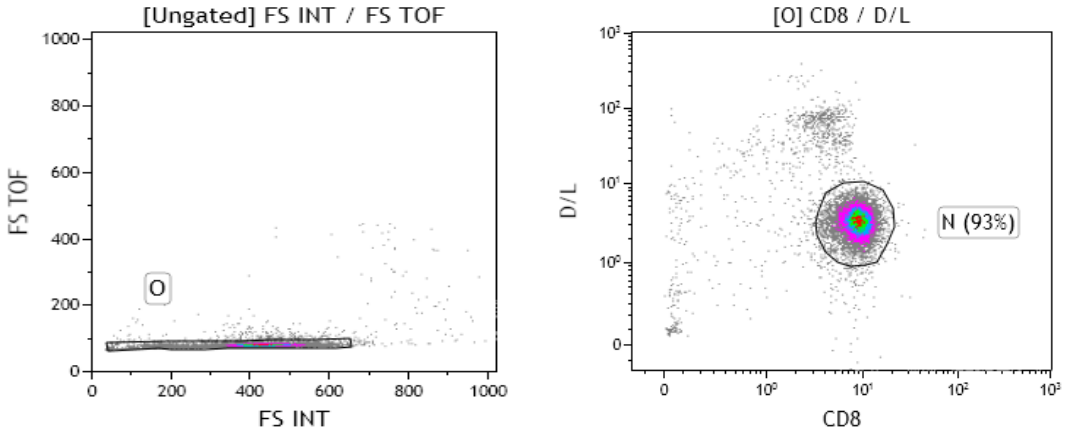


Figure 2.19. Flow cytometric density plots showing the percentage of purity of $CD8^+$ T cells sorting. $CD8^+$ T cells (gate N) that were gated on singlets cells (gate O). 93% of $CD8^+$ T cells were live after cell sorting. (n=4).

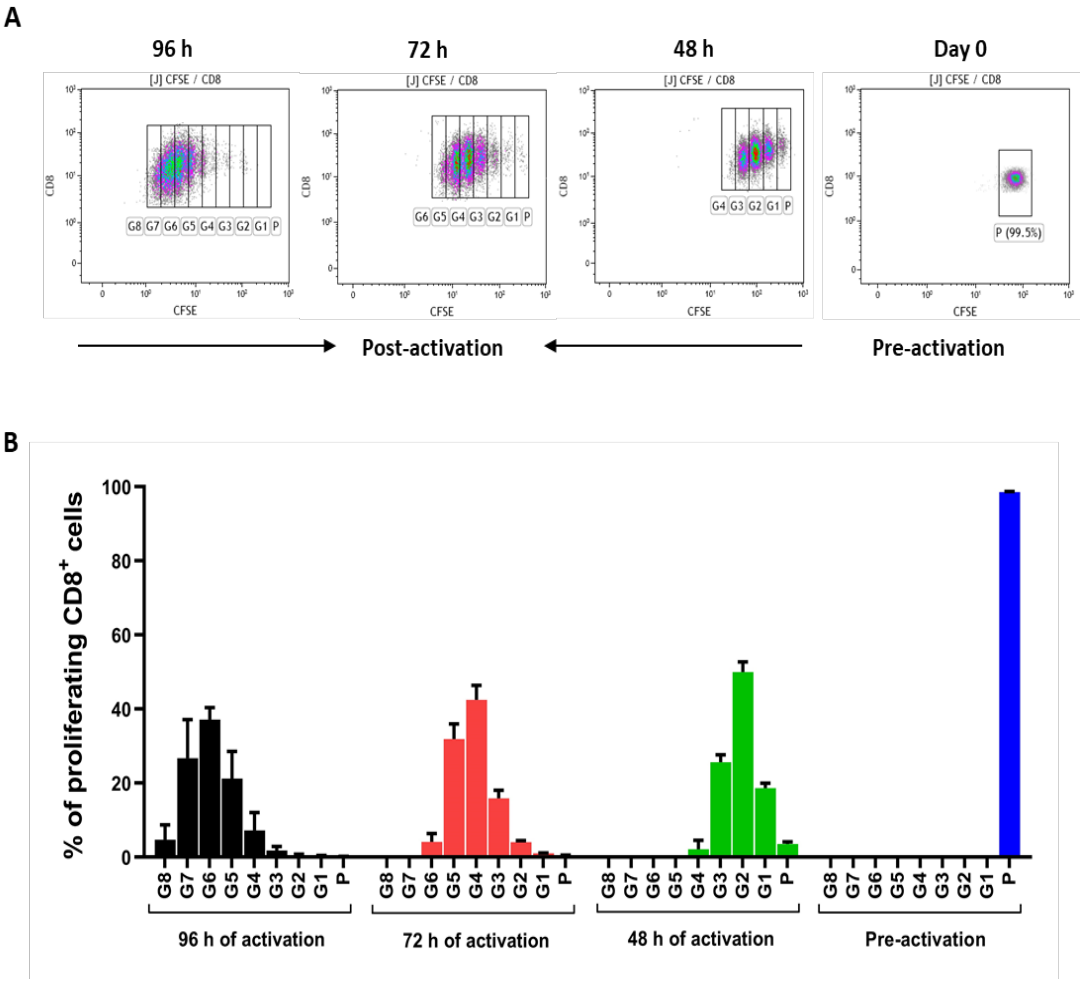


Figure 2.20. The percentage of proliferating CD8⁺ T cells from CFSE-stained paternal CD8⁺ T cells cultured in the presence of CD3/28 beads and IL-2 (30 U/mL) in vitro for 4 days. **(A)** Flow cytometric density plots showing the shift of CFSE-intensity of proliferating CD8⁺ T cells from higher intensity (right, log 10^3) to lower intensity (toward left, log 10^0), gate P (parental unstimulated CD8⁺ T cells), gate G1 (generation 1 of proliferating CD8⁺ T cells after CD3/28 TCR activation), gate G8 (generation 8 of proliferating CD8⁺ T cells). Kalusa software was used for analysing flow cytometric data. **(B)** Column bar graph illustrating the percentage of proliferating CD8⁺ T cells in each generation following staining with CFSE cell trace dye for 4 days of culture in vitro. GraphPad Prism 8.0.1 software was used for drawing the graph. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

As the trend of proliferation of $CD8^+$ T cells *in vitro* was tracked and achieved in this model. Next, the functional activity of the generated iTregs was assessed by co-culturing them with sorted, CFSE-stained $CD8^+$ T cells as mentioned above at different ratios. The results (figure 2.21) of the functional assay showed that the generated iTregs were highly immunosuppressive cells which could significantly inhibit the proliferation of $CD8^+$ T cells even in the presence of CD3/28 TCR activator and IL-2. The results showed that the percentage of $CD8^+$ T cells proliferation was only ~3% at 1:2 ratio ($CD8^+$ cells:iTregs), compared to the control (100%) where $CD8^+$ T cells were cultured without iTregs. Also, at ratio (1:1) the inhibition of $CD8^+$ T cells proliferation was ~95% by TGF-β1-derived iTregs and ~97% by TGF-β1/4T1-derived iTregs (figure 2.21). Even at ratio (16:1) the inhibition of $CD8^+$ T cells proliferation was 18% by TGF-β1-derived iTregs and 22.8% by TGF-β1/4T1-derived iTregs, which was significant compared to the control. The results also showed that iTregs induced in the presence of 4T1 were slightly more suppressive than iTregs induced by TGF-β1 only at specific ratios including (1:1), (2:1), (4:1) and 16:1), however these results were not statistically significant (figure 2.21).

2.3.2.8. The stability of Foxp3 expression of iTregs

As the generated iTregs were functionally immunosuppressive cells, it was assessed whether iTregs could maintain foxp3 expression after induction. The results showed (figure 2.22) both TGF-β1-derived iTregs and TGF-β1/4T1-derived iTregs lost about 5% of their Foxp3 expression after 5 days of removing TGF-β1 and CD3/28 TCR activator beads. Although there was a reduction in the expression of Foxp3 of iTregs, the reduction was not statistically significant (figure 2.22). These results suggest that the generation of iTregs might be based on the secretion of cytokines such as TGF-β1 with persistent TCR signals *in vivo*.

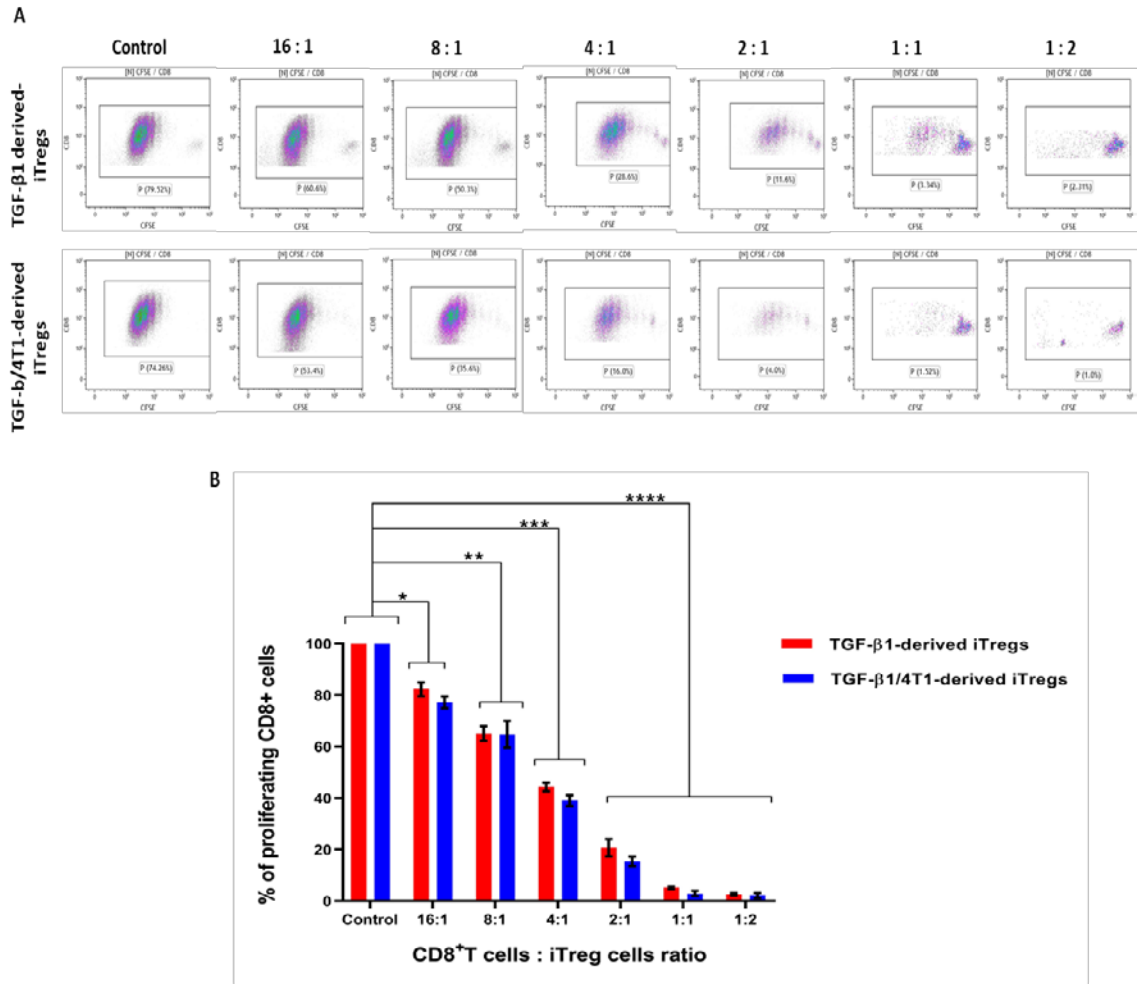


Figure 2.21. The percentage of proliferating CD8⁺ T cells from CFSE-stained parental CD8⁺ T cells cultured in the presence of TGF-β1/4T1-derived iTregs and TGF-β1-derived iTregs at different ratios for 4 days *in vitro*. CD3/28 beads (1:1 ratio (CD8:CD3/28 beads) and IL-2 (30 U/mL) were used for activating CD8⁺ T cells. **(A)** Flow cytometric density plots showing the percentage of proliferating CD8⁺ T cells (gate P) co-cultured with iTregs at different ratios (CD8 : iTregs). The proliferation of CD8⁺ T cells is illustrated by the change of CFSE staining intensity from higher intensity (right, log 10^3) to lower intensity (toward left, log 10^0). Kalusa software was used for analysing flow cytometric data. **(B)** Column bar graph illustrating the percentage of proliferating CD8⁺ T cells cultured in the presence of iTregs as mentioned above. By adding iTregs, the proliferation of CD8⁺ T cells was significantly reduced according to the culture ratio. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Control: naïve purely sorted CD8⁺ T cells were cultured in the complete T cell medium in the presence of IL-2, in the absence of iTregs. ($p \leq 0.0001$ (****), $p \leq 0.0010$ (***), $p \leq 0.0099$ (**), $p \leq 0.0301$ (*). Two-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software.

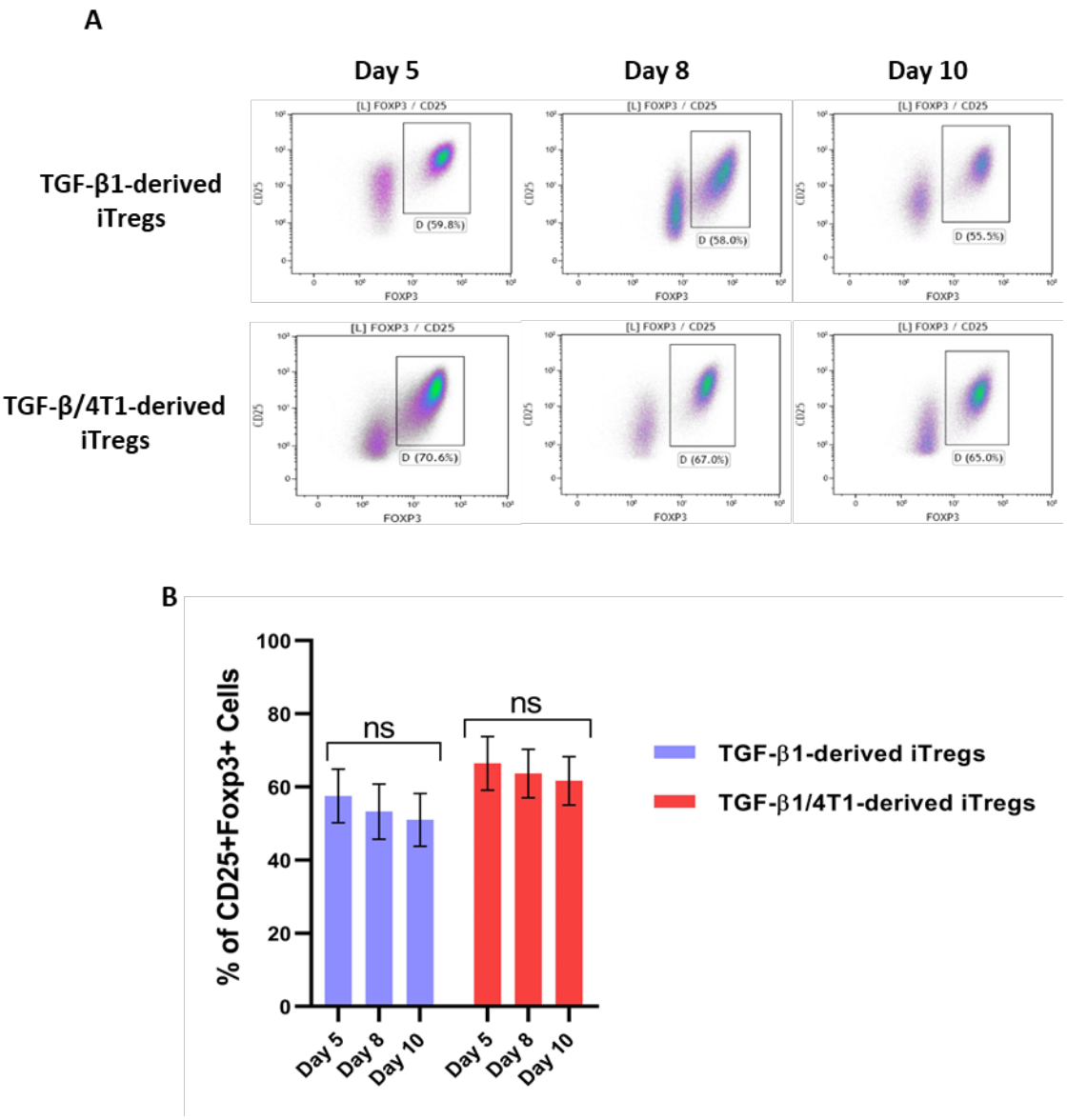


Figure 2.22. The percentage of Foxp3 expression of TGF-β1/4T1-derived iTregs and TGF-β1-derived iTregs after 3 and 5 days of removing TGF-β1 and CD3/28 TCR activator beads (which is 8 and 10 days of induction). iTregs were induced as mentioned above for 5 days. Then cells were cultured in a fresh complete T cell medium supplemented with IL-2 (30 U/mL) only. CD3/28 beads were removed magnetically. Then cells were cultured for another 5 days, and the expression of Foxp3 was assessed at day 3 and 5. **(A)** Flow cytometric density plots showing the percentage of Foxp3 expression of iTregs (gate D). gate D was gated on live $CD4^+$ T cells (gate L (not shown)). Kalusa software was used for analysing flow cytometric data. **(B)** Column bar graph illustrating the percentage of the expression of Foxp3 in both TGF-β1/4T1-derived iTregs and TGF-β1-derived iTregs at day 5, 8 and 10 of induction as mentioned above. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. Two-way ANOVA test was used for analysis using GraphPad Prism 8.0.1 software.

2.3.2.9. Foxp3-TSDR methylation status of the generated iTregs

The methylation status of Foxp3-TSDR of iTregs generated in this model was assessed. The results (figure 2.23) showed that all CpG regions were significantly hypermethylated within the Foxp3-TSDR of naïve and activated $CD4^+CD25^-$ T cells as the percentage of methylation was more than 90% for all CpGs compared to the percentage of methylation of the CpGs of naïve $CD4^+CD25^+Foxp3^+$ nTregs, which was between 2.5% and 8.5%. However, the percentage of methylation of the CpGs of TGF-β1-drived iTreg, TGF-β1/4T1-drived iTregs was between 71% and 78 %, which was significantly lower than the methylation status of naïve and activated $CD4^+CD25^-$ T cells (figure 2.23). However, the percentage of methylation of CpGs of iTregs was significantly higher than those of naïve nTregs. Surprisingly, the CpGs of activated nTregs were significantly hypermethylated (57-63%), compared to naïve nTregs (figure 2.23). Taken together, these results confirmed that the methylation status of 4 CpGs within Foxp3 TSDR was decreased in iTregs due to the presence of TGF-β1 mainly and the 4T1 cells, compared to naïve $CD4^+CD25^-$ T cells and activated $CD4^+CD25^-$ T cells. Although the generated iTregs were not as hypomethylated as naïve nTregs, they were slightly similar to activated nTregs.

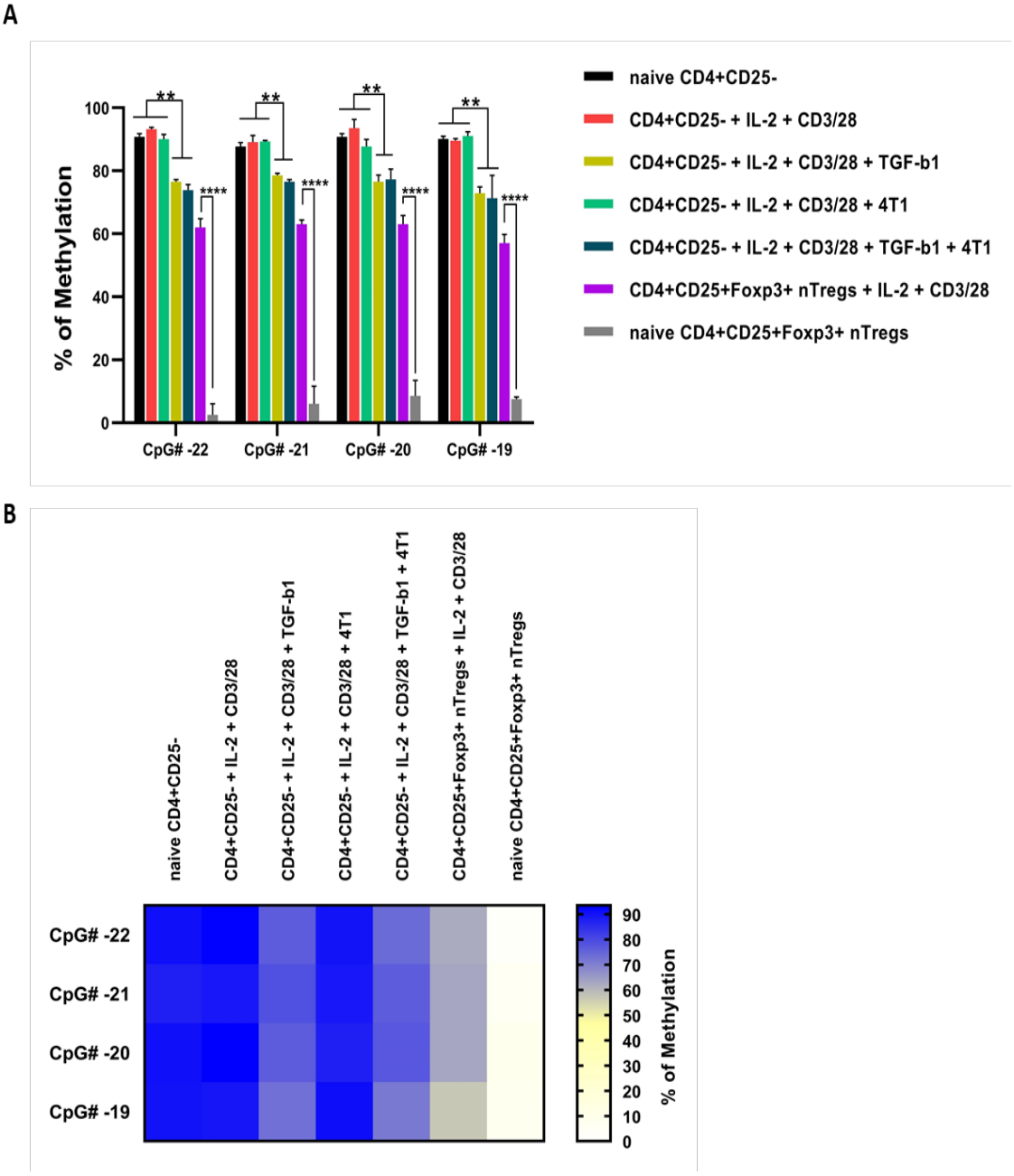


Figure 2.23. The percentage of methylation of 4 different CpG regions within Foxp3-TSDR locus of naïve $CD4^{+}CD25^{-}$ T cells that were cultured in different conditions, and naïve and activated $CD4^{+}CD25^{+}Foxp3^{+}$ nTregs. Here, CD3/28 TCR activator beads were used to activate cells. IL-2 (30 U/mL) and TGF-β1 (5 ng/mL) were added into the relevant culture conditions as shown in the legend of the graph and heatmap. **(A)** Column bar graph showing the percentage of methylation of 4 different CpG regions within mouse Foxp3-TSDR gene. Two-way ANOVA was used for statistical analysis using GraphPad Prism software 8.0.1. ($p \leq 0.0001$ (****), $p \leq 0.0056$ (**)). **(B)** Heatmap showing the percentage of methylation of 4 different CpG regions within Foxp3-TSDR locus. Heatmap was drawn using GraphPad Prism software 8.0.1. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

2.4. Discussion

The immune tolerance is one of key features of the immune system, which recognises and eliminates autoreactive immune cells that induce immune reactions against self-tissue. The T regulatory cells (Tregs) are professional cells with suppressive activity and employed by the immune system to maintain the immune tolerance by inhibiting and eliminating autoreactive immune cells. There are two main population of Tregs: nTregs and iTregs. nTregs are found to develop within the thymus and function to govern central tolerance. Whereas iTregs are generated in the periphery and function to maintain peripheral tolerance. Both subpopulations share similar phenotypic features by the expression of Foxp3 and CD25, however, they differ in the origin of development and functional stability within the local microenvironment. nTregs are originally generated as immunosuppressive cells by the immune system, whereas iTregs are generated from niave $CD4^+$ T cells in the periphery under certain conditions in the presence of cytokines such as TGF-β and IL-2. Therefore, in this study, a mouse model was developed to study the difference between nTregs and iTregs.

To develop the mouse model, both mouse strains; BALB/c and C57BL/6 were analysed and compared to select the appropriate strain for generating and studying iTregs. The BALB/c mice strain was selected to establish the model since it had significantly more nTregs (12% of total $CD4^+$ T cells) in the spleen, compared to C57BL/6 mice (~ 3% of total $CD4^+$ T cells). Similar findings were also found by a study conducted by Chen et al. the authors demonstrated that the percentage of nTregs in the lymphoid organs of BALB/c mice was significantly higher than in C57BL/6 mice. Also, nTregs of BALB/c mice were more suppressive than those of C57BL/6 mice (Chen, Oppenheim and Howard 2005). Moreover, Volgelsang et al. have demonstrated that BALB/c mice shown to have significantly more nTregs and plasmacytoid dendritic cells (pDCs) in the blood and spleen than C57BL/6 (Vogelsang, et al. 2009a). Another study also has reported that $Foxp3^+$ Tregs of BALB/c mice were significantly more functional to suppress IL-9-driven mast cells than those of C57BL/6 (Blankenhaus, et al. 2014a).

In this model, the results showed that the induction of iTregs was significantly enhanced following TCR activation of naïve $CD4^{+}CD25^{-}$ T cells in the presence of IL-2 and TGF-β1. However, there was no induction in the presence of IL-2 and TGF-β1 without TCR activation, highlighting the importance of TCR signals for initiating iTregs induction in the presence of TGF-β1 and IL-2. Vahl et al. have found that TCR signals are essential for the generation and activation of peripheral Tregs. Also, the authors have demonstrated that TCR-deficit Tregs were unable to supress T cells responses *in vivo* (Vahl, et al. 2014). Moreover, Vaeth et al. have found that suboptimal TCR signals were essential for the generation of iTregs from naïve $CD4^{+}CD25^{-}$ T cells by promoting the expression of Foxp3 via NFAT (nuclear factor of activated T cells) pathway (Vaeth, et al. 2012). Also, Ruan et al. have confirmed the importance of TCR stimulation in inducing Foxp3 expression of iTregs via activation NF-kB pathway in which c-Rel could bind Foxp3 promoter and then initiate iTreg differentiation (Ruan, et al. 2009). Taken together, the TCR engagement seems to play a crucial role in the development and differentiation of iTregs *in vitro* and *in vivo* by mediating the downstream signals for inducing the expression of Foxp3.

However, TCR signals alone were not sufficient for generating iTregs as the induction of iTregs following TCR activation in the presence of IL-2 did not occur. Whereas the induction of iTregs was significantly increased following adding TGF-β1, confirming the decisive role of TGF-β1 signalling in the development of iTregs. Li et al. have supported these findings. The authors demonstrated the critical role of TGF-β1 by modulating the expression of Foxp3 of iTregs, and how TGF-β1 antagonised TCR signalling via p38 pathway for maintain the expression of Foxp3 of iTregs. The author also demonstrated the mutual effort of both TCR and TGF-β signalling in controlling the methylation of Foxp3 locus during iTregs generation. (Li, Ebert and Li 2013). Recently, a study has demonstrated that the importance of both TGF-β receptor and IL-2 receptor in mediating Foxp3 expression via CNS2-5-azacytidine (5-aza-C) mechanisms (Freudenberg, et al. 2018). Another mechanism of TGF-β-induced Foxp3 expression was found to be mediated by the activation and recruitment of Smad2/3, a downstream transcription factor complex of TGF-β signalling, to a Foxp3 enhancer element which induces the expression of Foxp3 during iTregs development (Zheng, et al. 2010a, Tone, et al. 2008).

The results also showed that following receiving the same strength of TCR signals, the induction of iTregs from naïve $CD4^+CD25^-$ T cells was significantly increased by increasing the concentration of TGF-β1, confirming that TGF-β signalling could predominantly antagonise TCR signalling and mediate the Foxp3 induction. Moreover, the induction of iTregs was significantly reduced following prolonged TCR activation in the absence of TGF-β. These results confirmed the dominant role of TGF-β for generating iTregs in the presence of strong TCR signals, and the dominant role of TCR signals for resisting iTregs induction in the presence of TGF-β1 at lower concentrations. These finding also were supported by Li et al., as mentioned above (Li, et al. 2013). Another study has also found that CD3/28 co-stimulation could overcome TGF-β-mediated repression of the proliferation of activated T effector cells. This means that after persistent CD3/28 co-stimulation, effector activated T cells become more resistant to the influence of TGF-β signalling (Koehler, et al. 2007). These results have also been supported by the finding that activated T cells following CD28 engagement downregulate the expression of TGF-β receptor II (TGF-βRII), thereby bypassing the effect of TGF-β signalling (Sanjabi, et al. 2017).

The results also revealed that the induction of iTregs was significantly increased in the presence of 4T1 tumours cells, compared to those induced in the presence of TGF-β1 only. However, 4T1 cells could convert only ~ 3% of naïve $CD4^+CD25^-$ T cells to $CD4^+CD25^+Foxp3^+$ iTregs in the absence of TGF-β1. These results suggested that 4T1 cells might secrete insufficient amount of TGF-β1, so that could not induce iTregs or 4T1 cells might need a cellular interaction with other mediator cells such as dendritic cells and macrophages for inducing iTregs *in vitro*. Sun et al. have demonstrated the crosstalk between M2-macrophages and laryngeal squamous cancer cells for inducing iTregs. The authors have confirmed that cancer cell-activated M2-macrophages were able to induce iTregs from naïve $CD4^+CD25^-$ T cells. Moreover, it was found that M2-marcrophages-derived iTregs were able to distort the differentiation of monocytes towards M2-macrophages, thereby establishing a positive-feedback cycle (Sun, et al. 2017). These findings have also been found within the tumour microenvironment of nasopharyngeal carcinoma where tumour cells predominantly could induce M2-macrophages from monocytes via TGF-β1 and

IL-10, which leaded to the development of M2-macrophages-derived iTregs (Wang, et al. 2017a).

The results in this chapter also demonstrated that the percentage of methylation of 4 CpGs within Foxp3-TSDR of iTregs was significantly lower than those of naïve and activated $CD4^{+}CD25^{-}$ T cells. However, iTregs were significantly hypermethylated compared to naïve inactivated nTregs. These results were also found in a study conducted by Schmidt et al. where the authors assessed the methylation status of 15 different CpG regions within human Foxp3 gene locus and found that iTregs were significantly hypermethylated compared to unstimulated naïve nTregs. However, the iTregs and stimulated $CD4^{+}CD25^{-}$ T cells were similar in the methylation status (Schmidt, et al. 2016). Compared to the results in this model, however, the generated iTregs were significantly less methylated compared to activated $CD4^{+}CD25^{-}$ T cells. This also was demonstrated by the stability of the expression of Foxp3 of iTregs that were generated in this model. The results revealed that ~ 95% of the Foxp3 expression of iTregs was stable and maintained after 5 days of induction in the presence of IL-2 (30 U/mL) without re-activated with CD3/28 beads and TGF-β1. Although the demethylation status of Foxp3-TSDR is considered as a marker for the stability of Foxp3 expression, particularly in naïve nTregs, iTregs without TSDR demethylation could maintain their Foxp3 expression and suppressive activities (Hippen, et al. 2011, Lee, Lydon and Kim 2012, Gu, et al. 2014).

Interestingly, the results showed that naïve nTregs lost their demethylation status and became significantly hypermethylated following TCR activation. These results were also supported by Zhang et al.; where the authors demonstrated that nTregs lost their Foxp3 expression after TCR engagement and experienced re-methylation of the CNS1 region within Foxp3 locus (Zhang, et al. 2017). These findings were also found in an earlier study by Bailey-Bucktrout et al. in 2013. There the authors found that nTregs underwent Foxp3 instability and secreted pro-inflammatory cytokines following self-antigen-driven activation during inflammation in the CNS (Bailey-Bucktrout, et al. 2013). Taken together, it seems that both iTregs and nTregs can lose their stability *in vivo* after activation. This may stand with the idea that the immunosuppressive activity acquired by Tregs is induced upon the presence of stressful stimuli that promote prolonged acute or chronic inflammation, which can

make Tregs behaving with more flexibility based on the physiological condition in the milieu.

In summary, the BALB/c mouse model was successfully developed to generate and study the iTregs. In this model, iTregs ($CD4^+CD25^{++}Foxp3^+$) were generated *in vitro* from purely sorted $CD4^+CD25^-Foxp3^-$ T cells in the presence of TGF-β1 (5 ng/mL) and IL-2 (30 U/mL), and in the presence and absence of 4T1 tumour cells for 5 days. $CD4^+CD25^-Foxp3^-$ T cells were activated with CD3/CD28 TCR activation beads during the culture period for maintaining the induction of iTregs. The percentage of iTregs induction was 55-70% of total $CD4^+CD25^-Foxp3^-$ T cells after 5 days. The induction of iTregs was enhanced in the presence of both TGF-β1 and 4T1 tumour cells, confirming the role of tumour cells in the induction of iTregs. The iTregs were successfully generated by different concentrations of TGF-β1, while 0.05 ng/mL of TGF-β1 was the lowest concentration required for inducing iTregs from naïve $CD4^+CD25^-$ T cells. This study has also revealed that the generation of iTregs is based on a simultaneous synergy of TCR/CD28 and TGF-β1 signalling since activated $CD4^+C25^-$ T cells were less responsive to the effect of TGF-β1. The generated iTregs showed potent immunosuppressive activities by inhibiting the proliferation of $CD8^+$ T cells. In this model, the generated Tregs could maintain their Foxp3 expression for 5 days in the absence of TGF-β1, although their Foxp3-TSDR was hypermethylated compared to naïve nTregs. Finally, the percentage of iTregs induction and the percentage of nTregs in this mouse model was sufficient for profiling using mass spectrometry-based proteomics. Based on these results, optimisation for MS-based proteomic profiling of iTregs, nTregs and other subpopulation of naïve $CD4^+$ T cells will be discussed in the next chapter.

Chapter 3.

Optimisation of sample preparation for proteomic analysis

3.1. Introduction

Advances in 'omic' technologies, which include genomics, transcriptomics, proteomics and metabolomics, have contributed to the discovery of biomarkers that are currently used in the clinic as screening, diagnostic, prognostic and predictive biomarkers. Also, these technologies are increasingly being used in drug discovery and evaluation of their potency and toxicity (Quezada, et al. 2017). The omic based high throughput platforms have enabled the scientists to understand the relationship between the genotype and the phenotype in a scale which was never possible few decades ago. However, each of the above-mentioned technologies has its own advantages and disadvantages which will be discussed in the following section.

3.1.1. Profiling techniques used for biomarker identification

3.1.1.1. Genomics

Genomics is the comprehensive study of an organism's genome that includes the total DNA and its genes. There are several levels at which genome information can be derived such as structural and numerical alterations of chromosomes and sequence alterations at genome level. Chromosomal abnormalities have found to contribute to the emergence of many developmental diseases and malignancies (Theisen and Shaffer 2010). Numerical chromosomal abnormalities are characterised by the gain or loss of a single chromosome (Russo, et al. 2015). Structural abnormalities occur when there is an alteration in the structure of the chromosome arms and involve deletions, duplications, translocations, amplifications and inversions (Grade, Difilippantonio and Camps 2015). Advances in genome-based techniques such as next-generation sequencing, single-nucleotide polymorphism array, fluorescence in situ hybridisation, microarray-based comparative genome hybridisation and spectral karyotyping have contributed to the detection of chromosomal aberrations and mutations that cause genetic

diseases and malignancies (da Silva, et al. 2017, Ratan, et al. 2017, Braun, et al. 2019, Aleksandrova, et al. 2016, Song, et al. 2017). These chromosomal aberrations and mutations have been used as diagnostic biomarkers for several types of cancer including breast cancer, multiple myeloma, bladder cancer, endometrial cancer, colorectal cancer, cervical cancer, prostate cancer, renal cell carcinoma and acute lymphoblastic leukaemia (Yap, et al. 2015, Vargas-Rondón, Villegas and Rondón-Lagos 2017).

3.1.1.2. Transcriptomics

Transcriptomics is the global analysis of an organism's transcriptome that consists of all RNA transcripts expressed within cells. The transcriptome is the initial product of gene expression and plays as a biological bridge between DNA and functional proteins, the final products of the genome (Horgan and Kenny 2011). Different techniques have been developed for analysing gene expression including serial analysis of gene expression (SAGE), cap analysis of gene expression (CAPE), gene expression microarrays and RNA-sequencing (RNA-Seq). RNA-Seq is the most recent technology developed which superseded conventional array-based platforms and now widely adapted for transcriptome profiling with better dynamic range and accuracy. (Lowe, et al. 2017).

Transcriptomic profiling has been extensively used for biomarker discovery in many diseases such as cardiovascular diseases, tuberculosis, autoimmune diseases and cancer (Pedrotty, Morley and Cappola 2012, Van Rensburg and Loxton 2015, Ostrowski, et al. 2018, Sager, et al. 2015). However, information derived from transcriptomic profiling reflects the level of RNAs abundance, not proteins. Therefore, the interpretation of transcriptomic data is difficult to predict the phenotypic feature of a cell or tissue (Misra, et al. 2018). This limitation has been solved by the integration of genomic and transcriptomic datasets with data generated from proteomics and metabolomics that provide information about epigenetic and phenotypic variations (Misra, et al. 2018).

3.1.1.3. Proteomics

Proteomics is the field of studying the "proteome" that is defined as the collection of all proteins expressed within a cell, tissue or organism. One of the key aims of proteomics is to characterise phenotypic signatures of a cell or tissue through

proteins pathways and networks and describe the structure and function of proteins on a global scale (Theodorescu and Mischak 2007, Petricoin, et al. 2002). Like the transcriptome, the proteome is also dynamic in nature and can be easily altered or modified according to the condition of the biological microenvironment surrounding the cell or tissue. Proteins are the primary targets of most approved drugs, confirming the importance of identifying proteins that may have the potential for drug-target (Bull and Doig 2015) and also act as the mean for the most FDA-approved tests in the clinic, particularly plasma protein targets that constitute 45% of FDA-approved tests (Anderson 2010). However, the complexity of the proteome renders various challenges for comprehensive analysis (Merrick, et al. 2011). Proteins are more complex in nature than nucleic acids as they undergo post-translational modifications that are required to maintain the bioactivity of proteins within a cell or tissue. These modifications are one of the major reasons for the increased complexity of the proteome in any organism (Seo and Lee 2004). These complexities have contributed to the development of various proteomic platforms that facilitate the quantification and identification of different types of proteins and its post-translational modifications with comprehensive data analysis. These platforms will be described later in the following section (3.1.2).

3.1.1.4. Metabolomics

Similar to the transcriptomic characterisation of cellular RNAs and the proteomic characterisation of entire proteins expressed at a given time, cells can also be analysed for detecting metabolites that are final-end products of complex biochemical pathways that link genotype to phenotype (Goodacre, et al. 2004). The study of total low-molecular-weight compounds (metabolites) within a cell or organism is called “metabolomics” (Kamburov, et al. 2011). Theoretically, metabolomics has more advantages over the other omics technologies. First, since metabolites are the final downstream products, changes in the level of metabolites expression are correlated with changes in the transcriptome and proteome. Second, the cellular metabolome reflects a functional signature of the cellular proteome (Urbanczyk-Wochniak, et al. 2003, Fiehn 2002). The evolution of metabolomic technologies such as mass spectrometry (MS), high-resolution nuclear magnetic resonance spectroscopy (NMR), ultra-performance liquid chromatography (UPLC) and bioinformatics has enabled more comprehensive quantification of the

metabolome (Tsutsui, et al. 2011). Also advances in mass spectrometry have improved the accuracy, reliability and efficiency of metabolic profiling (Liang, Wang and Li 2015). Metabolomics-derived biomarkers can also be used as targets for clinical laboratory tests. For instance, recently, metabolic profiling using 1H NMR has been carried out to identify biomarkers that can predict renal failure in patients who received contrast medium before undergoing angiography. This analysis has identified a panel of six urinary metabolites that can predict renal damage including glutamic acid, uridine diphosphate, glutamine and tyrosine (Dalili, et al. 2019). Also, another study has analysed metabolomic changes in tissues derived from patients with adenocarcinoma lung cancer and squamous cell lung carcinoma. This study has identified 851 metabolites that were aberrated in lung cancer tissue compared to the normal tissues (Moreno, et al. 2018). Such studies have confirmed the importance of metabolomic profiling to identify robust biomarkers that can be used clinically for monitoring patients. However, metabolomics approach suffers from several challenges that make it more complex than other omics approaches. First, the cost of experiments and instrumentations as metabolomic profiling requires multiple analytical platforms to study the complexity of biochemical pathways (Wishart 2016, Pinu, et al. 2019). Second, metabolites are present with different molecular weights and lower concentrations and varied in their stability and turnover rates within cells, thereby measuring them using a single universal analysis method is difficult (Beale, et al. 2018, Fiehn, et al. 2007). Third, the accuracy and reproducibility of metabolite quantification is another challenge due to the variations of sample extraction and different instrumental standardisation steps (Broadhurst, et al. 2018). Finally, the validation of a metabolite fingerprint is also challenging due to the complexity of metabolites networks within cells (Zhang, et al. 2015).

3.1.2. Mass spectrometry-based proteomics

In general, proteomics is used to link genotype to phenotype through a large-scale determination of gene and cellular function at the protein level. The complexity of the cellular proteomes and the presence of many low-abundant proteins require highly sensitive techniques for proteomic analysis (Tyers and Mann 2003). Mass spectrometry (MS) has been evolved for analysis of complex protein samples. MS-

based proteomics has proven itself as a powerful technology to characterise the structure and function of proteins. To date, proteomic analyses such as protein primary sequencing, post-translational modifications, protein-proteins interactions and secretome (proteins that are secreted extracellularly) have been carried out using MS-based instruments (mass spectrometers) (Mishra 2010). In general, mass-spectrometers have three main parts (figure 3.1). The first is the ***ionisation source***, which converts peptides or proteins mixture into positively and negatively charged ions based on the nature of a mixture. The second is the ***mass analyser***, which sorts and separates ions based on their mass to charge (*m/z*) ratio. The third part is the ***detector***, which detects the ions sorted by the mass analyser (Lemière 2001). The development of soft-ionisation methods such as **MALDI** (matrix assisted laser desorption ionisation) and **ESI** (electrospray ionisation) has enabled the proteomics field to accurately measure peptide masses and identify their sequences.

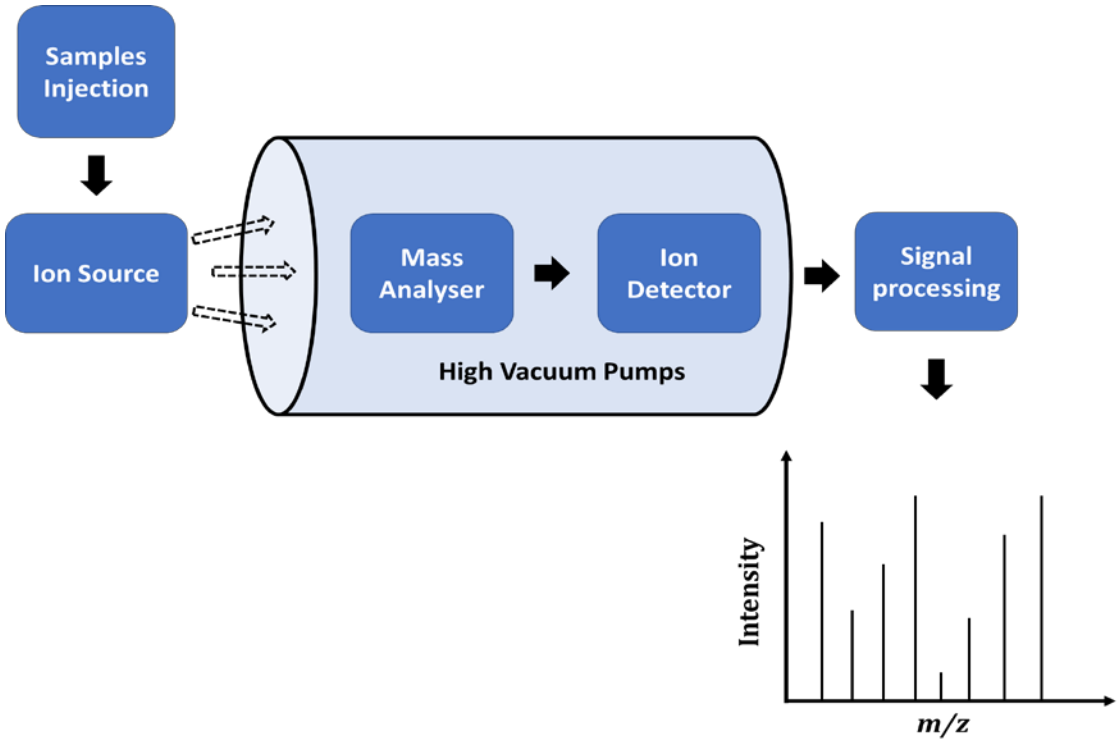


Figure 3.1. Schematic representation of the basic structure of a mass spectrometer. Upon injection, samples (peptides) are directed to the ion source where they become ionised with positive and negative charge based on their nature. After ionisation, peptide ions are eluted to travel through a high-pressure atmosphere to reach the mass analyser which sorts and separates peptide ions based on their *m/z* ratios. The sorted ions continue flying in a speed based on their *m/z* ratios to finally hit the ion detector which produces a signal which represents the time of flight for each ions to hit the detector. The signals are then processed and recorded as peaks or mass spectra.

3.1.2.1. MALDI-TOF MS Instruments

MALDI-TOF instruments represent the conventional MS technology which are generated to measure peptide masses. These instruments consist of the ionisation source "MALDI" and the mass analyser; time of flight (TOF). In the MALDI source, the sample of interest is mixed with matrix chemical compounds such as 2,5-dihydroxybenzoic acid, 3,5-dimethoxy-4-hydroxycinnamic acid and α-cyano-4-hydroxycinnamic acid, in addition to acetonitrile and water (Liebler 2001). The admixture of sample and matrix compounds is then loaded as spots onto a target plate and allowed to dry in air. Once dried, spots usually look like a powdery deposit or crystal lattice into which the peptide or protein sample is combined. The target plate is then placed into the source equipped with a laser which produces a beam of light at the target (figure 3.2). Once spots of sample are introduced to the laser beam, matrix compounds become electronically excited and transfer this energy to the peptides or proteins in the sample which are then expelled from the target plate into the gas phase(El-Aneed, Cohen and Banoub 2009). Each peptide can pick up a single proton, and therefore most of peptide ions produced by MALDI are singly charged ions and defined as $[M+H]^+$ ions; M is the actual mass of a peptide or protein and H is the proton (charge) bound to the peptide or proteins. Then the positive singly charged ions ($[M+H]^+$ ions) are directed to the TOF mass analyser (Demeure, Gabelica and De Pauw 2010) (figure 3.2).

The mass analyser measures the time needed for the ions to fly from the beginning to the end of the analyser until they hit the detector. The speed of ions flying through the analyser is proportional to their *m/z* values. Ions with high *m/z* values strike the detector faster than those with low *m/z* values (Demeure, et al. 2010). TOF analyser operates in two different modes; linear and reflectron mode. Linear mode is where ions are produced from the source and fly through the analyser tube until reach the detector. However, TOF analysers with linear mode are not sensitive as the resolution of MS analysis is very poor. The resolution in MS is defined as the ability of the instrument to discriminate between ions that are slightly different in their *m/z* values. The poor resolution in linear-mode TOF is due to differences in the velocities of ions having the same *m/z* values as they fly, which leads to create peak broadening during analysis (Kinter and Sherman 2005, Demelbauer, et al. 2004). In contrast, a TOF analyser operating in reflectron mode shows better resolution

where the reflectron acts as a speed controller for ions flying through the analyser, and therefore ions of the same *m/z* values are focused by the reflectron to reach the detector at the same time. This leads to better resolution as sharp, narrow peaks will be produced from analysis (Demelbauer, et al. 2004). Given the development of MALDI-TOF instruments with different modes, they have some limitations. First, these instruments are designed to measure peptide masses, and this type of information is not useful in obtaining true sequence data for peptides or proteins. Second, MALDI-TOF instruments are not linked directly with separation systems such as HPLC (high performance liquid chromatography) which separate contaminants such as urea, glycol, and Tris, and therefore the resolution and sensitivity of the analysis will be very poor (Byrd and McEwen 2000, Müller, et al. 2001, Liebler 2001), although offline chromatographic spotting/fractionation can overcome this it adds time and complexity to the analysis .

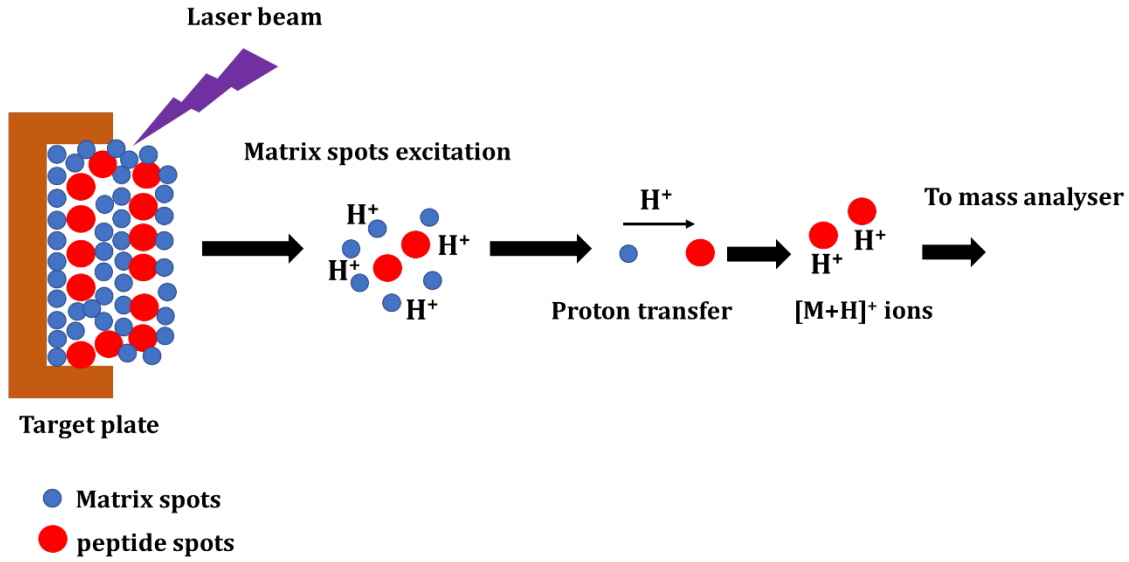


Figure 3.2. Schematic representation showing the principle of MALDI ionisation method. The peptides samples are mixed with the matrix compounds. The mixture of peptides and matrix compounds is then loaded into the target plate as spots. The target plate is then introduced into the source of ionisation supplied with a laser. Upon activation, the laser produces a beam directly to the target plate which makes matrix spots become electronically excited by acquiring protons (H^+). Excited matrix spots transfer protons to the peptides spots to produce positively-charged peptide ions ($[M+H]^+$ ions). The $[M+H]^+$ ions are then directed to the mass analyser for MS analysis.

3.1.2.2. ESI Tandem MS Instruments

ESI tandem MS or (ESI-MS-MS) instruments have been evolved to improve the sensitivity and reproducibility of proteomic analysis. ESI is electrospray ionisation, the source of ions production in the instrument. Tandem mass spectrometry (MS-MS) refers to mass analysers that are designed to carry out two- or multi-stage mass analyses of generated ions (Pitt 2009). The main principle of ESI method is the generation of multiple charged ions from a mixture of peptides or proteins. In this method, samples in liquid form are injected into the source via a flow stream that comes often from the HPLC linked with the source. Samples then enter a capillary, which has a shape of a cone, supplied with high electric voltages on its tip (figure 3.3). As the samples exit the capillary, they are sprayed out in a fine mist of droplets. The continuous application of electric voltages to the capillary's tip leads to the formation of charged droplets. These droplets are in a state of solvation and contain peptides or proteins in addition to ingredients (e.g. water, acetonitrile and acetic acid) from the HPLC mobile phase. All these steps occur under an atmospheric pressure (Ho, et al. 2003). The charged droplets (solvated droplets) are then dispersed from the capillary cone to the mass analyser which is held in high vacuum region. Before entering the mass analyser, ions of peptides of proteins will be purified and separated from the HPLC solvents by the source in a process called "ion desolvation". Ion desolvation can be achieved in different ways based on the type of the source. In some sources, the droplets pass through a heated capillary which assists evaporation the solvents (Banerjee and Mazumdar 2012). In other sources, a curtain of nitrogen gas cross through against the direction of the droplets' flow, which facilitates evaporation of the solvents and reduction in the size of the droplets. In both ways, the peptides or proteins ions are released from the source into the mass analyser, whereas the solvents from the droplets are bumped away by the vacuum system (Janusson, et al. 2015) (figure 3.3).

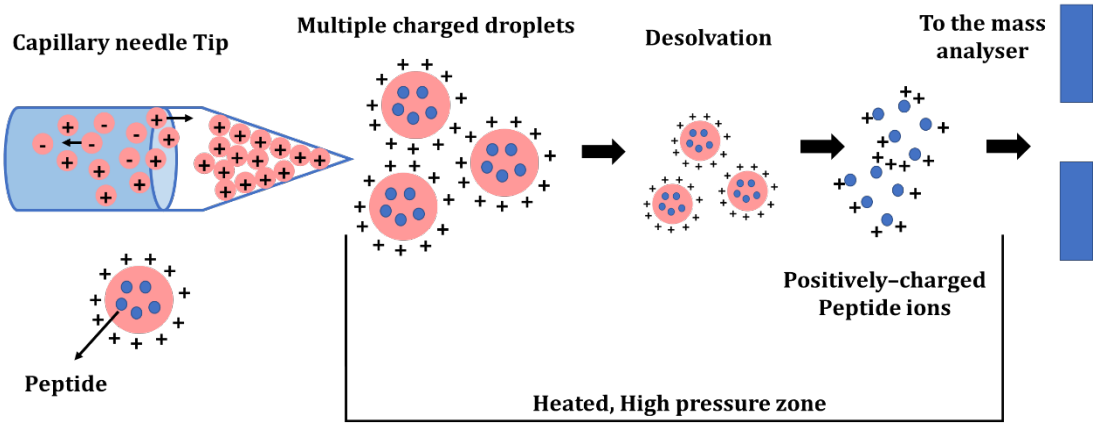


Figure 3.3. Schematic representation showing the principle of ESI ionisation method. The peptide samples are injected into the ionisation source via a flow-stream coming from the HPLC that is linked with the source. In the source, samples enter a capillary which has a shape of a cone supplied with high electric voltages on its tip where peptide ions become positively charged. The samples are then ejected from the tip in a fine mist of droplets with multiple charges. The multiple charged droplets contain peptides and other ingredients such as water and acetonitrile. These charged droplets are directed into a heated, high pressure zone in a process called desolvation that results in a reduction the size of droplets and liberation of free positively-charged peptides ions. The peptide ions are then directed into the mass analyser for MS analysis.

ESI-MS-MS instruments have been designed with different types of mass analysers including TOF, quadrupole, ion trap, Fourier transform mass spectrometry, Fourier transform-ion cyclotron resonance, orbitrap. Other instruments have hybrid mass analysers which are a combination of two or three mass analysers to obtain better performance. These instruments include triple/tandem quadrupole, quadrupole-TOF (Q-TOF), TOF-TOF, linear trap quadrupole-orbitrap (LTQ-orbitrap) and others. Each of these instruments has specific features for ions analysis, and therefore the use of such instruments is based on the proteomic approach required to sequence and characterise proteins of interest (Matthiesen and Bunkenborg 2013). All these analysers are the best examples of MS-MS technology which is now indispensable for any kind of proteomic approaches. There are three primary steps in any MS-MS mass analysers. First, selection of precursor ions (desolvated ions) that are produced by the ESI source. This step is mostly called "ions scanning or filtration" in which ions with a specific *m/z* value will proceed to the next stage of analysis. This is usually achieved by quadrupole analysers, specifically Q1, and ion trap analysers. Second, fragmentation of the selected precursor ions to generate fragment or

product ions. This step is achieved by a process called "collision-induced dissociation (CID)". This process occurs when the selected precursor ions are subjected to collide with "neutral collision gas", which is usually helium, nitrogen or argon, in a region called collision cell (Dawson 2013). This region of the analyser is usually supplied with radiofrequency voltages that facilitate fragmentation of the selected precursor ions by applying kinetic energy which exceeds the internal energy of the precursor ions, leading to dissociation of precursor ions into small fragments (product ions). Third, these product or fragment ions are further analysed based on their *m/z* values to generate a spectrum for each product ions. This MS-MS spectrum is usually used to deduce the sequence of a peptide (Matthiesen and Bunkenborg 2013).

3.1.2.3. Proteomic approaches for sequencing and characterisation of proteins

Sequencing of proteins can be accomplished either by analysing proteins using their intact from or by digesting them. Although modern MS instruments are still able to measure the molecular weight of intact proteins, it is still very difficult to sequence them using their intact structure, especially high molecular weight and hydrophobic or complex protein mixtures (Chait 2006). Also, the sensitivity of analysing intact proteins is not as high as sensitivity of analysing peptides. For this reason, proteomic analysis using digested proteins is more commonly carried out than using intact proteins (Feist and Hummon 2015).

There are three main proteomic approaches to sequence and identify proteins, including bottom-up approach, top-down approach and middle-down approach.

3.1.2.3.1. Bottom-up (BU) approach

BU approach is the most common technique used to identify proteins and their amino acid sequences by digesting proteins into small peptides prior to analysis using MS-based instruments. With regard to digestion, the proteins are either introduced to enzymatic proteolysis (e.g. Trypsin, Glu-C, Lys-C, Asp-N and Chymotrypsin) or chemical proteolysis (e.g. cyanogen bromide) (Kinter and Sherman 2005, Liebler 2001) . Each of these proteases can cleave proteins at specific amino acid residues. BU is a peptide-based approach for identification of a proteins, which means that complete digestion of proteins is ideally required to obtains

numerous fragments of peptides with length of ~ 7-20 amino acids, which is the ideal length for MS analysis. By this strategy, the identity of a protein is deduced by unequivocal detection of peptides that have unique sequences for that protein (Xie, Smith and Shen 2012). However, obtaining complete digestion of proteins is not straightforward as it depends on the complexity of the sample that will be analysed and the abundance of specific amino acid residues in the structure of proteins. For example, trypsin, which is the most common protease enzyme used for protein digestion in BU approach, cleaves proteins at the carboxyl side (C-terminal side) of the lysine and arginine amino acid residues. However, trypsin does not cleave proteins if the C-terminal of the lysine and arginine amino acids is followed by Proline residues (Olsen, Ong and Mann 2004). Therefore, the number of peptides fragments resulting from the digestion of a protein will be based on the number and site of lysine and arginine residues of that protein and whether the lysine and arginine are linked with Proline. (Tsiatsiani and Heck 2015). As a consequence, limited sequence coverage of proteins is a major obstacle in BU approach due to the incomplete digestion of proteins, which hampers detection of post-translational modifications and isoforms of proteins (Catherman, Skinner and Kelleher 2014).

3.1.2.3.2. Top-down (TD) approach

The TD approach is referred to identification and sequencing of proteins using their intact structure without introducing to any kind of digestion. The TD approach has been found more useful for studying and detecting post-translational modifications and isoforms of proteins than BU approach (Savaryn, et al. 2013). However, applying MS-based proteomics is relatively limited using a TD approach, although trials have been done to develop a realistic method for sequencing and characterising proteins using MS-based TD analysis (Toby, Fornelli and Kelleher 2016).

3.1.2.3.3. Middle-down (MD) approach

The MD approach is an emerging method for sequencing proteins by combining the principles of both BU and TD approaches. MD approach has been suggested and invented based on the advantages and disadvantages of BU and TD approaches (Garcia 2010). By MD strategy, digested proteins would be analysed instead of intact proteins by introducing them to enzymatic and chemical proteolysis as mentioned in BU approach. However, in MD approach, incomplete digestion of proteins is

typically required, compared with BU where complete digestion of proteins is needed. Therefore, the incomplete digestion would produce polypeptides fragments with length of ~20-100 amino acids, which are significantly longer than peptides resulted in BU approach (Cristobal, et al. 2017). Therefore, the number of peptides generated by MD approach would be fewer than the number of peptide produced by BU approach, which means that the complexity of the peptides mixture resulted in the MD approach would be significantly lower than the peptide mixture resulting from the BU approach (Pandeswari and Sabareesh 2019). Thus, the percentage of sequence coverage would be increased high enough to identify proteins and characterise their post-translational modifications (PMTs) and isoforms using MD compared to the BU approach where detection and identification PMTs and isoforms of proteins is very limited (Pandeswari and Sabareesh 2019).

3.1.2.4. Separation methods of peptides/proteins prior to MS analysis

Due to the complexity of proteomic samples that include a complex mixture of peptides or intact proteins, several methods have been used to reduce the complexity of proteomic samples, which allow the MS instruments to analyse maximum number of peptides and generate useful information for identifying proteins, especially low-abundant proteins (Shi, et al. 2004, Fonslow and Yates Iii 2009) . The separation methods can be classified into two main methods; gel-based and gel-free method. Gel-based methods include one-dimensional and two-dimensional gel electrophoresis that can separate proteins based on their isoelectric points and molecular weights, respectively (Baggerman, et al. 2005). Gel-free methods include liquid chromatography methods such as reverse phase high performance liquid chromatography (HPLC), ultra-high performance liquid chromatography (UHPLC) and multi-dimensional liquid chromatography (Mitulović and Mechtler 2006). In liquid chromatography, the degree of separation of peptides or intact proteins is based on the features of mobile phase (solvents) and the stationary phase (separation column). The mobile phase facilitates elution of analytes (peptides or proteins) that are attached to the stationary phase. The analyte's elution by the mobile phase is correlated with the strength of interaction between the analytes and the stationary phase (Pandeswari and Sabareesh 2019). Thus, the selection of columns and solvents is highly important to achieve great separation of complex mixture of peptides or intact proteins. The HPLC method is

mostly used in proteomic studies, where the mobile phase or solvents are water, acetonitrile and methanol (Ali, et al. 2010). These solvents are compatible to ESI-MS-MS instruments, and for this reason the HPLC technology has been successfully linked with most of ESI-MS-MS instruments, which is denoted as LC-MS or LC-MS-MS proteomics (Arrivault, et al. 2009). Another method is also used for separation peptides or intact proteins, which is called capillary electrophoresis (CE) (Simpson and Smith 2005). CE can be classified either gel-based or gel-free methods based on several operation modes including capillary gel electrophoresis, capillary zone electrophoresis, capillary isoelectric focusing, capillary isotachophoresis and capillary affinity electrophoresis (Sun, et al. 2016, Storms, et al. 2004, Nesbitt, Zhang and Yeung 2008).

3.1.2.5. Shotgun and targeted proteomics

3.1.2.5.1. Shotgun proteomics

Shotgun proteomics is a discovery-based method that is widely used in proteomic studies to acquire a broad overview of the proteins expressed in a sample that would be analysed (McDonald and Yates 2002). The shotgun method is broadly utilised as a high-throughput tool for biomarker discovery in which potent clinical validation may or may not be accomplished. Although the shotgun method is highly efficient in identifying hundreds or thousands of proteins in any biological samples, it is rarely used to analyse specific molecules within a cell or tissue at different conditions. This means that shotgun method is not suitable for hypothesising or answering a specific question (Faria, et al. 2017). Rather, the shotgun method has already been designed to provide complex answers for somewhat non-specific questions. For example, identifying a significant number of proteins that are upregulated and downregulated in various experimental conditions can be achieved using the shotgun method, however, knowing the main cause for the up- and down-regulation of the proteins will be very limited (Martins-de-Souza, et al. 2010, Castagnola, et al. 2012) . Also, shotgun proteomics has a very limited potential of identifying low-abundant proteins because it is based on the analysis of fragmented peptides that are usually identified relative to the abundance of each protein in a sample, thus only the highly enriched proteins will be identified (Wang, et al. 2017b). However, this does not mean that the shotgun approach is irrelevant or time-consuming in nature.

This approach has opened the doors for other methodologies such targeted proteomics that can investigate the proteome in a selective way (Faria, et al. 2017).

3.1.2.5.2. Targeted proteomics

Targeted proteomics is a question-based or hypothesis-based method that is widely used in proteomic studies to answer specific questions more quickly and more accurately than classic shotgun approach (Marx 2013). Targeted approach is mainly used to monitor or analyse a selection of proteins that are already known beforehand. Therefore, by using a targeted approach, it is compulsory to select proteotypic peptides (PTPs) that have unique amino acid sequences which constantly determine a specific protein in a given proteome investigated by MS (Domon and Gallien 2015). There PTPs play as a marker or signature for the selected protein of interest and are tracked throughout the experimental run on MS. Thus, the selection of PTPs is the most critical step in the targeted approach, which requires in-depth knowledge about the protein of interest (Gallien, Duriez and Domon 2011). There are two main strategies used in targeted proteomics including selection reaction monitoring (SRM) or multiple reaction monitoring (MRM) and parallel reaction monitoring (PRM). In SRM, a panel of peptides is selected to quantify the protein of interest and analysed on a mass spectrometer, usually triple quadrupole mass analysers in which the first mass analyser is used a mass filter with a narrow isolation window to isolate a specific peptide precursor ion generated from the selected peptide (Ebhardt 2014). Then the isolated precursor ion is further fragmented to generated small fragment ions, and one of the generated fragment ions is monitored by the next mass analyser which is set to filter a specific *m/z* value (Picotti and Aebersold 2012a). The double selection of the peptide precursor ion and fragment ion by the tandem mass analysers is called a transition process. This transition is highly specific and produces highly sensitive measurements for proteins of interest. In contrast, PRM method (such as from a Q-TOF instrument) operates in a different way where all fragments ions generated from the pre-selected peptide precursor ion are analysed simultaneously (Borràs and Sabidó 2017). In PRM method, a peptide precursor ion of interest is filtered and selected by a narrow isolation window setting of the first mass analyser. The selected peptide precursor ion is then fragmented, and the resulting fragment ions are together analysed by the next mass analyser set to screening mode (Bourmaud, Gallien and

Domon 2016). Then chromatographic peaks resulted from fragment ions are analysed to determine the best fragment ions for peptide identification and quantification. Therefore, PRM is more flexible method than SRM because it can analyse a broad range of fragment ions instead of monitoring only one fragment ion, and generate highly sensitive and specific data with dynamic range for peptide and protein quantification (Rauniyar 2015, Borràs and Sabidó 2017) . Thus, the PRM method can be specifically used for monitoring low-abundant proteins (Faria, et al. 2017). However, the limitation of using targeted proteomics in the clinic is the requirement of extremely sensitive and specific biomarkers for targeted analysis (Faria, et al. 2017).

3.1.2.6. Methods of quantifying proteins

There are two main methods commonly used for protein quantification in MS-based proteomics, including label-based and label-free methods.

3.1.2.6.1. Label-based method

Label-based method is a comparison analysis of proteins quantification between samples that are labelled with different mass tags which allow detection of proteins based on specific change in mass. These tags share the same chemical structure and are different in their isotopic affinity (Domon and Aebersold 2010). There are three different types of labelling, including metabolic, chemical and enzymatic labelling (Zhang, et al. 2010).

3.1.2.6.1.1. Metabolic labelling

Metabolic labelling is one of the most favoured method of labelling because samples of interest are labelled at the early stage of sample preparation, which leads to decrease experimental variabilities between samples (Bantscheff, et al. 2007). By metabolic labelling, different samples are introduced to an isotopically defined medium containing distinct labels that can incorporate with the proteome of samples during the process of protein metabolism. The samples are then equally pooled together and analysed on MS (Iliuk, Galan and Tao 2009). This generates two forms of each peptide with different isotopic structure, but with identical reaction to sample preparation and MS analysis, thus allowing quantification of proteins of the labelled samples without affecting their biochemical features (Geiger, et al. 2011). Initially ^{15}N-modified media was used for labelling, in which all ^{14}N

(nitrogen) isotopes were replaced by ^{15}N isotopes (Oda, et al. 1999). However, using this media showed several limitations including partial or inadequate labelling of samples and difficulties in data analysis as well as higher costs (Anand, et al. 2017). A more effective labelling technique known as SILAC (stable isotope labelling of amino acids in cell culture) was developed by Ong et al. (Ong, et al. 2002). SILAC is a non-selective method of labelling proteins where heavy isotopes of specific amino acids are used as labels. By this technique, two different medium are used to grow two different populations of cells *in vitro*; one medium contains normal (light) amino acids and the second contains isotopically-labelled heavy amino acids (Chen, et al. 2015). This labelling is achieved by replacing the natural atoms of H, ^{14}N and ^{12}C to ^{2}H, ^{15}N and ^{13}C, respectively. Leucine, lysine, methionine and arginine are the most suitable amino acids have been used for effective isotopic labelling (Mann 2006). The labelled amino acids can then integrate into the newly synthesised proteins of the labelled sample. Then samples are mixed and prepared to be analysed by MS. The MS analysis can distinguish between samples by the molecular weight of the heavy and light isotopes of the specific amino acids that were used during cell culture (Harsha, Molina and Pandey 2008). There are key advantages of using SILAC including an effective labelling due to the high integration rate of labelled amino acids into the proteome of interest, efficient labelling without any chemical manipulation to the samples, and reducing experimental variabilities as the samples are labelled in the early steps (Ong, et al. 2006). SILAC has been broadly used to study PTMs, protein signalling pathways and enzyme substrates, as well as identification of cancer-related biomarkers (Chen, et al. 2015). A major limitation of using SILAC is that the number of heavy forms of amino acids used for labelling is very limited, so that only 5 different biological states can be analysed within a single experiment (Anand, et al. 2017).

3.1.2.6.1.2. Chemical labelling

Chemical labelling is also another labelling technique which is mostly used for protein quantification. It is a flexible labelling technique as the isotopic labels can be selectively introduced into any of desired position in a peptide or a protein of interest (Chahrour, Cobice and Malone 2015). Chemical labelling is relatively similar to metabolic labelling except that isotopic labels of a peptide or a protein are mediated by chemical reactions (Kainosho, et al. 2006). Isotope-coded affinity

tagging (ICAT) is of the earliest chemical labelling methods, which has been introduced by Gygi et al. (Gygi, et al. 1999). ICAT reagent consists of thiol-reactive iodoacetamide group that labels cysteinyl thiols residues in proteins, a linker region that contains heavy or light deuterium atoms and biotin tag used for affinity purification (Smolka, et al. 2001). The limitation of ICAT is that the rate of peptide recovery during purification is very low due to the presence of deuterium atoms in the linker region (Zhang, et al. 2001).

There are other techniques of chemical labelling including tandem mass tag (TMT) and isobaric tag for absolute and relative quantification (iTRAQ). These techniques are amine-based labelling in nature, which target only amine functional group of the peptides (Savitski, et al. 2013). Using these techniques, the proteins are essentially labelled with chemical groups which are equal in mass (isobaric) and are fragmented under the pressure of tandem MS to produce reporter ions in a various mass range (Evans, et al. 2012). TMT is designed to analyse a maximum of ten samples, whereas iTRAQ allows analysis of up to eight samples in a single experiment (Latosinska, et al. 2015, Zhang, et al. 2016) . TMT and iTRAQ have been widely used due to their flexibilities of analysing multiple samples with a relative accuracy of protein quantification (Huang, et al. 2017). However, there are some limitations to using TMT/iTRAQ including the requirement of MS instruments with higher resolving powers, higher costs, non-specificity of reporter ions for peptides due to the impact of peptide co-isolation process during the MS run, and inaccuracy between peptide quantity and identity (Karp, et al. 2010).

3.1.2.6.1.3. Enzymatic labelling

Enzymatic labelling is another labelling technique used for protein quantification. ^{18}O labelling technique is an example of enzymatic labelling which uses protease enzymes such as trypsin to integrate two ^{18}O atoms instead of ^{16}O atoms in peptides, making a shift in peptides mass (Reynolds, Yao and Fenselau 2002). However, it is not widely used in quantitative proteomics due to instability of enzymatic labels (Ramos-Fernandez, Lopez-Ferrer and Vazquez 2007, Zhang, et al. 2013) .

3.1.2.6.2. Label-free quantitation methods

Label-free quantitation methods have been widely used in proteomic studies as it provides easy options for global analysis of biological samples. Unlike label-based

methods, label-free methods allow analysis and comparison of many samples that are injected individually into MS (Wang, et al. 2008). Label-free approach is cost-effective, does not require expensive labelling reagents, and is not time consuming as compared to label-based methods (Megger, et al. 2014). All these advantages have allowed label-free method of the favoured approaches for large-scale protein quantification (Wang, et al. 2008) . Using label-free method, thousands of proteins can be identified from even complex samples such as blood, plasma, saliva, urine, cell line and tissues. Moreover, label-free approach produces highly sensitive MS analysis and is less vulnerable to technical or experimental errors (Megger, et al. 2013).

3.1.2.6.2.1. Spectral counting

Spectral counting that determines the relative abundance of a protein by measuring the frequency of peptide identified of a protein. In spectral counting, the amount and size of a protein is directly equivalent to frequency of peptide spectra identified that belong to a protein (Arike and Peil 2014). Protein abundance index (PAI) is also used in spectral counting method to measure the abundance of a protein in a given sample (Cutillas and Vanhaesebroeck 2007). PAI is defined by the number of identified peptides divided by the number of theoretically observable peptides for each protein. However, the spectral counting method is biased for MS analysis due to the frequent detection of peptides with various physico-chemical features (Zhang, et al. 2013).

3.1.2.6.2.2. Ion intensity

Ion intensity that measures the MS chromatographic signal intensity of peaks for peptides of a protein. This intensity-based method, which is also known as “the area under the curve” (AUC), quantifies proteins by integration and measuring of all chromatographic peak areas of identified peptides of a protein (Neilson, et al. 2011, Podwojski, et al. 2010) . However, several factors should be considered before using this method to ensure the accuracy and reproducibility of MS data analysis. These factors are co-eluting peptides that can affect peptides retention time, multiple signals of the same peptide, MS speed and sensitivity, and background noise due to chemical overlapping (Listgarten and Emili 2005, Cui, et al. 2013, Zhang, et al. 2013).

3.1.2.6.2.3. Data-dependent acquisition

Data-dependent analysis or acquisition (DDA) has also been used for protein quantification. DDA is a conventional mode of the most tandem MS instruments, where the first mass analyser (MS1) is set to run a scan survey for all peptide-precursor ions that are eluted from the LC column for multiple cycles (Wu and MacCoss 2002). In each cycle, the most abundant ions represented by "LC-MS peaks" are then recorded, selected and consecutively isolated for fragmentation by the second mass analyser to identify structural features and identity of proteins (Bateman, et al. 2014). MS1 quantification can be carried out in two different ways based on the method used for chromatogram processing. First, LC-MS feature-based MS1 quantification, where every LC-MS peak (feature) is detected and characterised from the MS1 data, and then all peaks are annotated using the peptide identifications from the MS2 data (Bilbao 2019). Examples of open source software packages implementing this method include SuperHirn (Mueller, et al. 2007), VIPER (Monroe, et al. 2007) and MzMine-2 (Pluskal, et al. 2010). Second, extracted ion chromatograms (XIC or EIC)-based MS1 quantification, this method is targeted analysis, where precursor ion signals from each confidently-identified peptide are extracted from MS1 data. Then XICs are listed to determine *m/z* targets of the confidently-identified peptides (Bilbao 2019). Skyline software package is used process the targeted MS analysis (Pino, et al. 2017, MacLean, et al. 2010) . However, DDA method is biased as only abundant peptides are selected for further analysis to identify and quantify proteins, whereas low-abundant peptides which may represent important low-abundant proteins are excluded from the analysis, thereby leading to irreproducible results specially for complex proteomic samples (Bateman, et al. 2014).

3.1.2.6.2.4. Data-independent acquisition

The limitations of DDA for protein quantification and identification have been overcome by the recent developments in hybrid MS instruments and data analysis software by which all precursor ions of peptides eluted from the LC-MS are fragmented and analysed regardless of their abundance levels and intensities in an operation mode called "data-independent acquisition (DIA)" (Hu, Noble and Wolf-Yadlin 2016). DIA is unbiased method which comprehensively and frequently analyses all peptide precursor ions of a given sample in a wide precursor isolation

window , generating a complex set of mass spectra (peaks) (Chapman, Goodlett and Masselon 2014). Also, DIA provides a broad range of detected peaks (signals) with different *m/z* values of detected ions, and therefore increases the coverage rate of peptides and proteins identification with highly sensitive and reproducible results (Bilbao, et al. 2015). In addition, the efficient performance of DIA has enabled label-free shotgun approach to be as a valid alternative to isotope-labelling-based methods (Navarro, et al. 2016). The aforementioned successes of DIA has also been applied in another more specialised proteomics approach; middle-down approach, wherein the size of proteolytic peptides is much larger than those resulting from bottom-up approach. Carvalho et al. have developed another version of DIA, which is called XDIA (extended data-independent acquisition) for protein quantification using middle-down approach. The authors have confirmed that applying XDIA has increased the number and accuracy of peptide spectra and improved protein sequence coverage using middle-down approach which targets large molecules and detects post-translational modifications and various isoforms of proteins (Carvalho, et al. 2010). However, one of the considerations of using DIA is the complexity of data yielded post MS analysis which contains ions spectra of all detected peptides. Accordingly, several informatics acquisitions have been evolved to process these sophisticated MS datasets, including SWATH-MS (Gillet, et al. 2012), DIANA (Teleman, et al. 2014)and Skyline (Pino, et al. 2017). Another obstacle of DIA in the increased probability of interference due to the overlapping of fragments ion yielded from co-fragmented precursor ions, which can mislead the identification and quantification of proteins (Zhang, et al. 2015). However, several computational approaches have been developed to handle this issue to further extend the benefits of DIA, including SwathTUNER and NOFI (non-outlier fragment ion) (Bilbao, Lisacek and Hopfgartner 2016).

3.1.2.7. Subcellular fractionation

The complexity of the proteome within organisms is due to the presence of various types of tissues in which each cell undergoes a specific-lineage differentiation and thereby acquires a specific functional specialisation. A cell is also compartmentalised in nature, which means that each cell contains different organelles (subcellular components) performing specific functions within a cell. This compartmentalisation provides distinct and appropriate environments for

synthesis, localisation and interaction of subcellular proteins (Dreger 2003). The proteome of subcellular organelles within a cell is significantly less complex than the proteome of whole cells. Thus, this has allowed development different methods for Isolation and enrichment of subcellular components of a cell, which is called "subcellular fractionation" (Yates Iii, et al. 2005). Subcellular fractionation method has been widely used to reduce the complexity of proteomic samples, thereby leading to identify and quantify more proteins or interest. Subcellular fractionation has also enabled detection of peptides that belong to low-abundant proteins, thereby increasing the coverage rate of protein quantification using MS-based proteomics (Lee, Tan and Chung 2010). Subcellular fractionation has also enabled identification and quantification of plasma membrane proteins which are the key target of antibody-based therapies (Leth-Larsen, Lund and Ditzel 2010). There are several methods have been developed for subcellular fractionation, including differential centrifugation, density-gradient centrifugation, differential detergent fractionation, free-flow electrophoresis, immunoaffinity purification and fluorescent-assisted organelle sorting (Satori, Kostal and Arriaga 2012). Differential centrifugation is one of the conventional methods for subcellular fractionation and involves subsequent centrifugation of the cell or tissue homogenate to isolate organelles such as nuclei, mitochondria and lysosomes. The isolation or separation of organelles in this method is based on variations in size and density of organelles. However, the high organelles contamination with poor separation purity is the major disadvantage of differential centrifugation (Lee, et al. 2010). Density-gradient centrifugation is the most common conventional method for subcellular isolation, which separate organelles using various density-gradient media with different osmolarities, viscosities and densities. There are many types of media used in this method including Ficoll, Percoll, Nycodenz, Metrizamide and Sucrose. Sucrose is the most common medium used in this method as it is inexpensive and biologically inert (Araùjo, Hube and Stasyk 2008). The separation of organelles in density-gradient centrifugation is based on the differences between organelle and medium density which is affected by the organelle content and lipid to protein ratio (Graham 2001). However, the disadvantages of density-gradient centrifugation are time-consuming centrifugation steps and poor isolation purity. To overcome the disadvantages of deferential and density-gradient centrifugation methods, differential detergent

fractionation (DDF) method has been developed to isolate proteins in their native state from four subcellular localisations (cytosolic, membrane, membrane-associated and nuclear and cytoskeleton proteins) using different detergent-containing buffers (Abdolzade‐Bavil, et al. 2004, McCarthy, et al. 2005). Free-flow electrophoresis (FFE) is an alternative method which has been developed to segregate organelles such as peroxisomal membranes, secretory vesicles, plasma membrane vesicles and mitochondria (Satori, et al. 2012). By FFE, such organelles are separated according to their net isoelectric charges. FFE can also isolate proteins in their native state which is ideal for functional proteomic analysis (Weber, et al. 2004, McDonald, et al. 2006). However, FFE can result in co-migration of organelles that have similar isoelectric charges, thereby leading to isolation contamination and low isolation purity (Lee, et al. 2010). Immunoaffinity purification and immunoprecipitation are robust methods used to isolate organelles based on the interaction between antibodies (ligands) that are immobilised on solid plate and organelles of interest (targets). This affinity-based method yields high isolation purity as the purification process can be repeated for optimal organelle enrichment. However, immunoaffinity purification method is expensive and requires more time and effort for optimal purification (Fang and Zhang 2008). Fluorescent-assistance organelle sorting (FAOS) is a cell sorting-based method which has been used to sort and isolate subcellular organelles such as mitochondria, phagosomes, secretory granules and endocytic vesicles (Satori, et al. 2012). FAOS requires a well-known protein biomarker that is specific for the organelle of interest. This specific biomarker must be tagged with green fluorescent protein (GFP) and then transfected into cells, enabling cell sorting based on its fluorescence (Cao, et al. 2008).

3.1.3. Aims of this study

The aim of this chapter is to optimise methods of sample preparation to generate global and subcellular proteomic profile of biologically different subsets murine of $CD4^+$ T cells for identifying distinct proteins which can distinguish natural and induced Tregs from other subsets of $CD4^+$ T cells either individually or in combination.

The objectives of this chapter are:

1- Determination of minimum cell number required for the generation of quantitative proteomics data using murine T cells.
2- Develop a suitable cell lysis condition to yield maximum amount of proteins from the optimised cell numbers
3- Optimisation of tryptic digestion conditions for the proteins prior to mass spectrometry analysis.
4- Determine cell sorting parameters for different subsets of $CD4^+$ T cells populations including iTregs, naïve nTregs and naïve CD4 T cells.
5- Optimisation of subcellular fractionation of $CD4^+$ T cells subpopulation prior to mass spectrometry-based analysis.

3.2. Material and Methods

3.2.1. Optimisation protocols for global proteomic profiling of mouse splenocytes using LC-MS-MS proteomics

3.2.1.1. Determination of optimal cell number for global MS profiling

Mouse splenocytes were harvested from the spleen and counted using NucleoCounter machine (NucleoCounter® NC-250™, Chemometec). Then, different numbers of cells were used as a part of optimisation to determine the best number of cells that yields more proteins, including $10x10^6$, $5x10^6$, $2.5x10^6$, $1x10^6$ and $0.5x10^6$. Cells were centrifuged at 350 x g for 8 min, washed twice using DPBS. Then, cell pellets were resuspended and lysed by adding Ericka's lysis buffer (EB) which contains 9.5M urea (Sigma Aldrich), 2% of DTT (0.5M dithiothreitol) and 1% N-Octyl-Beta-Glucopyranoside. 100 µL of EB were added into samples containing $10x10^6$ cells, whereas 50 µL of EB were added into the other samples. MS-SAFE Protease and Phosphatase Inhibitor (PI, Sigma Aldrich) was added into cell lysates at ratio 1 to 100 dilution (1 µL of PI / 100 µL of EB). Cells were mixed thoroughly by vortexing. Cell lysates were sonicated for 5 minutes and placed on ice for other 5 minutes to ensure complete lysis of cells. Sonication and ice incubation cycles were repeated until cells were completely lysed. Bradford protein assay (Dye Reagent Concentrate, Bio-Rad, Cat No. 500-0006) was conducted to estimate the protein concentration in each cell lysate. 25 ug of proteins were taken from each cell lysate for mass spectrometry sample preparation. Prior to trypsinisation, proteins were reduced by adding 1 µL of 0.5 M dithiothreitol (DTT, Sigma Aldrich) into protein solutions and incubated at 56^0C for 20 min for reduction. Alkylation of proteins was followed by adding 2.7 µL of 0.55 M iodoacetamide (IAA, Sigma Aldrich) and incubated at room temperature in the dark for 15 min. After the reduction and alkylation, the volume of proteins solutions was increased up to 100 µL by adding 50 mM of triethyl ammonium bicarbonate (TEAB) solution to reduce the concentration of urea in the protein solution below 1 M before trypsinisation. Protein digestion was carried out by adding trypsin (Trypsin Gold, Promega, UK) at a ratio 1:20 (1 µg trypsin : 20 µg protein) and incubated at 37^0C with constant mixing for 16 h on a Bioer thermomixer. After the trypsinisation, the solutions of digested proteins were dried at 45^0C using vacuum concentrator (Eppendorf). Dried pellets

were resuspended in 20 µL of 5% acetonitrile in 0.1% formic acid (LCMS grade, Sigma Aldrich, UK) and mixed well by pipetting several times. The resuspended samples were incubated at 37^0C with constant mixing for 5 min and then centrifuged at 12000 x g for 3 min. The supernatants of samples were carefully transferred into 300 µL high recovery liquid chromatography (LC) vials (Chromatography Direct, UK) to be analysed using a 6600 TripleTOF mass analyser (SCIEX). Table 1A (shown in the appendix) summarises the amount and concentration of reagents that were used to prepare samples for proteomic MS analysis.

3.2.1.2. Effect of RBC lysis on total protein yield and identification

The results of experiments detailed in section 3.2.1.1 showed that the number of proteins identified by MS analysis was very low in all samples irrespective of the cell number used for that study. Accordingly, it was hypothesised that the reduction of protein identification might due to the presence of red blood cells (RBCs) in murine splenocytes. Herein, mouse splenocytes were harvested from the spleen. Then RBCs of splenocytes were lysed and eliminated using RBC lysis buffer (BD Pharm Lyse™, BD Biosciences, Cat no. 555899) according to the manufacturer's instructions. Subsequently, splenocytes were counted and prepared as per the protocol explained in section 3.2.1.1. Table 2A (shown in the appendix) summarises the concentration of proteins yielded in each sample, and the amount and concentration of reagents that were used to prepare samples for MS analysis.

3.2.1.3. Optimisation of the volume of cell lysis buffer for improved protein recovery

The results of experiments discussed in section 3.2.1.2 confirmed that the use of RBC lysis buffer was beneficial as it increased the number of proteins identified by MS analysis. However, the concentration of proteins yielded from cell lysis was not improved. To address this, it was suggested to reduce the volume of cell lysis buffer (EB) to increase or improve the concentration of proteins in each sample (cell lysate). Herein, splenocytes were harvested from the spleen. RBCs were lysed by RBC lysis buffer. Then splenocytes were counted and aliquoted into 5 different samples based on cell number as shown in table 3A (shown in the appendix). Then 60 µL, 40 µL, 30 µL, 24 µL and 18 µL of EB were added to lyse samples containing $10x10^6$, $5x10^6$, $2.5x10^6$, $1x10^6$ and $0.5x10^6$ cells, respectively. Then samples were

prepared following the same steps mentioned in section 3.2.1.1. Table 3A (shown in the appendix) summarises the concentration of proteins resulted in each sample and the amount and concentration of reagents used for samples preparation.

3.2.1.4. Effect of addition MS-compatible surfactant agent (protease max) on total protein yield and identification

To assess the effect of a surfactant (Protease Max) to enhance protein solubilisation and enzymatic performance of trypsin on the identification of proteins, optimised experiments carried out in section 3.2.1.3 were repeated with the addition of 1.25 µL of 1% protease max solution immediately after adding trypsin. The samples were further processed following the same protocol mentioned in section 3.2.1.1. Table 4A (shown in the appendix) illustrates the concentration of proteins resulted in each sample and the amount and concentration of reagents used for samples preparation.

3.2.1.5. Optimisation of enzymatic digestion of total proteins

To improve the identification of proteins by MS analysis, the enzymatic performance of trypsin and trypsin Lys-C was compared in the presence and absence of protease max (P.Max). Mouse splenocytes were isolated from the spleen. RBCs were lysed using RBCs lysis buffer, splenocytes were counted and aliquoted into $2.5x10^6$ cells aliquotes. Samples were processed for MS analysis following the protocol mentioned in the previous sections. However, herein, samples were diluted using 50mM TEAB to reduce the final concentration of urea below 1 M prior to protein digestion. Also, two different proteases, trypsin (Sigma Aldrich) and trypsin-Lys-C (a mixture of trypsin and Lys-C, Promega, UK) were used to enzymatically digest proteins in samples with and without the addition of surfactant (P.Max). In each experiment, 4 samples containing the same quantity (25 µg) of proteins were prepared. One set of sample was digested with trypsin only, another sample was digested with trypsin in addition to P.Max, third set was digested with trypsin-Lys-C, and the last was digested with trypsin-Lys-C in addition to P.Max. The total number of proteins identified by MS analysis was evaluated and compared between the 4 samples to determine the best enzymes cocktail for protein digestion by using the number of proteins identified as the readout. Table 5A (shown in the appendix) outlines the concentration of proteins yielded in each sample in all experiments, and the amount and concentration of reagents used for preparation samples.

3.2.1.6. Effect of urea concentration on tryptic digestion of total proteins

To evaluate the impact of urea concentration on tryptic digestion of proteins, 6 different lysis buffers were prepared and used for cell lysis. In this experiment, Ericka's buffer (EB) and Urea solution buffer were used with titration in the molarity of urea in both buffers. The standard Ericka's buffer which was used in the experiments mentioned above contains 9.5 M urea. However, in this section, EB with 8 M urea and 7 M urea were prepared individually. In addition, urea 9.5 M, 8 M and 7 M buffers were also prepared separately and used for cell lysis. Herein, samples containing $2.5x10^6$ cells were lysed with the prepared buffers. Then they were toped up with 50 mM TEAB to reach 200 µL of volume as mentioned in section 3.2.1.5. Then samples were further processed for reduction, alkylation and trypsinisation following the same protocol mentioned in section 3.2.1.1. Herein, trypsin was only used for protein digestion. Table 6A (shown in the appendix) shows a summary of the concentration of proteins of each sample, and the amounts and concentrations of each reagent used for processing samples.

3.2.1.7. Effect of peptide clean-up prior to MS analysis on total protein identification.

To assess whether clean-up of peptides prior to MS analysis can efficiently purify tryptic peptides with higher recovery rates and improve the identification of proteins. Herein, samples were processed as mentioned in section 3.2.1.6. However, in this experiment, tryptic peptides of each sample were purified immediately after trypsinisation using C_{18} Hypersep columns (Thermo Scientific, UK) which were used to separate tryptic peptides from other interfering contaminants present in the digestion solution. Reagents required for the C_{18} Hypersep columns are as follows; releasing solution consists of 95% acetonitrile (ACN) and 0.1% formic acid (FA), the binding solution contains 0.1% FA. Labelled C_{18} columns were placed into 1.5 mL centrifuge tubes. C_{18} columns were then activated by washing with 50 µL of releasing solution and centrifuged at 4.5 x g for 45 sec at room temperature. This step was repeated three times. Flow-through solution was discarded, and the columns were equilibrated with 50 µL of binding solution and centrifuged again at the same settings. This step was also repeated three times. C_{18} columns were placed into new 1.5 mL centrifuge tubes to avoid contamination of flow-through solution. Then, samples were loaded into the C_{18} columns. Samples were loaded into the

columns in multiple batches; 50 µL of sample each time and centrifuged at 4.5 x g for 45 sec. This step was repeated until total volume of sample was completely loaded. After loading samples, columns were washed again with 50 µL of binding solution and centrifuged three times as mentioned above. C_{18} columns were then placed into fresh Lo-Bind Eppendorf tubes (Eppendorf, UK, Cat No 0030108116). Then, samples (peptides fragments) were eluted or released from C_{18} columns by washing with 50 µL of releasing buffer and centrifuged at 4.5 x g for 45 seconds. This step was repeated two times to ensure full recovery of samples. After elution, samples were dried in Eppendorf vacuum concentrator on V-AQ (vacuum-aqueous) setting at 60°C. Dried samples were then resuspended in 15 µL of 5% ACN in 0.1% FA solution and incubated at 37°C with constant mixing for 5 min. Then, samples were centrifuged at 12000 x g for 3 min. Supernatants of samples were transferred into LC (liquid chromatography) tubes for running on MS. Table 7A (shown in the appendix) shows a summary of protein concentrations of each sample and the concentrations and amounts of reagents used for preparation samples.

3.2.1.8. Global shotgun proteomic profiling of purified murine T cells

Mouse T cells were purified from splenocytes using Dynabeads™ Untouched™ Mouse T Cells Kit (Invitrogen, Thermo Fisher Scientific, catalog No. 11413 D). T cells were counted and aliquoted into several samples, each contains $2.5x10^6$ cells. Samples were washed twice with DPBS and centrifuged at 350 x g for 8 min. Then, samples were lysed with Ericka's buffer containing 8 M urea. Protease inhibitor (MS Safe Protease and Phosphate Inhibitor, Sigma Aldrich) was added into samples at a ratio of 1:100 dilution. Samples were sonicated and incubated on ice three times (5 min each) to ensure full lysis. The concentration of total proteins in sample was measured using Bio-Rad Bradford assay (Dye Reagent Concentrate, Bio-Rad). After that, 25 µg of proteins from each sample were transferred into fresh 1.5 mL centrifuge tubes. According to the concentration of proteins, samples were topped up with 50mM TEAB to reach 200 µL total volume. After that, samples were reduced with 1 µL of DTT and incubated at 56°C for 20 min in a shaking thermomixer. The samples were then alkylated with 2.7 µL of IAA and incubated at room temperature in the dark for 15 min. The total volume of each sample before trypsinisation was 203.7 µL and the concentration of urea was maintained between 0.49 M and 0.56 M prior to digestion to enhance the performance of trypsin. The samples were digested

with trypsin at a ratio 1:20 (1 µg trypsin: 20 µg protein) and incubated at 37°C with constant mixing for 16 h in a thermomixer. After trypsinisation, peptides of samples were cleaned up from interfering contaminants and purified using C_{18} Hypersep columns and prepared for MS analysis as mentioned in section 3.2.1.7. Table 8A (shown in the appendix) shows a summary of protein concentrations of each sample and the concentrations and amounts of reagents used for preparation samples.

3.2.2. Optimisation protocols for subcellular proteomic profiling of mouse $CD4^+$ T cells using LC-MS-MS proteomics

3.2.2.1. Optimisation a strategy for sorting a pure population of induced Tregs (iTregs) and natural Tregs (nTregs) prior to MS profiling.

Induced Tregs (iTregs) were generated *in vitro* as mentioned in chapter 2 (section 2.2.4). Naïve $CD4^+CD25^-$ T cells were sorted and rested overnight in complete T cell medium at 37^0C, 5% CO_2. A 24-well culture plate was prepared by adding 900 µL of complete T cell medium in each well with a hanging cell culture insert After that, $2.5x10^5$ of sorted naïve $CD4^+CD25^-$ T cells resuspended in 200 µL of complete T cell medium were transferred into each of hanging cell culture inserts. Then, cells were activated using Dynabeads™ Mouse T-Activator CD3/CD28 for T-Cell Expansion and Activation Kit according to the manufacturer's instruction with IL-2 (30 U/mL) and TGF-β1 (5 ng/mL). The cultures were further incubated for 5 days at 37^0C, 5% CO_2. At day 5, cells were harvested, washed with DPBS twice and counted. Prior to sorting, $0.5x10^6$ of cells were stained with anti-CD4, anti-CD25 and anti-Foxp3 antibodies to assess the induction of Foxp3 of iTregs and to gate $CD4^+CD25^{++}Foxp3^+$ iTregs based on the intensity of CD25 staining. Herein, the strategy of sorting iTregs was based on the intensity of CD25 staining, not on the intensity of Foxp3 staining as the latter requires cell fixation and permeabilisation which damages the surface membrane of cells. Cells were stained with anti-CD4 and anti-CD25 antibodies for sorting $CD4^+CD25^{++}Foxp3^+$ iTregs from activated $CD4^+CD25^+Foxp3^-$ using MoFlo XDP cell sorter, sorting mode: purify. Cell sorter was set to sort $CD4^+CD25^{++}Foxp3^+$ iTregs based on the same gate that was used for identifying $Foxp3^+$ iTregs on flow cytometry. After sorting, sorted $CD4^+CD25^{++}Foxp3^+$ iTregs were rested for three hours in complete T cell medium at 37^0C, 5% CO_2. Cells were then centrifuged at 350

x g for 8 min and counted. $0.2x10^6$ of iTregs were stained again with anti-CD4, CD25 and Foxp3 antibodies to assess the purity of iTregs sorting.

The sorting of nTregs was mainly based on the intensity of CD25 staining. To achieve this, naïve $CD4^+$ T cells were purified from splenocytes as mentioned in chapter 2 (section 2.2.2). purified $CD4^+$ T cells were stained with anti-CD4 and CD25 antibodies and sorted into two separate populations $CD4^+CD25^-$ T cells and $CD4^+CD25^+$ nTregs.

3.2.2.2. Determination of the number of spleens required to obtain $2.5x10^6$ of purely sorted $CD4^+CD25^{++}Foxp3^+$ iTregs and $CD4^+CD25^+Foxp3^+$ natural Tregs (nTregs) prior to subcellular fractionation.

According to the optimised protocol used for global profiling of T cells, $2.5x10^6$ was the optimal number of cells for identification higher number of proteins. Therefore, to obtain $2.5x10^6$ of purely sorted iTregs and nTregs, splenocytes were harvested from two and three spleens in two separate experiments. After harvesting, splenocytes were pooled together and processed for isolation $CD4^+$ T cells which were then processed for sorting naïve $CD4^+CD25^+$ nTregs and $CD4^+CD25^-$ T cells as mentioned above. Then $CD4^+CD25^-$ T cells were then cultured for generating iTregs *in vitro* as mentioned above.

3.2.2.3. Optimisation subcellular fractionation of $CD4^+$ T cells using Mem-PER Plus membrane protein extraction kit (pilot study).

Subcellular fractionation of $CD4^+$ T cells was carried out using Mem-PER™ Plus membrane proteins extraction kit (Thermo Fisher Scientific, Cat No. 89842) which has been designed to isolate and enrich integral membrane proteins and membrane-associated proteins from cells. Following manufacturer's instructions, $2x10^6$ of $CD4^+$ T cells were harvested and centrifuged at 350 x g for 8 min. The cell pellet was washed with 2 mL of cell wash solution which was included in the kit and centrifuged at 350 x g for 6 min. After washing, supernatants were carefully discarded, and the cell pellet was washed again with 2 mL of cell wash solution and centrifuged at 300 x g for 6 min. Following the final wash, 600 µL of permeabilisation buffer, which was included in the kit, was added to the cell pellet. Cells were resuspended by brief and gentle vortexing. After that, 6 µL of protease inhibitor (MS Safe, Sigma Aldrich) was added into the cell suspension at ratio 1 in 100 dilution.

The homogenous cell suspension was incubated for 20 min at 4^0C with constant mixing using a rotary mixer. Permeabilised cells were centrifuged at 16000 x g for 15 min at 4^0C. The supernatants containing cytosolic or cytoplasmic proteins were carefully removed and transferred into a new 2 mL LoBind Eppendorf tube. The pellet containing membrane proteins fraction was resuspended with 300 µL of solubilisation buffer which was included in the kit to solubilise membrane proteins. The solution was mixed well by careful pipetting up and down for 2 min. Then, 3 µL of protease inhibitor was added into the protein solution at ratio 1 in 100 dilution. The protein solution was incubated at 4^0C for 60 min with constant mixing. The protein solution containing solubilised membrane proteins was centrifuged at 16000 x g for 15 min at 4^0C. The supernatants containing membrane and membrane-associated proteins were transferred to a new 2 mL LoBind Eppendrof tube for further downstream analysis.

3.2.2.4. Protein quantitation Assay.

To quantify proteins in the cytoplasmic and membrane fractions enriched from cells, Pierce™ 660 nm protein assay was carried out according to the manufacturer's instructions.

To purify and concentrate proteins of the cytoplasmic and membrane fractions isolated from the cells, proteins were precipitated using cold acetone. Acetone precipitation of proteins was also used to remove all interfering substances and detergents included in the permeabilisation and solubilisation buffers, which are incompatible with MS analysis. To achieve this, acetone was cooled to -20^0C for a day. Then five times the sample volume of cold acetone (-20^0C) was added into the cytoplasmic and membrane fractions. For cytoplasmic proteins fraction, the samples (600 µL volume) were transferred to a new 5 mL LoBind Eppendorf tubes, and then 3mL of cold acetone was added into the samples. For membrane proteins fraction, the samples (300 µL volume) were transferred to a new 2 mL LoBind Eppendorf tubes, and then 1.5 mL of cold acetone was added into the samples. Then both membrane and cytoplasmic proteins samples were incubated overnight at -20^0C. Samples were then centrifuged at 16000 x g for 30 min at 4^0C. Supernatants were carefully discarded, and the protein pellet was precipitated again by adding 1.5 mL of cold acetone and 300 µL of nano-pure distilled water to ensure removing

all interfering detergents in the protein solution. Proteins samples were then incubated for 60 min at -20^{0}C. proteins samples were centrifuged at 16000 x g for 15 min at 4^{0}C. the protein pellet was then prepared for downstream process.

3.2.2.5. Preparation of precipitated cytoplasmic and membrane protein samples for MS analysis.

8 M urea dissolved in nanopore water was prepared. 1%, 0.1% and 0.2% protease max surfactant in 50 mM TEAB solutions were prepared. For membrane protein samples, the precipitated protein pellet was solubilised by adding 20 µL of 0.2% protease max in 50 mM TEAB and 15 µL of 8M urea. Samples were then incubated for 60 minutes at 4^{0}C with constant mixing using rotary mixer. After the vortexing, 60 µL of 50mM TEAB were added into samples. The samples were reduced by adding 1 µL of 0.5 M DTT into samples, and samples were incubated at 56^{0}C for 20 min. Alkylation steps were performed by adding 2.7 µL of 0.55 M IAA into samples which were incubated at room temperature in the dark for 15 min. Then proteins were digested by adding trypsin at a ratio 1:20 (1 µg trypsin: 20 µg protein) and 1 µL of 1% protease max surfactant. Samples were incubated overnight at 37^{0}C with constant shaking at medium speed (200 rpm) on a thermomixer.

For cytoplasmic protein samples, the precipitated protein pellet was solubilised by adding 20 µL of 0.1% protease max in 50mM TEAB. Samples were then incubated for 60 min at 4^{0}C with constant mixing by vortexing at 70 rpm. Then 75 µL of 50 mM TEAB were added into samples. Samples were then prepared similar to membrane proteins samples as mentioned above. Table 3.1 summarise the concentrations and volumes of reagents that were used for preparation membrane and cytoplasmic protein samples for MS analysis.

After trypsinisation, samples were dried in an Eppendorf vacuum concentrator on V-AQ setting at 60°C. Dried samples were then resuspended in 15 µL of 5% ACN in 0.1% FA solution and incubated at 37°C with constant mixing for 5 min. Then, samples were centrifuged at 12000 x g for 3 min. Supernatants of samples were transferred into LC (liquid chromatography) tubes for running on MS.

Table 3.1: Summary of in-solution solubilisation and digestion reaction volumes for membrane and cytoplasmic proteins.

Reagents	**Membrane proteins**	**Cytoplasmic proteins**
0.1% protease max in 50 mM TEAB	-	20 µL
0.2% protease max in 50 mM TEAB	20 µL	-
8 M Urea	15 µL	-
50 mM TEAB	60 µL	75 µL
0.5 M DTT	1 µL	1 µL
0.55 M IAA	2.7 µL	2.7 µL
Trypsin (1 µg/µL)	1.5 µL	1.5 µL
1% protease max in 50 mM TEAB	1 µL	1 µL
Total	101.2 µL	101.2 µL

3.3. Results

3.3.1. Optimisation protocols for global proteomic profiling of mouse splenocytes and T cells using LC-MS-MS proteomics

3.3.1.1. Determination of optimal cell number for global MS profiling

To determine the minimum cell number required to obtain maximum amounts of total proteins in calorimetric protein assay and mass spectrometry-based identification, five different cell numbers were tested. Highest amounts of total proteins were obtained from samples containing $10x10^6$ cells (mean = 195.3 µg, SEM 24.16), compared with samples containing $2.5x10^6$ cells (mean = 42.1 µg, SEM 7.15), $1x10^6$ cells (mean = 16.8 µg, SEM 1.05) and $0.5x10^6$ cells (mean = 7.5 µg, SEM 2.81). There was no significant difference in the amounts of total proteins between samples containing $10x10^6$ and $5x10^6$ cells (mean = 77.8 µg, SEM 3.56) (figure 3.4 A) as measured by the protein assay. Samples with $1x10^6$ and $0.5x10^6$ cells were not processed for MS analysis as they yielded less than 25 µg of total proteins. However, the number of proteins identified by MS analysis was not relatively correlated with the number of cells within samples. The number of proteins identified in samples with $10x10^6$ cells was between 120 and 806 proteins (mean = 415 proteins, SEM 158.9), whereas the mean of the number of proteins was 256 (SEM 38.18) and 226 (SEM 35.44) identified in samples with $5x10^6$ and $2.5x10^6$ cells, respectively (figure 3.4 B). There was no significant difference in the number of proteins identified by MS between samples with $10x10^6$, $5x10^6$ and $2.5x10^6$ cells. These results also were seen in the number of distinct peptides and spectra as there was no significant difference in the number of distinct peptides and spectra between samples with $10x10^6$ (mean = 2615 (SEM 878.3) and 5364 (SEM 1631)), $5x10^6$ (mean = 1667 (SEM 219.3) and 3602 SEM 400.1)) and $2.5x10^6$ (mean = 1534 (SEM 178.2) and 3469 (SEM 350.7); peptides and spectra, respectively) (figure 3.4 C and D). However, compared to samples of $10x10^6$ and $5x10^6$ cells, samples containing $2.5x10^6$ cells provided more reasonable results according to their cell number. These results confirmed that samples need further optimisation for MS analysis as the number of proteins, peptides and spectra identified in all samples was significantly low compared to the number of cells especially samples with higher number of cells.

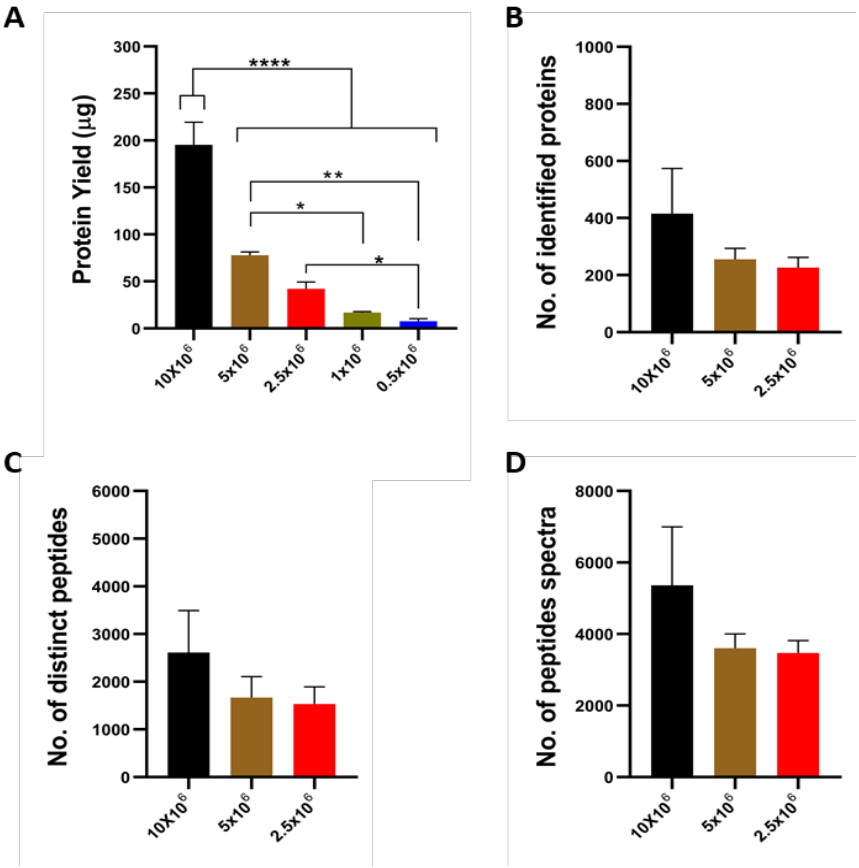


Figure 3.4. Bar graphs showing the amount and total number of proteins obtained from samples containing different cell number**. (A)** The amount of proteins in micrograms yielded from samples containing different number of cells. **(B)** The number of proteins that were identified in samples containing different number of cells using LC-MS-MS analysis. **(C)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(D)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (****; P-value <0.0001, **; p-value <0.0046, *; p-value <0.0138).

3.3.1.2. Effect of RBC lysis on total protein yield and identification

One of the potential problems with using total splenocytes for MS analysis is the presence of RBC in the cell preparation which may interfere with cell lysis, protein assays and other downstream processing. To improve the identification of total proteins by MS, the RBC lysis buffer was used to remove RBCs from splenocytes. The results showed that the amounts of proteins yielded after using RBC lysis buffer were decreased in samples with $10x10^6$ cells (mean = 147.5 µg, SEM 32.5) and $5x10^6$ (mean = 52.50 µg, SEM 2.5), whereas the amounts of proteins were significantly increased in samples with $1x10^6$ cells (mean = 39 µg, SEM 1) and $0.5x10^6$ (mean = 32.5 µg, SEM 0.5) (figure 3.5 A), compared to the results in section 3.3.1.1. The results also showed that there was no significant difference in the amounts of proteins yielded between all samples, especially samples with $5x10^6$, $2.5x10^6$, $1x10^6$ and $0.5x10^6$ cells as they showed approximately similar values. These results suggested that there might be a background noise within samples, which might affect the accuracy of proteins quantitation assay. However, the results showed a significant increase in the number of proteins, distinct peptides and spectra in all samples after using RBC lysis buffer, compared to the results in section 3.3.1.1 where samples were processed without using RBC lysis buffer. According to the results, the mean of the number of proteins identified by MS was higher in samples with $5x10^6$ cells (1336, SEM 19) compared to samples with $10x10^6$ (1133, SEM 43), $2.5x10^6$ (1202, SEM 28.5), $1x10^6$ (473.5, SEM 84.5) and $0.5x10^6$ cells (565, SEM 75) (figure 3.5 B). Also, the number of distinct peptides and spectra detected by MS was proportional to the number of proteins in all samples. The mean of the number of distinct peptides and peptides spectra was higher, of course, in samples with $5x10^6$ cells (7250 (SEM 103) and 11463 (SEM 162.5), respectively) compared to other samples with $10x10^6$ (6183 (SEM 79) and 9541 (SEM 80)), $2.5x10^6$ (6678 (SEM 158.5) and 10567 (SEM 250.5)), $1x10^6$ (1821 (SEM 324.5) and 2786 (SEM 496.5)) and $0.5x10^6$ (2612 (SEM 347) and 3954 (SEM 525.5), respectively) (figure 3.5 C and D). Nevertheless, the number of proteins, peptides and spectra detected by MS was not correlated to the number of cells within samples. For example, the number of proteins identified in samples with $0.5x10^6$ cells was ~50% less than number of proteins identified in samples with $10x10^6$ cells which have 20 times more cells than

samples with 0.5x10^6 cells, although the same amounts of proteins (25 µg) were digested from both samples.

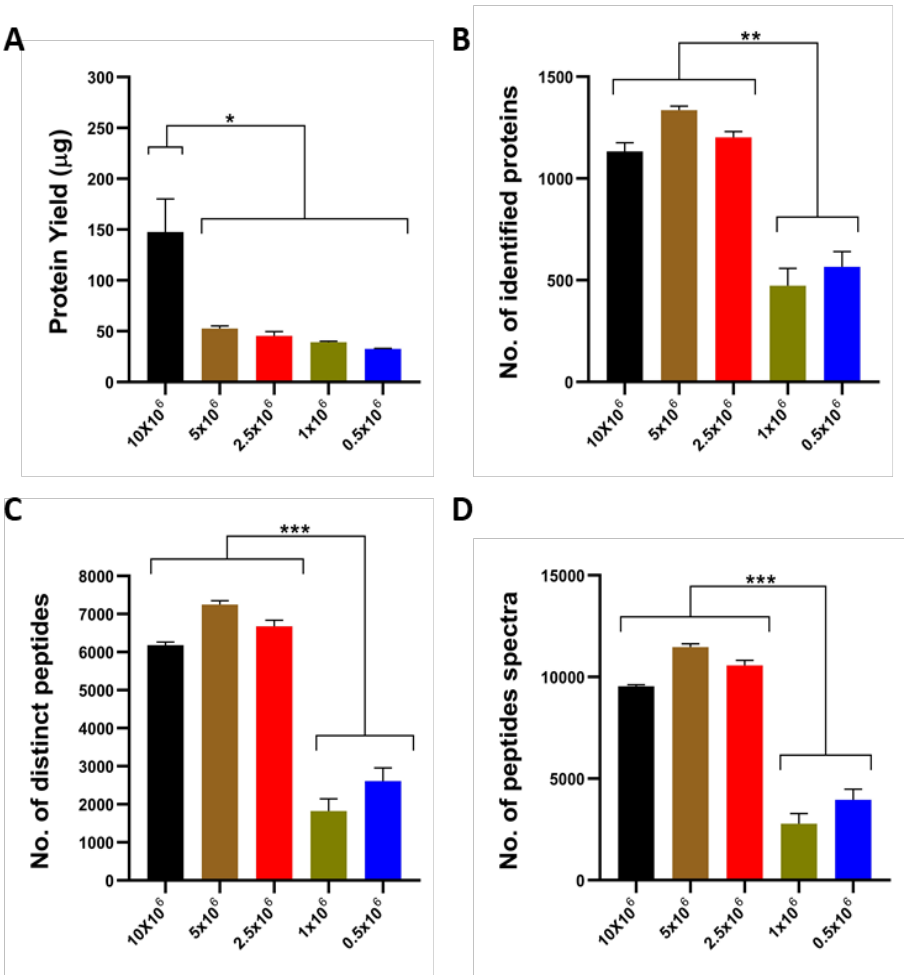


Figure 3.5. Bar graphs showing the amount and total number of proteins obtained from samples containing different cell number after using RBC lysis buffer. **(A)** The amount of proteins in micrograms yielded from samples containing different number of cells. **(B)** The number of proteins that were identified in samples containing different number of cells using LC-MS-MS analysis. **(C)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(D)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.0225, **; p-value <0.0044, ***; p-value <0.0006).

3.3.1.3. Optimisation the volume of cell lysis buffer for improved protein recovery

To improve the yield of total proteins, samples were lysed with different volumes of cell lysis buffer based on the number of cells they contain. The results showed the amounts of proteins yielded in all samples were relatively correlated to the number of cells after titrating the volume of lysis buffer, compared to the results in section 3.3.1.2 where there was no significant difference in the amount of proteins obtained from samples containing $5x10^6$, $2.5x10^6$, $1x10^6$ and $0.5x10^6$ cells. The results revealed that the amount of proteins yielded were higher, of course, in samples containing $10x10^6$ cells (mean = 149 µg, SEM 11), compared to samples with $5x10^6$ (mean = 60 µg, SEM 4), $2.5x10^6$ (mean = 38.3 µg, SEM 2.3), $1x10^6$ (mean = 24.4 µg, SEM 1.6) and $0.5x10^6$ (mean = 14.5 µg, SEM 0.8) (figure 3.6 A).

For proteins identification by MS, the results generally showed an increase in the mean of the number of total proteins identified in all samples after reducing the volume of lysis buffer during preparation samples, except samples containing $5x10^6$ cells which showed a reduction in the mean of the number of total proteins (mean = 1096, SEM 4) compared to the results in section 3.3.1.2. The results also showed that samples with $2.5x10^6$ cells specifically had the highest number of proteins (mean = 1597, SEM 74) identified by MS, compared to samples with $10x10^6$ (mean = 1316, SEM 74), 5x10 (1096, SEM 4), $1x10^6$ (979, SEM 8) and $0.5x10^6$ cells (mean = 803, SEM 50) (figure 3.6 B). The number of distinct peptides and spectra detected by MS analysis was correlated with the number of proteins identified in all samples (figure 3.6 C and D).

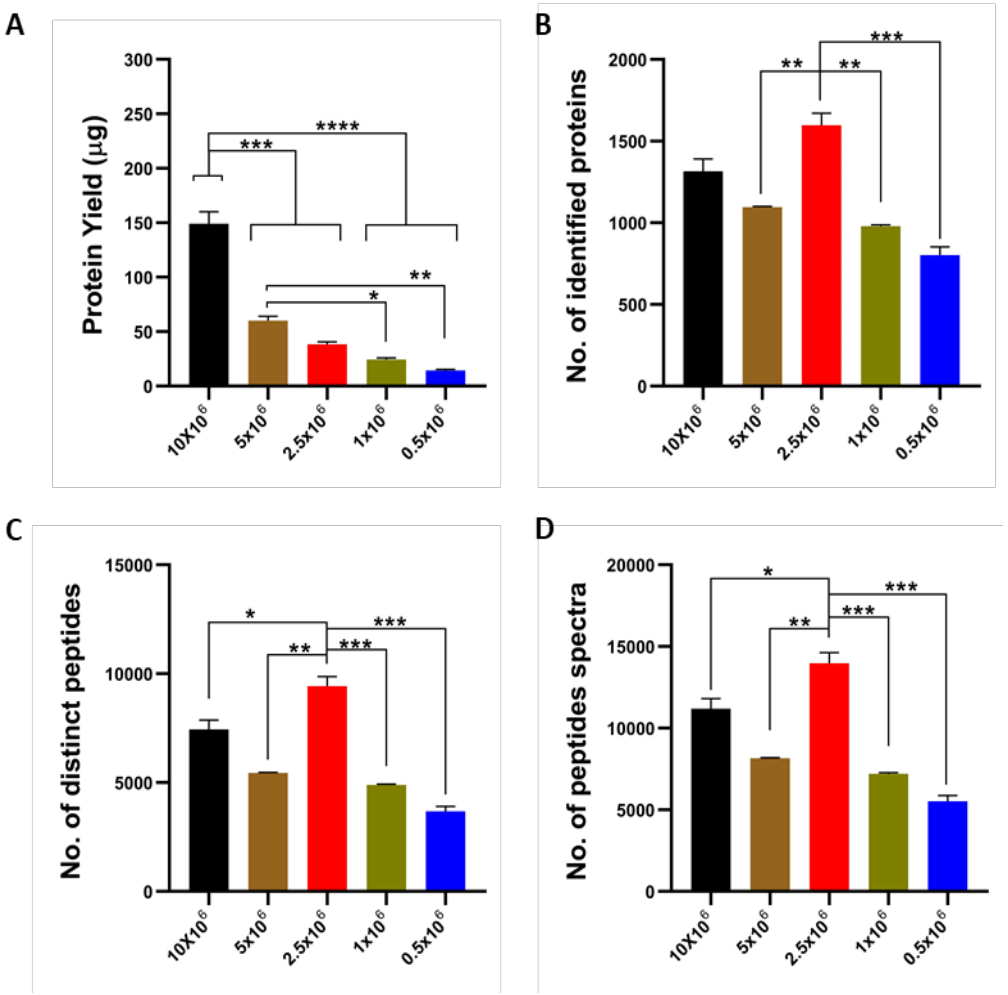


Figure 3.6. Bar graphs showing the amount and total number of proteins obtained from samples with different cell number which were lysed using different volumes of cell lysis buffer. **(A)** The amount of proteins in micrograms yielded from samples containing different number of cells. **(B)** The number of proteins that were identified in samples containing different number of cells using LC-MS-MS analysis. **(C)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(D)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.0275, **, p-value <0.0099, ***; p-value <0.0004, ****; p-value < 0.0001).

3.3.1.4. Effect of addition MS-compatible surfactant agent (protease max) on the protein identification

To achieve higher identification of proteins by MS, protease max (P.Max), a surfactant agent compatible with MS, was added into samples immediately after addition of trypsin to enhance the enzymatic function of trypsin. The results showed a reduction in the number of total proteins identified by MS in samples containing $10x10^6$ and $2.5x10^6$ cells, whereas samples with $5x10^6$, $1x10^6$ and $0.5x10^6$ cells showed a slight increase in the number of total proteins (figure 3.7 A), compared to the results shown in section 3.3.1.3. Despite of the reduction and increase in the number of proteins after using P.Max , samples with $2.5x10^6$ cells provided the highest number of proteins (mean = 1467, SEM 35.5) compared to samples with $10x10^6$ (mean = 1172, SEM 52), $5x10^6$ (mean = 1160, SEM 30.5), $1x10^6$ (mean = 1121, SEM 41) and $0.5x10^6$ (mean = 820, SEM 44) (figure 3.7 A). The number of distinct peptides and spectra was also correlated with the number of proteins (figure 3.7 B and C).

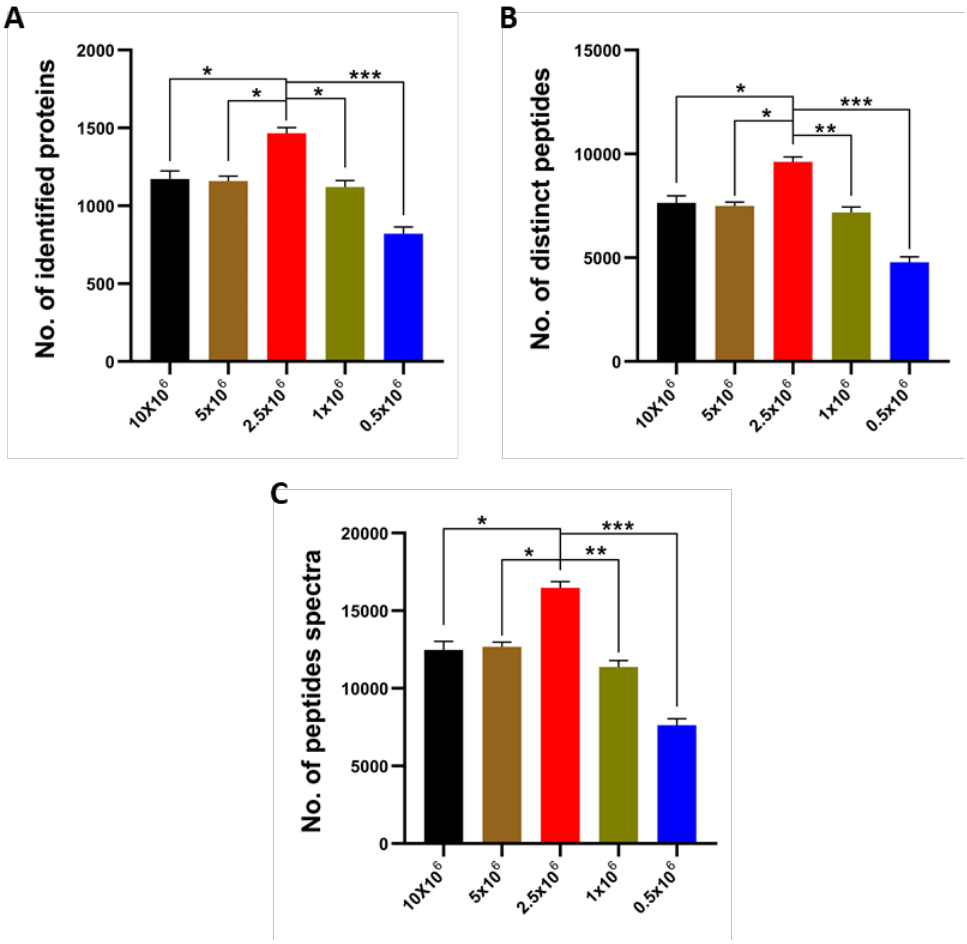


Figure 3.7. Bar graphs showing the total number of proteins, distinct peptides and peptides spectra identified in samples with different cell number which were digested in the presence of surfactant protease max. **(A)** The number of proteins that were identified in samples containing different number of cells using LC-MS-MS analysis. **(B)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(C)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.0202, **, p-value <0.0069, ***; p-value <0.0004).

According to the results shown in sections 3.3.1.1, 3.3.1.2 and 3.3.1.3, samples with $2.5x10^6$ cells showed a significant increase in the number of total proteins identified by MS analysis and were corresponding to the optimisation of experimental protocols as mentioned in sections 3.3.1.1, 3.3.1.2 and 3.3.1.3 (table 3.2). Also, the mean of the amounts of total proteins yielded from samples with $2.5x10^6$ was in range between 34.8 and 45.25 µg (range 10.45, SEM 5.23), which means that there were no significant variations in the amounts of proteins yielded from $2.5x10^6$ cells in different experiments. For these reasons, $2.5x10^6$ cell number was selected as an optimal cell number for optimising a protocol for global proteomic profiling of mouse T cells.

Table 3.2: Summary of results representing the mean of the number of proteins identified by MS in samples containing different number of cells.

Optimisation methods			**The mean of the number of identified proteins by MS analysis**				
Lysis Buffer	**RBC lysis Buffer**	**Digestion**	**$0.5x10^6$**	**$1x10^6$**	**$2.5x10^6$**	**$5x10^6$**	**$10x10^6$**
Ericka's	NO	Trypsin	-	-	226.3	255.5	415
Ericka's	Yes	Trypsin	565	473.5	1202	1336	1133
Ericka's (low volume)	Yes	Trypsin	803	979	1597	1096	1316
Ericka's (low volume)	Yes	Trypsin+P.Max	820	1121	1467	1160	1172

3.3.1.5. Optimisation of enzymatic digestion of total proteins

To achieve higher identification of proteins by MS, two different proteases, trypsin and trypsin-Lys-C, were used for protein digestion with and without addition of P.Max. The results showed that the number of proteins identified by MS analysis was significantly higher in samples in which proteins were digested using trypsin alone (mean = 1484, SEM 58.5) compared to samples in which proteins were digested using trypsin-Lys-C alone (mean = 1249, SEM 20.5) and trypsin-Lys-C with addition of P.Max (mean = 1044, SEM 26.8) (figure 3.8 A). The results also showed that there was no significant difference in the number of proteins identified between samples digested with trypsin and samples digested with trypsin in addition to P.Max (mean

= 1265, SEM 15.9), however samples digested with trypsin showed the highest number of identified proteins (figure 3.8 A). The results also showed a reduction in the number of proteins identified in samples which were digested in the presence of P.Max, compared to samples digested in the absence of P.Max (figure 3.8 A). The number of distinct peptides and spectra was also correlated with the number of proteins (figure 3.8 B and C).

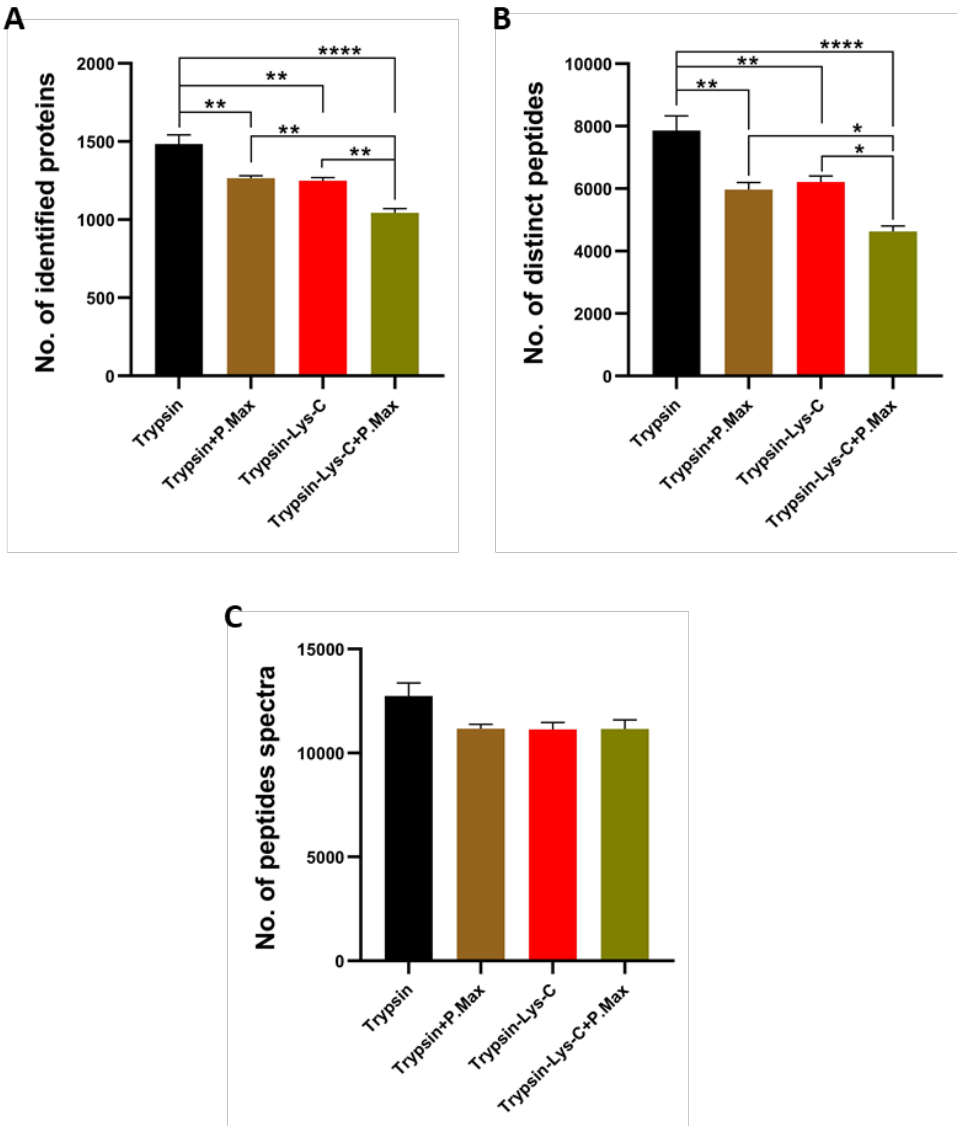


Figure 3.8. Bar graphs showing the total number of proteins, distinct peptides, peptides spectra identified from protein samples that were digested in different conditions. **(A)** The number of proteins that were identified in samples in which proteins were digested using different proteases with and without addition of P.Max. **(B)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(C)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (**, p-value <0.0083, ****; p-value <0.0001).

3.3.1.6. Effect of urea concentration on tryptic digestion of total proteins

To assess the impact of urea concentration on the enzymatic performance of trypsin, cells were lysed using two different lysis buffers; Ericka's buffer (EB) and urea buffer. Both buffers were prepared using different concentrations of urea. The results showed that there was no significant difference in the amounts of proteins yielded in samples lysed with different lysis buffers containing different concentrations of urea including EB 9.5 M, EB 8 M, EB 7 M, urea 9.5 M, urea 8 M and urea 7 M (figure 3.9 A).

The results also showed that the number of proteins identified in samples that were lysed by EB containing 8 M of urea concentration was the highest (mean = 1927 proteins, SEM 28.5) compared to other samples lysed with different lysis buffers (figure 3.9 B). The results showed that the number of identified proteins was significantly higher in samples lysed by EB with different concentrations of urea than samples lysed by urea buffer with different concentrations, especially at concentration 9.5 M and 7 M of urea (figure 3.9 B). The mean of number of proteins identified in samples lysed with EB 9.5 M and EB 7 M was 1650 (SEM 20) and 1429 (SEM 78.5), respectively, whereas it was 1089 (SEM 84) and 972 (85.5) in samples lysed with urea 9.5 M and 7 M, respectively (figure 3.9 B). However, although the difference in the number of identified proteins between samples lysed with EB 8 M (mean = 1927 proteins, SEM 28.5) and urea 8 M (mean = 17347, SEM 67) was not significant, samples lysed by EB 8 M yielded more proteins than samples lysed with urea 8 M. Moreover, according to the results, the concentration 8 M in both buffers was sufficient to yield more soluble proteins from cells than concentrations 9.5 M and 7 M. The number of distinct peptides and spectra was also correlated with the number of proteins (figure 3.9 C and D).

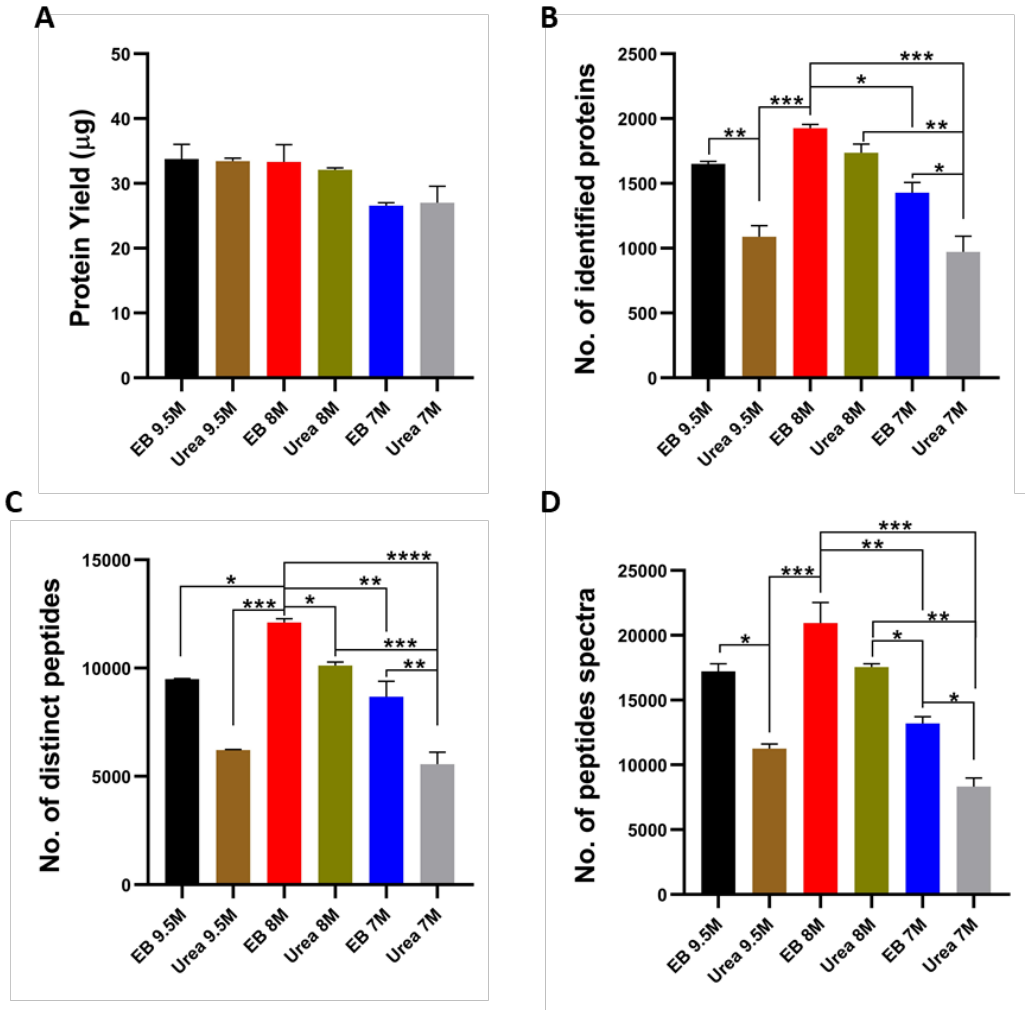


Figure 3.9. Bar graphs showing the amount and total number of proteins obtained from samples lysed with different lysis buffers containing different concentrations of urea (in molarity (M)). **(A)** The amount of proteins in micrograms yielded from samples that were lysed with different lysis buffers containing different concentrations of urea (in molarity (M)). **(B)** The number of proteins that were identified in samples lysed with different lysis buffers containing different concentration of urea using LC-MS-MS analysis **(C)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(D)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.0275, **, p-value <0.0099, ***; p-value <0.0004 ****; p-value < 0.0001).

3.3.1.7. Effect of peptide purification prior to MS analysis on total protein yield and identification

To assess the impact of peptide purification prior to MS analysis on identification of total proteins. In this experiment, tryptic peptides were purified and isolated from contaminants (such as urea) within digestion solution prior to MS analysis using C_{18} Hypersep columns. In general, the results showed that the number of proteins identified by MS analysis was significantly increased after using peptides purification columns compared to the results mentioned in section 3.3.1.6 where samples were analysed on MS without peptide purification (figure 3.10 A). The increase in the number of proteins was in range between 116 and 628 more proteins identified in samples after using C_{18} columns compared to samples prepared without using C_{18} columns (section 3.3.1.6) (figure 3.10 A). The results also showed that samples lysed with EB containing 8 M urea concentration yielded the highest number of proteins (mean = 2427, SEM 133.5) identified after using C_{18} columns compared to other samples that were lysed by different lysis buffers (figure 3.10 A). The results also showed that the number of distinct peptides was significantly increased after peptides purification in samples that were lysed with EB 9.5 M, urea 9.5 M, EB 8 M, urea 8 M and urea 7 M lysis buffer. However, there was no significant increase in the number of peptides before and after purification in samples lysed with urea 7 M lysis buffers (figure 3.10 B). For peptide spectra, the results showed that there was a significant increase in the number of peptides spectra after peptide purification in samples that were lysed with EB 9.5 M, urea 9.5 M and EB 7 M only, compared to other samples lysed with EB 8 M, urea 8 M and urea 7 M (figure 3.10 C). These results confirmed the efficiency of the C_{18} columns in peptide purification which could enhance the identification of proteins.

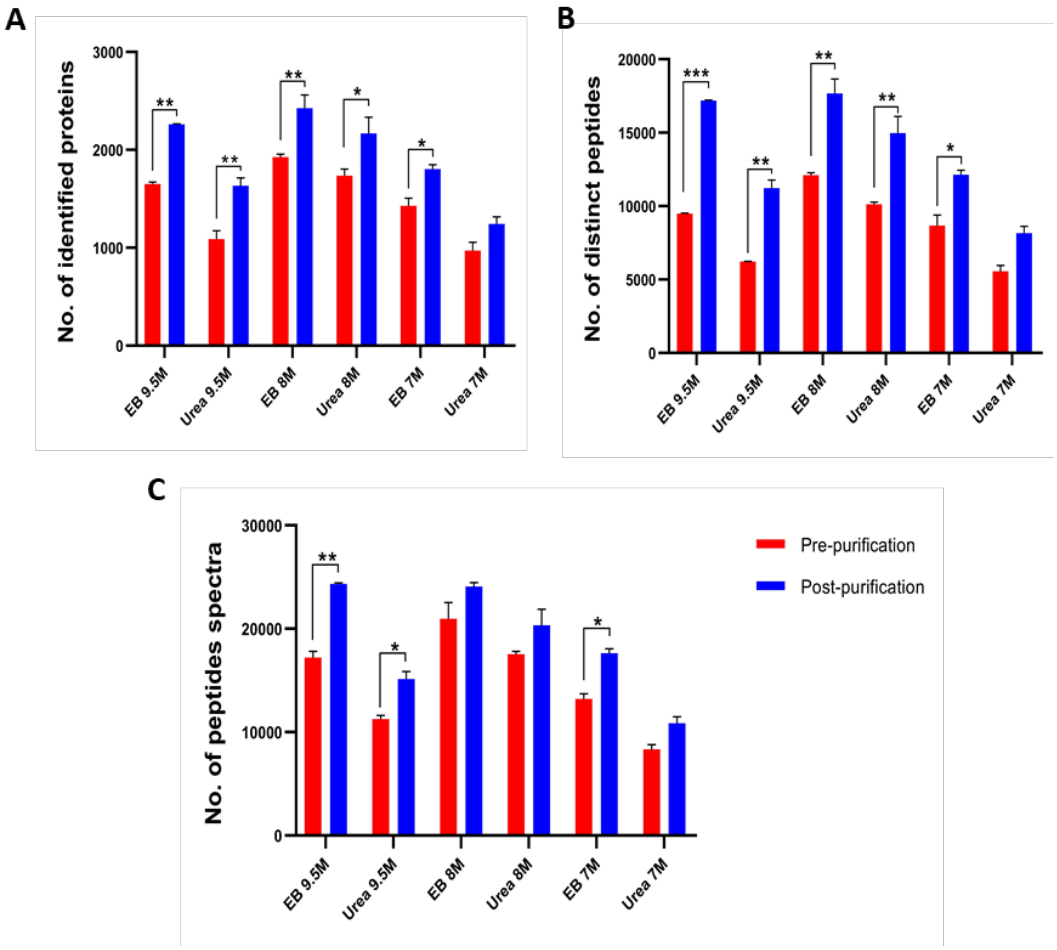


Figure 3.10. Bar graphs showing the total number of proteins, distinct peptides and peptides spectra identified pre- and post-purification of peptides derived from samples lysed with different lysis buffers. **(A)** The number of proteins that were identified before and after purification of peptides using LC-MS-MS analysis. **(B)** The number of distinct peptides that uniquely match or belong to the identified proteins. **(C)** The number of peptides spectra that represent the MS-MS peaks of peptides fragments used to identify distinct peptides of proteins. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.0275, **, p-value <0.0099, ***; p-value <0.0004 ****; p-value < 0.0001).

3.3.1.8. Global shotgun profiling of purified murine T cells

To create a library of proteins derived from mouse T cells using MS shotgun analysis, mouse T cells were isolated from splenocytes and lysed by EB 8 M lysis buffer to extract proteins. Proteins were digested and processed prior to MS analysis according to the optimised protocol mentioned in section 3.2.1.8. The results showed that the number of proteins identified from $2.5x10^6$ of purified T cells was in range between 1053 and 1232 at 1% of false discovery rate (FDR) and 50% of confidence score based on four separate experiments as shown in table 3.3. Proteins showing confidence score ≥ 50% were defined as "confident proteins". Functional information and subcellular location of proteins were retrieved from Uniprot database, an online resource of protein sequence and functional information (www.uniprot.org). Figure 3.11 shows an example of how biological information of identified proteins was acquired from Uniprot database online.

Table 3.3: Summary of the results of the total number of proteins, distinct peptides and peptides spectra that were identified from global proteomic profiling of murine T cells in four separate experiments.

Experiment	Cell Type	Cell No.	Proteins No.	Distinct peptides	Spectra
1	Purified T cells	$2.5x10^6$	1225	6723	12940
2	Purified T cells	$2.5x10^6$	1101	5553	11585
3	Purified T cells	$2.5x10^6$	1232	6795	12907
4	Purified T cells	$2.5x10^6$	1053	5254	11009

For the first experiment, the results showed that the total number of proteins identified at 1% FDR was 1225. 272 (22.2%) of them showed confidence score ≥50%, whereas 953 (77.8%) of proteins had confidence score < 50%. Of 272 confident proteins, 177 proteins (14.45% of total) were cytoplasmic and nuclear, while 95 (7.75%) proteins were identified as membrane proteins. 52 (4.24%) of membrane proteins were characterised as plasma membrane proteins or cell

surface membrane proteins, whereas the remaining 43 (3.51%) were characterised as organelles membrane proteins (Figure 3.12).

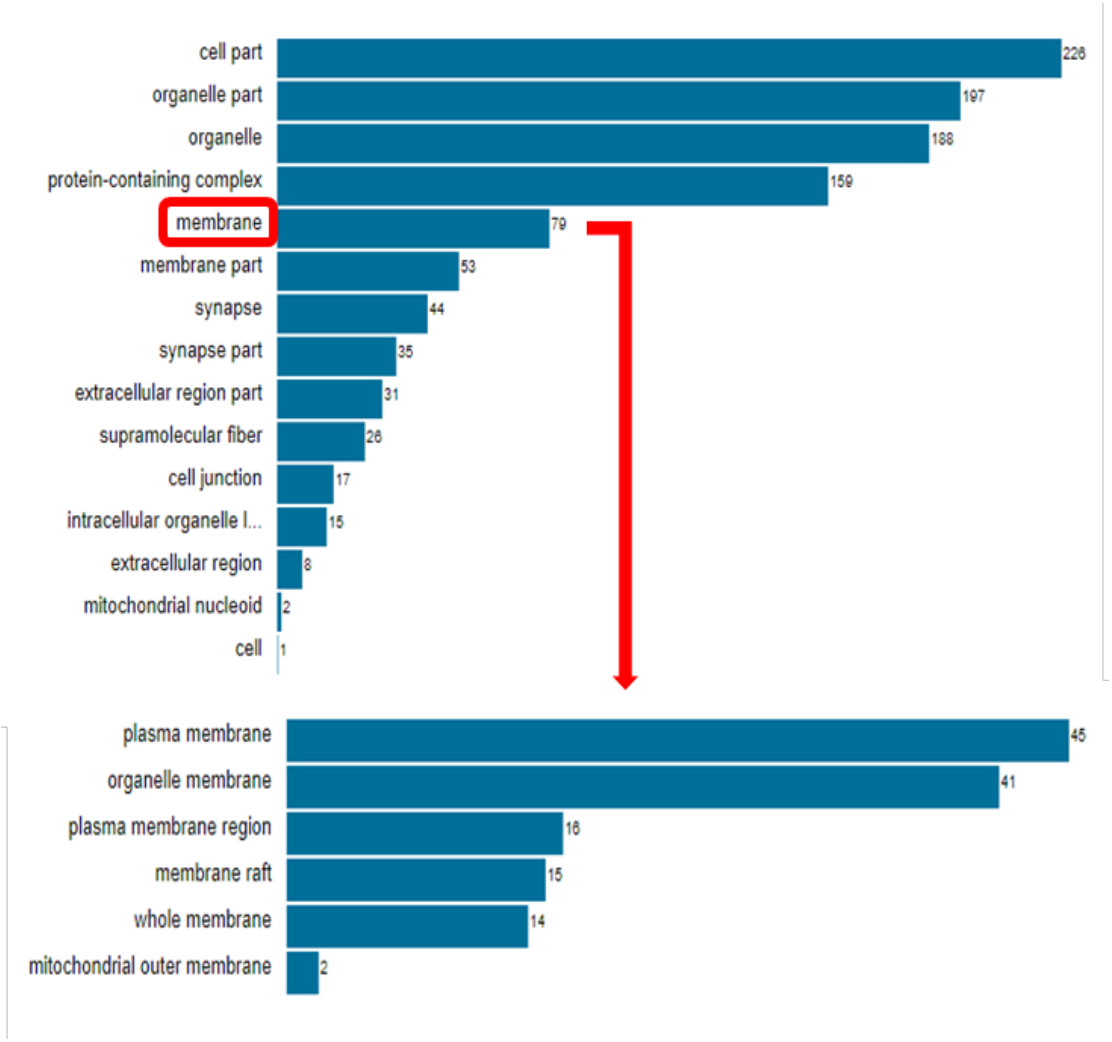


Figure 3.11. An example of Bar graph showing subcellular locations of 226 confident proteins, which was obtained from Uniprot database. Membrane proteins highlighted by Red rectangle can be sorted by Uniprot into different subcellular locations of membrane-bound proteins including plasma membrane proteins and intracellular membrane proteins (organelles membrane and mitochondrial membrane).

In the second experiment, the results showed that the total number of proteins identified at 1% FDR was 1101. Out of which 265 (24.07%) showed confidence score ≥ 50%, whereas 836 (75.93%) of proteins had confidence score < 50%. Of 265 confident proteins, 171 proteins (15.53% of total) were cytoplasmic and nuclear,

while 94 (8.54%) proteins were identified as membrane proteins. 54 (4.91%) of membrane proteins were characterised as plasma membrane proteins or cell surface membrane proteins, whereas the remaining 40 (3.63%) were characterised as organelles membrane proteins (Figure 3.12).

In the third experiment, the results showed that the total number of proteins identified at 1% FDR was 1232. 289 (23.46%) of them showed confidence score ≥ 50%, whereas 943 (76.54%) of proteins had confidence score < 50%. Of 289 confident proteins, 187 proteins (15.18% of total) were cytoplasmic and nuclear, while 102 (8.28%) proteins were identified as membrane proteins. 56 (4.55%) of membrane proteins were characterised as plasma membrane proteins or cell surface membrane proteins, whereas the remaining 46 (3.73%) were characterised as organelles membrane proteins (figure 3.12).

In the last experiment, the results showed that the total number of proteins identified at 1% FDR was 1053. 226 (21.46%) of them showed confidence score ≥ 50%, whereas 827 (78.54%) of proteins had confidence score < 50%. Of 226 confident proteins, 147 proteins (13.96% of total) were cytoplasmic and nuclear, while 79 (7.50%) proteins were identified as membrane proteins. 45 (4.27%) of membrane proteins were characterised as plasma membrane proteins or cell surface membrane proteins, whereas the remaining 34 (3.23%) were characterised as organelles membrane proteins (figure 3.12).

According to the results of four experiments mentioned above, the percentage of plasma membrane proteins was in range between 4.24% and 4.91% in samples that were prepared using the optimised protocol detailed in section 2.3.1.8. Identification of proteins that are exclusively expressed on the plasma membrane was relatively limited compared to other proteins (cytoplasmic and nuclear). Therefore, it was suggested to attempt subcellular fractionation to specifically enriched membrane-bound proteins. The overall results were summarised in figure 3.12.

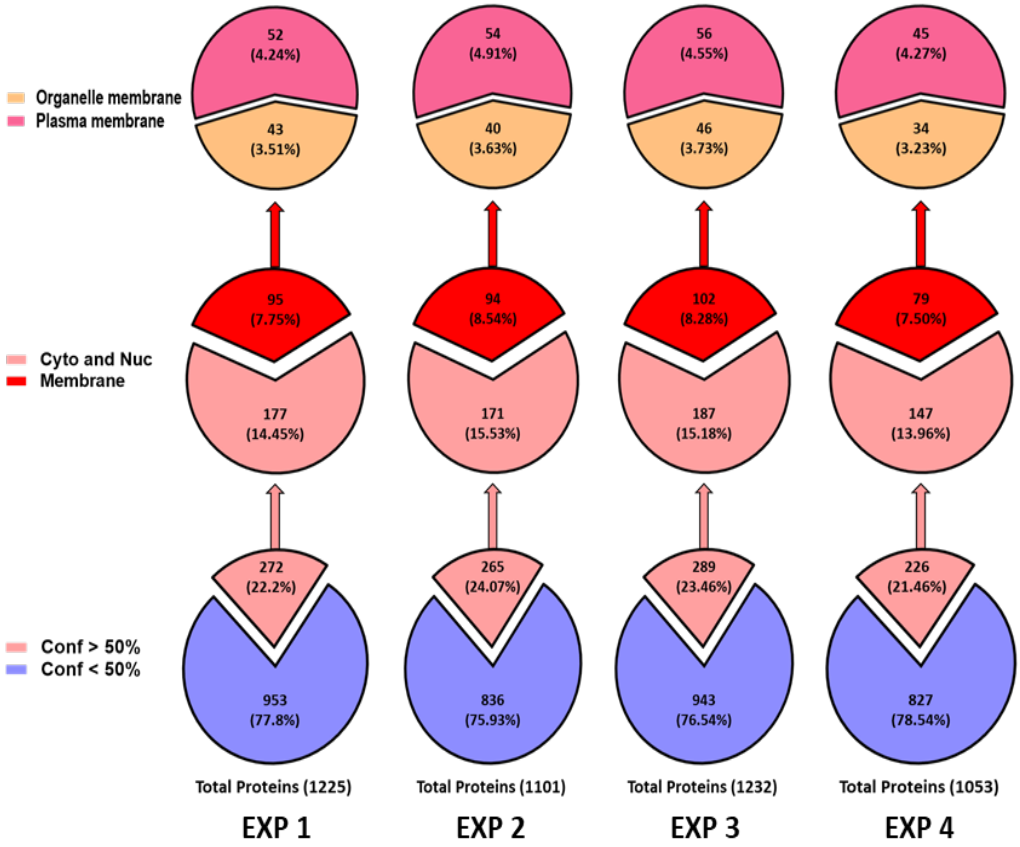


Figure 3.12. Pie of Pie graphs showing the percentages of proteins identified from global profiling of purified murine T cells based on their confidence score and subcellular locations in four separate experiments. EXP: experiment, Conf: Confidence rate of identification, Cyto: Cytoplasmic, Nuc: Nuclear.

3.3.2. Optimisation protocols for subcellular proteomic profiling of mouse CD4⁺ T cells using LC-MS-MS proteomics

3.3.2.1. Optimisation a strategy for sorting a pure population of induced Tregs (iTregs) and natural Tregs (nTregs) prior to MS profiling.

The sorting of iTregs was mainly based on the intensity of CD25 staining. The results showed that over 90% of cells with higher intensity of CD25 expression (> $\log10^1$) were positive for Foxp3 expression and defined as "pure iTregs population ($CD4^+CD25^{++}Foxp3^+$)" (figure 3.13 A), whereas cells with intensity of CD25 lower than $\log10^1$ were mixed populations of $CD4^+CD25^+Foxp3^+$ and $CD4^+CD25^+Foxp3^-$ T cells (figure 3.13 A). The percentage of pure iTregs (gate R3) for sorting was in range between 20% and 23% of total activated $CD4^+CD25^+$ T cells, whereas the percentage

of total iTregs induced from activated $CD4^+CD25^+$ T cells was between 50% and 55% (Figure 3.13 A). Figure 3.15 B showed the gating strategy for sorting the pure iTregs population from total activated $CD4^+CD25^+$ T cells using MoFlo XDP cell sorter. The results showed that the purity of iTregs sorting was in range between 90% and 92%, which means that 90-92% of sorted $CD4^+CD25^{++}$ T cells positive for Foxp3 expression based on the gating strategy which was followed (figure 3.13 B). The percentage of iTregs purity was significantly increased from 50-55% before sorting to 90-92% after sorting (figure 3.13 C).

The sorting of naïve nTregs was also based on the expression of CD25. The results showed that the purity of sorting naïve nTregs from total purified $CD4^+$ T cells was 99.9%, which means that 99.9% of nTregs of total $CD4^+$ T cells were effectively sorted (figure 3.14 A) following the optimised gating strategy (figure 3.15 A). The results also showed that 98% of sorted $CD4^+CD25^+$ nTregs were positive for Foxp3 expression (figure 3.14 A). Similar results were obtained for sorting $CD4^+CD25^-$ T cells as the purity of sorting was also 99.9%. after sorting, naïve $CD4^+CD25^-$ T cells were completely negative for CD25 expression, confirming the effectiveness of the optimised gating strategy (figure 3.14 A). The results also showed that the percentage of purification of nTregs and $CD4^+CD25^-$ T cells was significantly increased from 13% to ~100% and from 84% to ~100%, before and after sorting, respectively (figure 3.14 B).

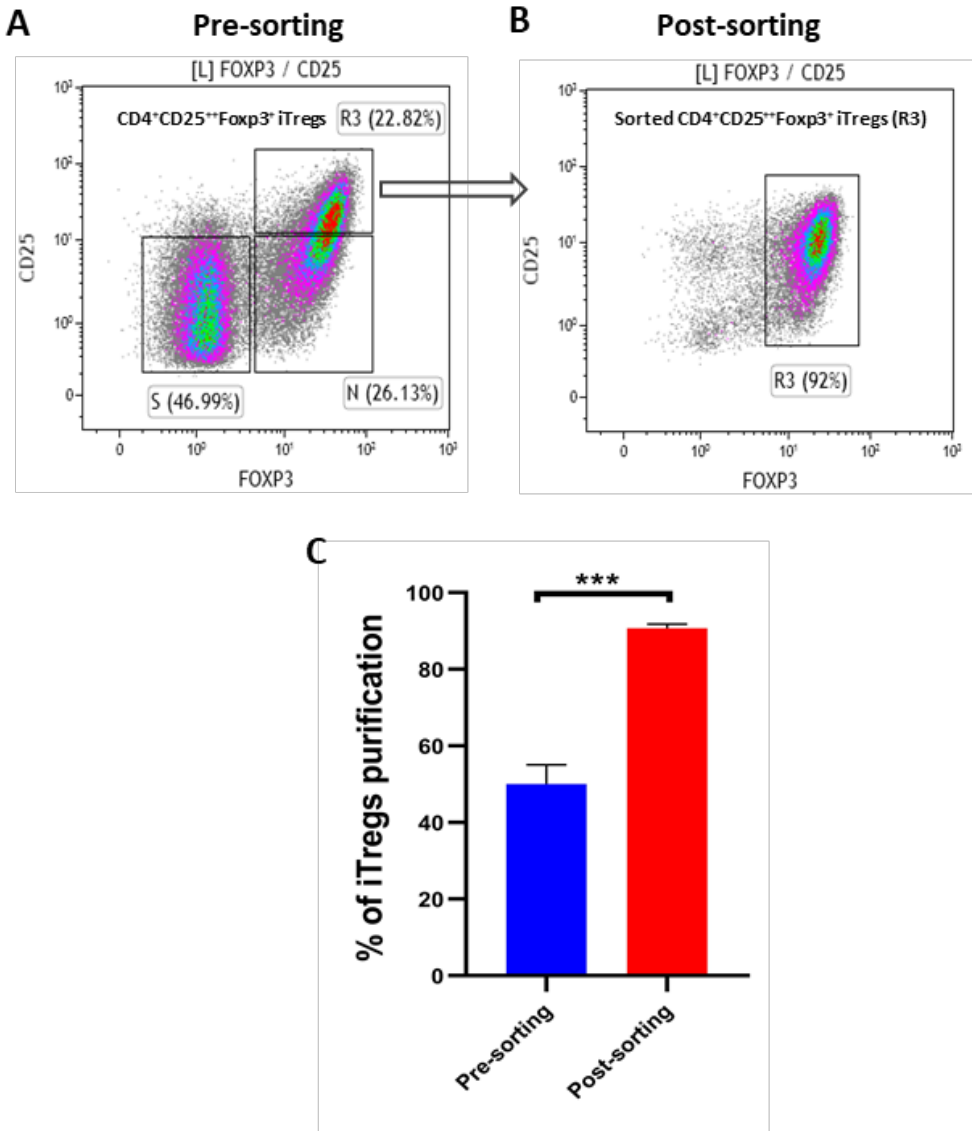


Figure 3.13. Gating strategy for sorting a pure population of iTregs. **(A)** Flow cytometry density plots showing the gating strategy for sorting a pure population of iTregs from total activated $CD4^+CD25^+$ T cells. Gate R3 represents $CD4^+CD25^{++}Foxp3^+$ iTregs (22.82% of total activated $CD4^+CD25^+$ T cells) that were selected for sorting based on the intensity of CD25 staining at $log10^1$. Gate N represents $CD4^+CD25^+Foxp3^+$ iTregs (26.13%) that were mixed with $CD4^+CD25^+Foxp3^-$ T cells (gate S (46.99%). Gates R3 and N represent the percentage of total iTregs (48.95%) induced from the total activated $CD4^+CD25^+$ T cells. Gate S represents the percentage of activated $CD4^+CD25^+$ T cells (46.99%) that were not converted to iTregs. All gates were gated on live $CD4^+CD25^+$ T cells (gate L). the percentage of total $Foxp3^+$ iTregs (gate R3 and N) induced from total activated $CD4^+CD25^+$ T cells (gate L). **(B)** Flow cytometry density plots showing the percentage of $Foxp3^+$ iTregs (R3 (92%)) sorted from $CD4^+CD25^{++}$ T cells based on the intensity of CD25 staining. **(C)** Bar graph representing the percentage of iTregs purification before and after cell sorting. Unpaired T test was used for statistical analysis using GraphPad Prism software 8.0.1. Three independent biological experiments were carried out (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (***; p-value < 0.0002).

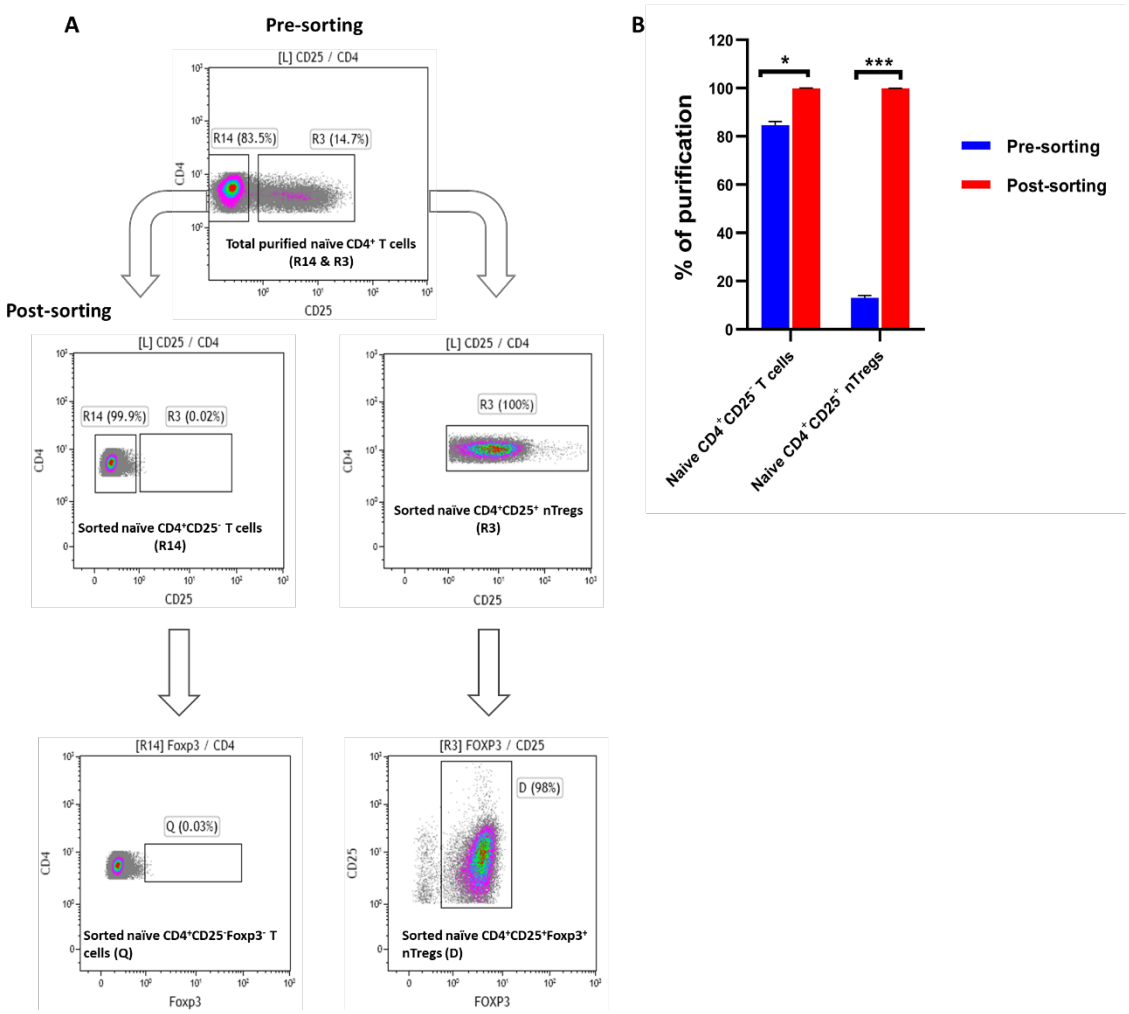


Figure 3.14. Gating strategy for sorting a pure population of naïve $CD4^+CD25^-$ T cells and naïve $CD4^+CD25^+Foxp3^+$ nTregs. **(A)** Flow cytometric density plots showing the gating strategy for sorting naïve $CD4^+CD25^-$ T cells and $CD4^+CD25^+$ nTregs from total population of purified naïve $CD4^+$ T cells. Gate L represents live $CD4^+$ T cells. Gate R14 represents naïve $CD4^+CD25^-$ T cells. Gate R3 represents naïve $CD4^+CD25^+$ nTregs. Both R14 and R3 were gated on gate L. Gate D represents the percentage of $Foxp3^+$ expression gated on the sorted naïve $CD4^+CD25^+$ nTregs. Gate Q represents the percentage of $Foxp3^+$ expression gated on the sorted naïve $CD4^+CD25^-$ T cells. **(B)** Bar graph showing the percentage of purification of both naïve $CD4^+CD25^-$ T cells and $CD4^+CD25^+$ nTregs before and after cell sorting. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=4) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.019, ***; 0.0009). 2way-ANOVA test was usded for statistical analysis using GraphPad Prism software 8.0.1.

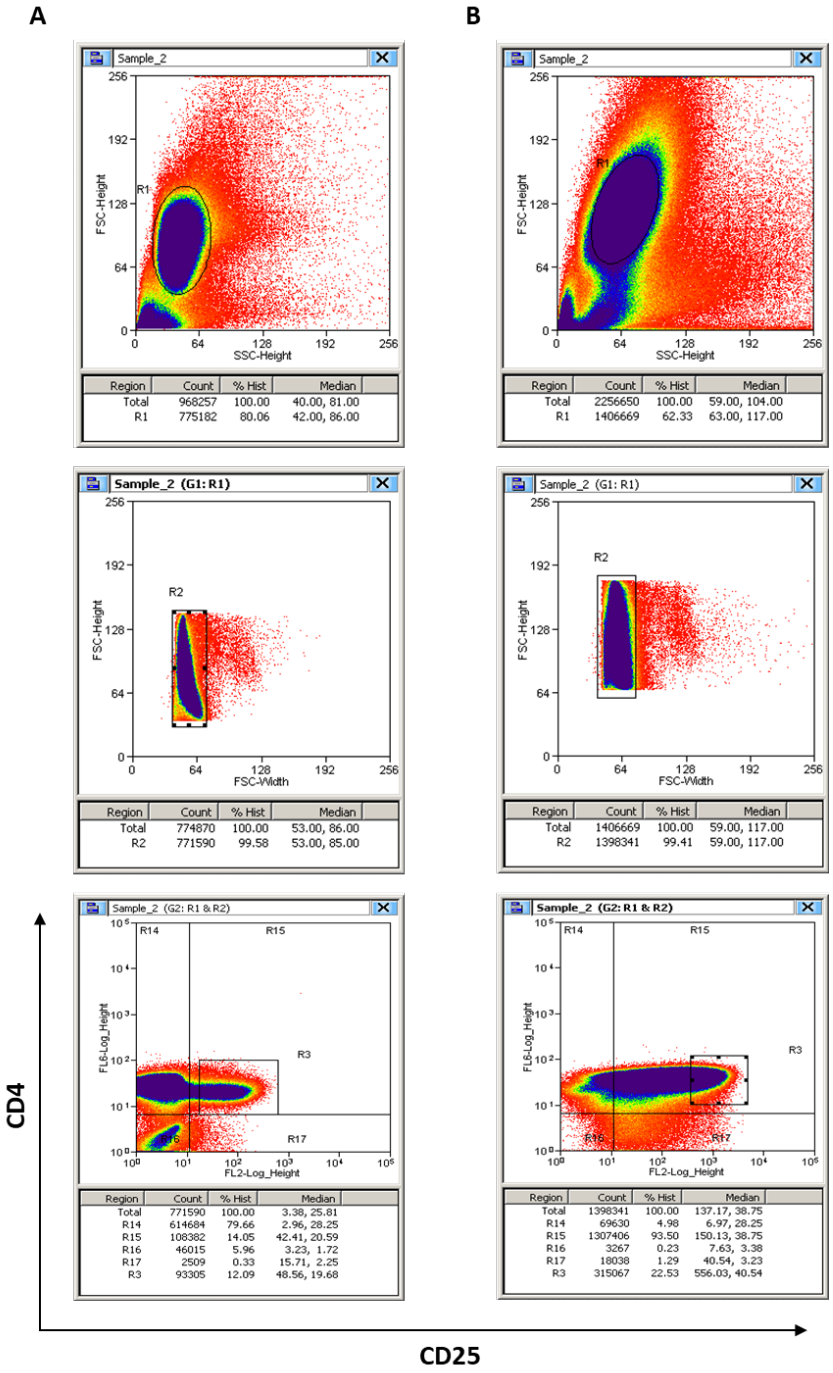


Figure 3.15. Density plots showing the gating strategy that was followed for sorting $CD4^+$ T cell subsets. **(A)** The gating strategy for sorting naïve $CD4^+CD25^-$ T cells (gate R14) and naïve $CD4^+CD25^+$ nTregs (gate R3) from total population of purified naïve $CD4^+$ T cells. **(B)** The gating strategy for sorting $CD4^+CD25^{++}$ iTregs (gate R3) from total population of activated $CD4^+CD25^+$ T cells. All gates were gated on gates R1 and R2 which represent total live cells and total singlets cells, respectively. Density plots were taken from cell sorting software (Summit 5.4).

3.3.2.2. Determination of the number of spleens required to obtain 2.5x10^6 of purely sorted CD4$^+$CD25^{++}Foxp3$^+$ iTregs and CD4$^+$CD25$^+$Foxp3$^+$ natural Tregs (nTregs) prior to subcellular fractionation.

Despite the high purity of sorting, the number of iTregs and nTregs recovered after sorting was very low and insufficient for subcellular fractionation. For instance, the pure population of iTregs represented only 20-23% of total activated CD4$^+$ T cells at 50% of total iTregs induction, which theoretically means that ~ 12.5x10^6 of activated CD4$^+$ T cells are required to obtain 1x10^6 of purely sorted iTregs. For this reason, the number of spleens was increased to obtain enough number of cells after sorting for subcellular fractionation. The results showed that the mean of number of sorted iTregs and nTregs derived from one spleen was 0.53x10^6 each. Whereas the number of sorted iTregs and nTregs was significantly increased to 1.3x10^6 and 1.4x10^6 after processing two spleens. The highest number of purely sorted iTregs and nTregs obtained was 2.6x10^6 and 2.4x10^6, respectively, after processing three spleens (figure 3.16).

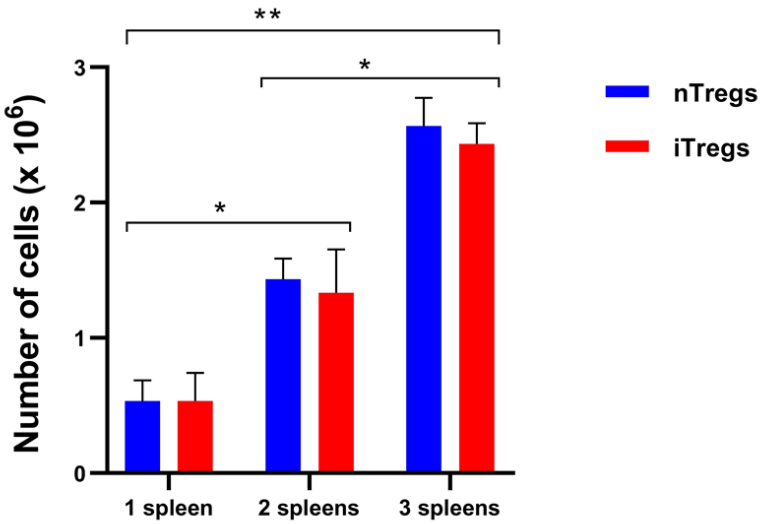


Figure 3.16. Bar graph showing the cell number of purely sorted nTregs and iTregs derived from different number of spleens. 2way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1. Three independent biological experiments were carried out, (n=3). The statistical analysis was carried out on a limited data set (n=3) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (*; p-value < 0.025, **; p-value 0.0016).

3.3.2.3. Subcellular fractionation of CD4+ T cells using Mem-PER Plus membrane protein extraction kit (pilot study).

A pilot study was conducted to assess the efficiency of the Mem-PER Plus extraction kit for subcellular fractionation of CD4+ T cells. The results showed that the percentage of total purified membrane proteins was significantly increased using the fractionation kit. The results showed that the percentage of membrane proteins identification was significantly increased to 57% (mean) using the fractionation kit compared to the whole cell lysate protocol (section 3.3.1.8) where only ~ 8% (mean) of total membrane proteins were identified (figure 3.17). The results also showed a significant increase in the identification of plasma membrane proteins specifically; as 26% (mean) of plasma membrane proteins were identified using subcellular fractionation, whereas only 4.5% (mean) of plasma membrane proteins were identified using whole cell lysate without fractionation (figure 3.17). Subcellular location and functional information of identified membrane proteins were obtained from Uniprot database. Figure 3.18 displays an example of how total membrane proteins were searched and sorted using Uniprot database.

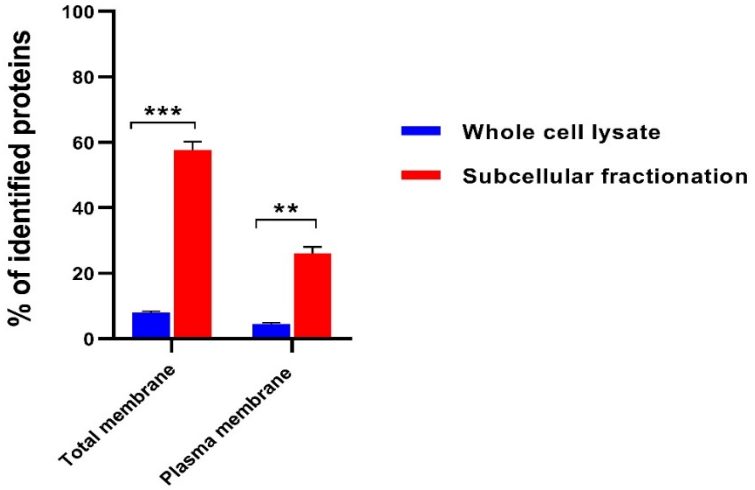


Figure 3.17. Bar graph showing the percentage of identification of total membrane and plasma membrane-bound proteins using two different approaches including whole cell lysate and subcellular fractionation. Four independent biological experiments were carried out, (n=4). The statistical analysis was carried out on a limited data set (n=*4*) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean. (***; p-value < 0.0007, **; p-value < 0.0016). 2way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1.

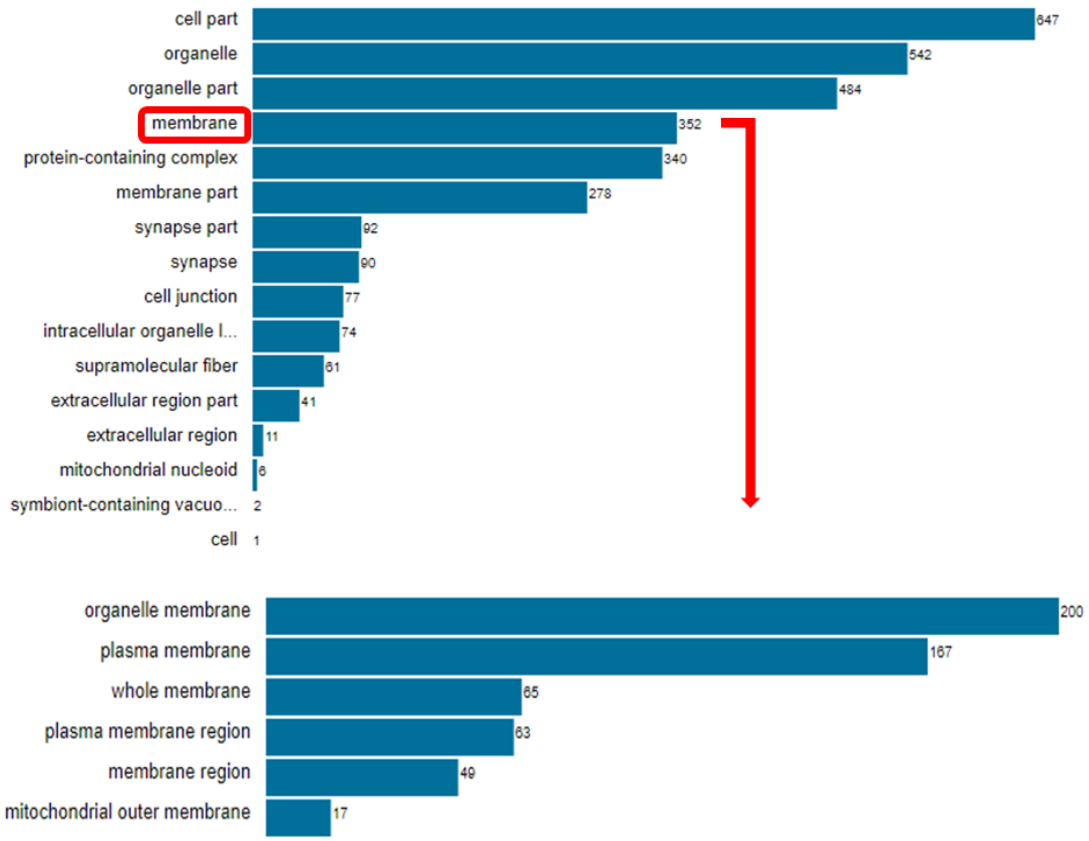


Figure 3.18. An example of Bar graph showing subcellular locations of 647 confident proteins, which was obtained from Uniprot database. 352 (54% of total confident proteins) membrane proteins highlighted by Red rectangle were sorted by Uniprot into different subcellular locations of membrane-bound proteins including plasma membrane proteins (26%) and intracellular membrane proteins (31%) (organelles membrane and mitochondrial membrane).

3.3.2.4. Protein quantitation Assay.

The limitation of using the Mem-PER Plus membrane protein extraction kit was that the enriched cytoplasmic and membrane proteins were resuspended in a high volume of permeabilisation buffer (600 µL) and solubilisation buffer (300 µL), respectively. Therefore, the results of protein assay showed that the concentration of both cytoplasmic and membrane proteins was too low and even out of the experimental linear range of the protein assay. As a result, it was suggested to concentrate the proteins using acetone precipitation and subsequent resuspension in a small volume buffer prior to mass spectrometry.

3.3.2.5. Acetone precipitation of proteins.

The cytoplasmic and membrane proteins were precipitated using cold HPLC-grade acetone to concentrate proteins and replace permeabilisation and solubilisation buffers with MS-compatible reagents. Proteins pellet was clearly seen from the both proteins fractions, however the protein assay was not carried out after acetone precipitation to avoid losing more proteins. Therefore, subcellular proteomic profiling was carried out based on the number of cells not on the concentration of proteins. The number of cells used in each group was uniform and all samples were uniformly processed thereafter.

3.4. Discussion

Advances in the MS-based proteomics instrumentations have enabled a global identification of proteins from complex biological samples, thereby elucidating the overall picture of the expression of functional proteins and their post-translational modifications that regulate protein cellular interactions and functions within a cell or tissue (Aebersold and Mann 2016). In contrast, although the development of transcriptomic profiling technologies, especially RNA-sequencing (RNA-Seq), has provided invaluable insights into discovering networks regulating gene expression at the RNA level, most of transcriptomic data have failed to translate into the clinic. These failures can be attributed to the poor correlation of transcriptome data to the protein expression. This highlights the importance of proteomic analysis as a powerful technology to investigate and discover the proteins as biomarkers and also druggable targets in the clinical settings (Kumar, et al. 2016).

In the last decade, several studies have been carried out to the study the development and differentiation of human and murine $CD4^+$ T cells using high throughput transcriptomic and proteomic analysis (Elo, et al. 2010, Tuomela, et al. 2012, van den Ham, et al. 2013, Tuomela, et al. 2016, Proserpio, et al. 2016, Mohammad, et al. 2018, van den Ham, et al. 2019). However, all of these studies were focused only on the differentiation of T helper 1 (Th1), T helper 2 (Th2) and Th17 cells. Limited number of studies has analysed the development and function of Tregs using global proteomic analysis. Procaccini et al. have carried out a global comparative proteomic analysis between human $CD4^+CD25^{hi}$ $Foxp3^+$ $CD127^-$ Tregs and $CD4^+$ $CD25^-Foxp3^-$ T conventional (T conv) cells at the metabolic level. The authors have identified a set of distinct metabolic proteins which have a pivotal role in the proliferation of Tregs and T conv cells (Procaccini, et al. 2016). Another study by Duguet et al. has applied global label-free quantitative proteomics to identify a proteomic signature of both $CD4^+Foxp3^+$ Tregs and $CD4^+Foxp3^-$ T cells. The authors have identified a specific nuclear protein called “Themis-1” that regulates the suppressive function of Tregs (Duguet, et al. 2017). A recent study by Cuadrado et al. has been conducted to identify human Tregs cell identity at the molecular levels using label-free quantitative proteomics in a combination with transcriptomic analysis. The authors have compared naïve subpopulations of $CD4^+CD25^{hi}CD45RA^+$

Tregs and $CD4^+CD25^-CD45RA^+$ T conv cells with $CD4^+CD25^{hi}CD45RA^-$ effector Tregs and $CD4^+CD25^-CD45RA^-$ memory T conv cells and identified distinct proteins that govern and maintain cellular identity of each of the subpopulations (Cuadrado, et al. 2018). Another recent study by Schmidt et al. has also applied global transcriptomic and proteomic analysis to study the differentiation of human induced Tregs (iTregs) from naïve $CD4^+CD25^-$ T cells based on the induction of the expression of Foxp3 at different time points. The authors have found a panel of 37 distinct nuclear molecules that function as regulators of Foxp3 (Schmidt, et al. 2018). A final recent study by Mohammad et al. has analysed and compared the proteome signature of murine Th17 and iTregs during their differentiation *in vitro* from naïve $CD4^+CD62L^+$ T cells using global label-free proteomic profiling. The authors have only focused on the metabolic pathways that are implicated in the differentiation of Th17 and iTregs (Mohammad, et al. 2018).

However, all of these studies mentioned above have not analysed the difference in the phenotypic features between natural Tregs (nTregs) and induced Tregs (iTregs) at the level of cell-surface membrane proteins, particularly plasma membrane proteins that are the main targets in the development of therapeutic agents, using subcellular fractionation. To date, no studies have reported the identification of membrane biomarkers that can distinguish induced Tregs from other subsets of $CD4^+$ T cells, therefore it is an area which required further investigation. Identification of such biomarkers would help facilitate the development of therapeutic agents that can selectively target iTregs which could be generated as a result of highly inflamed tumour microenvironment.

Accordingly, in this study, two different methods; global (total cell lysate) and subcellular fractionation were optimised to identify membrane biomarkers that could differentiate iTregs from other subpopulations of $CD4^+$ T cells using a state-of-the-art MS-based instrumentation TripleTOF 6600 mass spectrometer with SWATH-MS acquisition, SCIEX. The novelty of this study is represented by optimising a protocol for subcellular fractionation of membrane and cytoplasmic proteins derived from five different subpopulations of $CD4^+$ T cells for label-free quantitative proteomic analysis.

The results of optimisation of global proteomics profiling showed that the total number of proteins identified from whole cell lysate containing 2.5x106 T cells was in range between 1053 and 1232 at 1% FDR. Out of which, the total number of membrane proteins with confidence score ≥ 50% was in range between 79 (7.5%) and 102 (8.3%) (n=4), respectively. The percentage of plasma membrane proteins identified was in range between 4.2% and 4.9% from the total number of proteins. These results were not sufficient to meet the main aim of this stud as the optimisation of proteins isolation using whole cell lysate lysed by Erika's or urea buffer has failed to purify sufficient percentage of membrane proteins. There are several limitations associated with the purification of plasma membrane proteins including low recovery rate, protein insolubility and hydrophobicity (Orsburn, Stockwin and Newton 2011). These limitations could justify the low percentage of membrane proteins identified by the global whole cell lysate protocol. As a result, subcellular fractionation was approached to overcome the limitation of the whole cell lysate method for isolation membrane proteins.

As mentioned in the introduction (section 3.1.2.7), there are several methods for subcellular fractionation such as density-gradient centrifugation with sucrose medium as a common conventional method. However, some of these methods require a long processing with multiple centrifugations that might result in the loss of proteins during the isolation process. Also, the number of cells ($2x10^6$) from which membrane proteins were isolated has put another barrier for the selection of an appropriate method for the subcellular fractionation. For these reasons, the Mem-PER™ Plus membrane proteins extraction kit was selected for the isolation of membrane and cytoplasmic proteins from $CD4^+$ T cells since it has been optimised for isolation membrane and cytoplasmic proteins in 2 hours using permeabilisation and solubilisation buffers. The time required for processing samples was of the main aims for isolation membrane proteins from *in vitro* cultured cells, especially induced Tregs, to prevent experimental variations in the processing of samples. This kit was the most appropriate among other commercially available kits since it has been optimised to extract both membrane and cytoplasmic proteins from $5x10^6$ T cells, whereas other kits require at least $10x10^6$ or more for extraction membrane proteins which is 5 times higher than the cell number used in this study.

The results showed that the mean of the percentage of membrane proteins identified following subcellular fractionation was significantly increased to 57% compared to the results of whole cell lysate where the mean of the percentage of membrane proteins was only 8%. The results also revealed that the mean of the percentage of the plasma proteins was significantly increased to 26% of the total proteins identified following subcellular fractionation, while it was only 4.5% using the protocol of whole cell lysate. These results were relatively satisfactory since only $2x10^6$ cells were processed for the isolation of membrane proteins.

In summary, in this chapter, a subcellular fractionation method has been successfully optimised to isolate membrane proteins from $2x10^6$ of murine $CD4^+$ T cells for label-free quantitative proteomic analysis. 57% of membrane proteins and 26% of plasma membrane proteins were successfully identified from $2x10^6$ cells. These results are sufficient to conduct a comprehensive quantitative proteomic profiling of five different subpopulations of murine $CD4^+$ T cells using LC-MS-MS and SWATH-MS proteomics which will be discussed in more details in the next chapter (chapter 4).

Chapter 4.

Identification, selection and verification of novel membrane markers of natural and induced T regulatory cells

4.1. Introduction

T regulatory cells (Tregs) play a pivotal role in maintaining immune tolerance by eliminating autoreactive T cells that induce immune response against self-tissues. However, in cancer, Tregs play a detrimental role by inhibiting the proliferation and activity of $CD8^+$ effector T cells, thereby promoting the tumour progression and metastasis. The presence of tumour-infiltrating Tregs in the tumour microenvironment is of main obstacles for successful immunotherapy (Liu, Workman and Vignali 2016b). It has been proposed that the tumour-infiltrating Tregs might be nTregs which are recruited into the tumour microenvironment or iTregs that are developed from circulating naïve $CD4^+CD25^-$ T cells (Chaudhary and Elkord 2016, Deng 2018). To date, distinct biomarkers that can differentiate between thymic-derived natural Tregs(nTregs) and peripheral-derived induced Tregs (iTregs) are still required to be identified for specifically targeting iTregs. This chapter will discuss the process of identification and verification of biomarkers using SWATH-MS proteomics.

4.1.1. SWATH-MS based proteomics

SWATH-MS (sequential window acquisition of all theoretical mass spectra) acquisition was initially developed to process the complicated datasets yielded by data independent acquisition (DIA)-MS analysis. SWATH-MS acquisition combines the high reproducibility and sensitivity of a DIA-MS approach and a targeted data extraction strategy to read out the complex datasets of fragments ions (Gillet, et al. 2012) (figure 4.1).

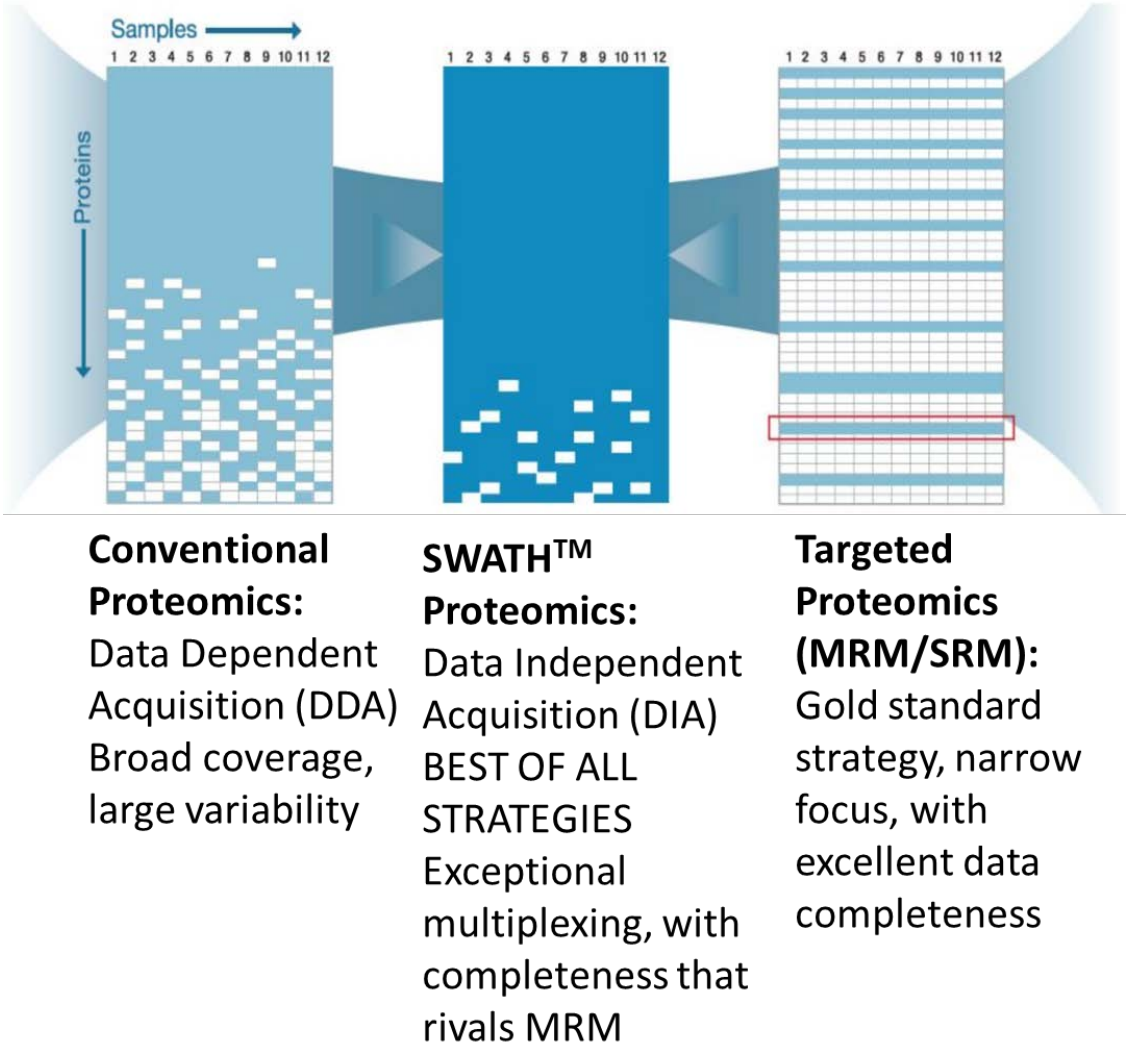


Figure 4.1. Data Dependent versus Data Independent versus Selected Reaction Monitoring. A comparison between the three different strategies for quantitative identification of proteins. DDA is biased untargeted acquisition that is widely used in conventional proteomics. MRM/SRM is targeted strategy used in targeted proteomics which is used to selectively analyse proteins of interest. Targeted proteomics require prior knowledge about the protein of interest; therefore, it is not suitable for biomarker discovery despite of its high sensitivity and accuracy. SWATH-MS has been developed to combine the features of conventional proteomics in terms of biomarker discovery with unlimited identification of proteins and the feature of targeted proteomics by which SWATH-MS allows to selectively analyse the identified proteins in a targeted data processing and analysis. Image was taken from Sciex.com website. Copyright permission was obtained from AB Sciex team.

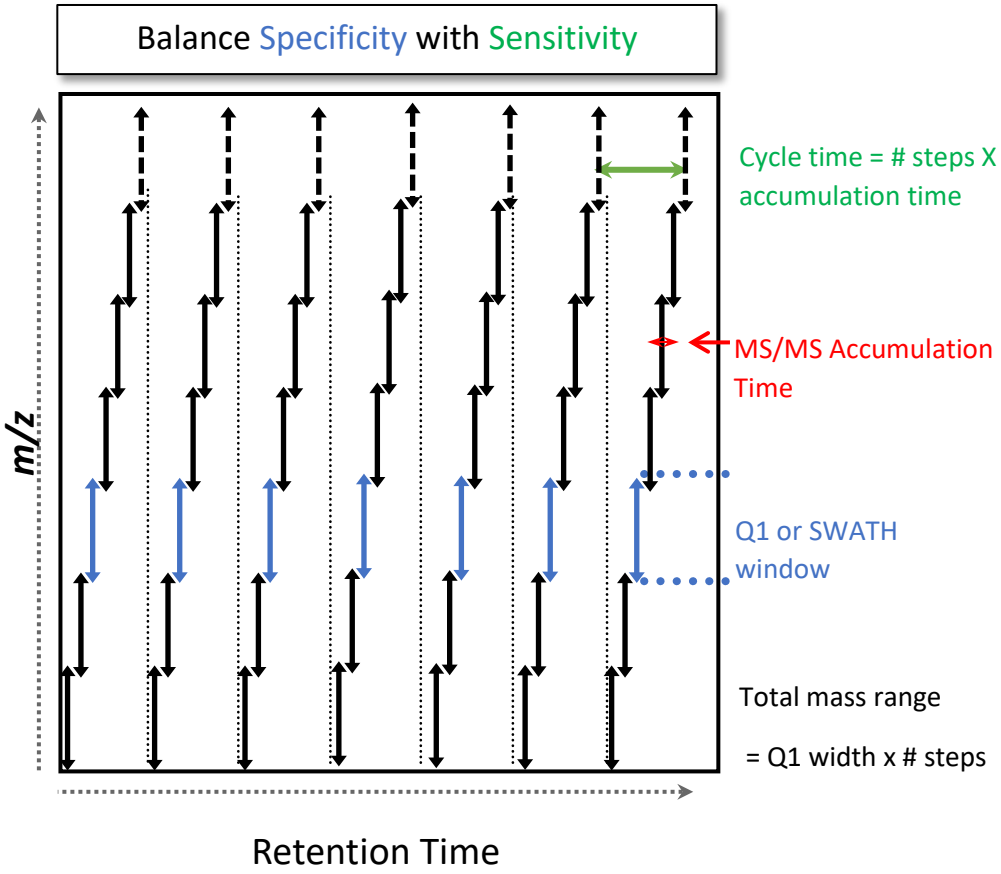


Figure 4.2. Q1 Isolation Strategy in SWATH. The width of the SWATH or Q1 window in blue has a correlation with the specificity of the analysis. In a fixed SWATH window strategy, the smaller SWATH, window the more specific (fewer different peptides will be fragmented) the analysis will be and vice-versa. Sensitivity of the analysis directly correlates with the cycle time (how long the instrument takes to fragment all the SWATH windows) and accumulation time (how long the instrument dwells on fragmenting each packet of precursor ions (contained in one SWATH window). These can be optimised depending on the chromatographic conditions of the analysis, for analyses where each peptide peak from the chromatographic column is around 20 seconds wide, then approximately 2-3 seconds cycle time would be optimal to get enough data points for quantitation across a peak. Typical accumulation times would be 25-50 msec. . Image was taken from Sciex.com website. Copyright permission was obtained from AB Sciex team.

The original framework of SWATH-MS acquisition consisted of 32 sequential, slightly overlapping precursor isolation windows with isolation width of 25 *m/z* each. These windows were constructed to cover the mass range 400 – 1200 *m/z* (figure 4.2).

The key principle of SWATH-MS is to obtain optimal correlation between precursor and fragment isotopes peaks at any time they elute from LC. This is achieved by setting consecutive isolation windows with (+1 *m/z*) overlap between windows to

ensure the transfer of any given precursor ions in at least one window. More recently SWATH methodologies have been optimised to use a higher number of variably sized windows (typically 100 windows), with larger windows in peptide sparse *m/z* regions and smaller windows over the *m/z* range where more peptides exist in the sample (figure 4.2).

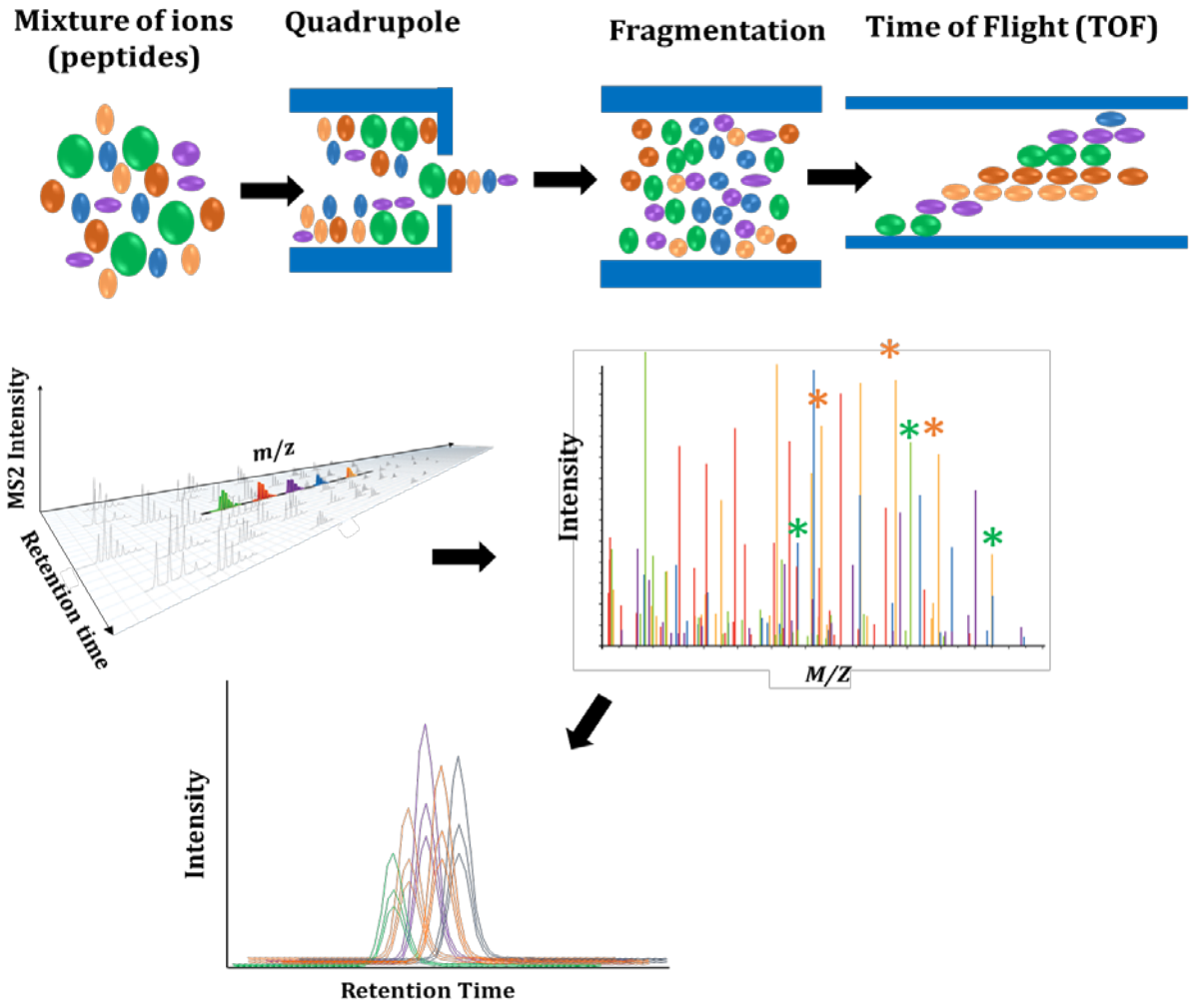


Figure 4.3. Schematic representation showing the workflow of SWATH-MS acquisition. In SWATH-MS mode, once peptides become ionised by the ion source, they enter the first mass analyser (quadrupole Q1) which scans and records the *m/z* values of eluted peptides. Ions of peptides within a "SWATH window" are then fragmented (second quadrupole or Q2/collision cell) to produce small fragment ions that enter the second mass analyser (TOF) which records the *m/z* values of the fragment ions with respect to the retention time and their intensity. By plotting the fragment ions within a SWATH-MS window over time, specific ions spectra (coloured fragments) can be deconvoluted and assigned to a "parent" or precursor ion (peptide), based on the theory that the same type of molecule eluting off the chromatographic column at the same time will share the same peak apex. These peptide ion fragments can then be matched to sequences and retention times using a pre-generated spectral or ion library derived by more standard Information/Data Dependent Acquisition (IDA/DDA) mode. This figure was adaptive from (Ludwig, et al. 2018).

SWATH-MS acquisition can also deal with complex samples in which many peptides of different size are able to co-elute simultaneously from the LC in a process called "coelution". These co-eluting peptides are then differentially selected in the precursor ion selection window and undergo co-fragmentation by the scheme of SWATH-MS acquisition, generating multiplexed and complicated spectra of fragment ions (Ludwig, et al. 2018) as schematically represented in figure 4.3. To handle this issue, a novel data analysis approach has been proposed by Gillet et al., which is based on peptide centric scoring that requires previous information of a peptide of interest. This is called "peptides query parameters (PQPs) which can be acquired from previously generated spectral libraries (Gillet, et al. 2012). PQPs include the sequences of peptides that belong to proteins, the optimal or dominant precursor ion *m/z* values of peptides, four to six fragment ions with highest *m/z* values and intensities, precursor and fragment ions charges, types of fragment ions and normalised retention time of both precursor and fragment ions (Collins, et al. 2017). The primary strategy of SWATH-MS as a targeted data extraction method using information from spectral libraries has been modified, automated and improved by other tools such as OpenSWATH and SWATHProphet which accurately identifies and quantifies peptide fragment ions and automatically uncovers and removes interfering precursor ions (Röst, et al. 2014, Keller, et al. 2015). The main advantage of SWATH-MS acquisition is its ability to quantitatively analyse peptides covering thousands of proteins with a high consistency and accuracy. It is perfectly tuned for projects that require proteomic analysis of a large number of samples with more accuracy and reproducibility (Ortea, et al. 2016). However, there are two main drawbacks of SWATH-MS. First, it is still significantly less sensitive in peptide quantification compared to the conventional methods of targeted proteomics (SRM and PRM) which are the best option for quantifying low abundant proteins. Second, SWATH-MS, until recently required a prerequisite spectral library and/or PQPs for peptides and proteins quantification, which adds additional effort to the data analysis (Schmidlin, et al. 2016). While SWATH/DIA data can now be searched directly against known proteomes by various software in a way similar to IDA/DDA data (PEAKS Studio X from Bioinformatic Solutions Inc, Spectronaut Pulsar from Biognosys AG) it is more optimal and much faster to reduce the search space with a reduced list of peptide and peptide fragments corresponding to proteins at known

retention times in the mass spectrometer's acquisition file. This is termed a spectral library. The most basic spectral library contains the precursor ion *m/z* (Q1 *m/z* or peptide ion), the fragment *m/z* (Q3 *m/z* corresponding to the peptide sequence/amino acids), retention time of the peptide, and the protein accession it is part of. A spectral library may also contain other optional information such as charge state, intensity, modifications etc. The key point to remember if searching using a spectral library – if the protein/peptide is not in the library, then no quantitative data can be determined for that peptide/protein. It is therefore of paramount importance to have as comprehensive a spectral library as possible. To this end there are various projects, such as the SWATHAtlas project at the Institute of Systems Biology, Seattle, USA that are generating large high quality spectral libraries that are publically available (http://www.swathatlas.org/) to complement locally generated libraries from the same samples that are being quantitified. This is useful as the nature of SWATH/DIA means it can be inherently more sensitive than traditional IDA/DDA workflows which alone may not generate a good quality comprehensive library.

4.1.2. Biomarker validation – journey toward the clinic

The evolution of candidate biomarkers undergoes rigorous steps in order to be clinically approved. The first step is pre analytical validation which focuses on preparation and processing of samples that will be tested using a proposed assay. Several factors can affect the preparation of samples including time and storage conditions between sample collection and processing, the efficiency of reagents used for sample processing and the time of sample processing (Henry and Hayes 2012). The second step is analytical validation which tests the technical efficiency of the biomarker assay through assessing sensitivity, specificity, limit of detection, linearity and potency of the assay. In addition to the accuracy and reproducibility of the biomarker assay (Masucci, et al. 2016). The final step is clinical validation which assesses the clinical validity and utility of the biomarker. To achieve this, several parameters must be measured including clinical sensitivity and specificity, false positive and false negative values as well as relative risk vales of the biomarker (Dobbin, et al. 2016).

4.1.3. Proteomics biomarker pipeline: from discovery to the clinic

The validation of proteomic biomarkers is a step-wise process consisting of long-term distinct steps (figure 4.4). First, the discovery or screening step where candidate biomarkers are identified using shotgun proteomics. This step is usually carried out on small number of samples in which proteins will be extracted and analysed by LC -MSMS proteomics. Two main methods are used to quantify proteins in samples including label based and label free methods which have been mentioned in chapter 3 (section 3.1.2.6) (Surinova, et al. 2010). Second, the verification step by which proteins with high fold change values between the samples are selected for verification using targeted proteomics analysis. This step requires more samples to be analysed than the first step (Hathout 2015). Different methods can be used for targeted proteomics analysis including SRM MS and MRM MS (mentioned in chapter 3, section 3.1.2.5.2) and SWATH-MS. The final step is the validation of candidate biomarkers in the clinic. In the validation step, large cohort of samples is needed to assess the clinical validity and utility of the biomarker (Parker and Borchers 2014). Due to these long, expensive and uncertain steps, a large number of proteomics biomarkers has failed to reach the final stage of clinical validation (Chauvin and Boisvert 2018).

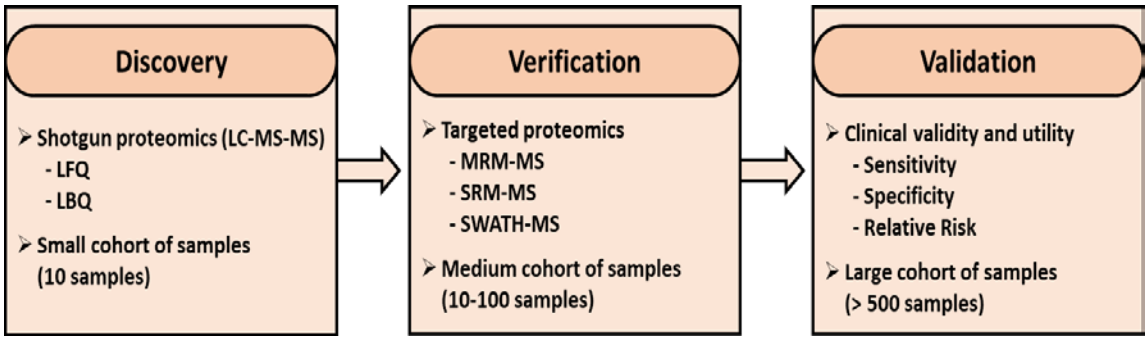


Figure 4.4. Schematic steps for the development of proteomics biomarkers. LC-MS-MS: liquid chromatography-tandem mass spectrometry, LFQ: label-free quantification, LBQ: label-based quantification, MRM: multiple reaction monitoring, SRM: selected reaction monitoring, SWATH: sequential window acquisition of all theoretical fragment ion spectra. This figure was adapted from (Chauvin and Boisvert 2018) and (Parker and Borchers 2014).

4.1.4. Methods for biomarker verification

4.1.4.1. Targeted proteomics

SRM (Selected Reaction Monitoring, also known as MRM, multiple Reaction Monitoring) is one of the LC-MS-MS targeted proteomics methods that have been

developed complementary to untargeted shotgun proteomics. SRM has substantial potential to close the chasm between biomarker discovery and the translation to clinical use (Harlan and Zhang 2014). It involves simultaneous targeted analysis of peptides that play as surrogates for the protein targets of interest which are usually selected from discovery proteomics assays. SRM based assays are highly specific, accurate and they can be multiplexed assays in which enormous number of peptides can be quantified in one assay (Thomas and Zhang 2016). However, SRM assay have shown some challenges including the detection sensitivity for the targeted proteins, establishment of a standardised methods for samples preparation and maintaining reproducibility of targeted peptides (Thomas and Zhang 2016, Picotti and Aebersold 2012b). Also, SRM requires tremendous optimisation for selecting the optimal ionised peptides that carefully represent the protein of interest, assessing linearity and quantitative dynamic range coverage of targeted peptides (Hathout 2015).

SWATH-MS, which is not considered as targeted proteomics approach, but is targeted data analysis method, has developed with potential to overwhelm some of SRM challenges. SWATH-MS is a powerful method for biomarker verification as a broad range of identified peptides that uniquely represent the identity of a protein can be selected for further analysis to measure the fold change of a proteins between samples (Thomas and Zhang 2016, Narasimhan, et al. 2019).

4.1.4.2. Enzyme linked immunosorbent assays (ELISA)

ELISA technique is the most common assay for biomarker validation as it can efficiently measure many samples simultaneously with low experimental variation. This technique is mainly used to validate secreted proteins (Katsila, et al. 2014, Chauvin and Boisvert 2018). As antibody based technique, ELISA requires highly specific antibodies that bind with high affinity to the target biomarker to assess the reliability and accuracy of the biomarker assay (del Campo, et al. 2015). Four different assays of ELISA are available including direct, indirect, sandwich and competitive ELISA. Among these assays, sandwich ELISA is the most accepted assay used in biomarker verification and validation as it is highly specific and sensitive (Aydin 2015). By this assay, the target biomarker will be detected by two different antibodies; capture and detection, where each antibody is specific for a particular epitope of the target biomarker (Aydin 2015). However, for numerous candidate

biomarkers, commercially specific antibodies are not available and required to be developed initially in order to design a new ELISA assay. The development of a novel ELISA assay entails an attentive design to provide more accurate measurements for the target biomarker (Kelley and DeSilva 2007). The workflow for designing ELISA assay includes antibody design, ELISA development, assay validation and clinical assessment. The most important step is the antibody design which requires optimal selection of the epitope (unique peptides) of the candidate biomarker detected in the proteomics discovery stage (Cox, et al. 2004). Selection the optimal epitope improves the specificity and sensitivity of the developed antibody and therefore increases the robustness of the biomarker assay. Thus, it is essential to have information about biological traits of the protein (marker) of interest, including its 3D structure, binding sites, hydrophobicity and post translational modifications. For examples, if the selected epitope of the protein of interest belongs to a highly glycosylated or hydrophobic portion of the protein, it may not be appropriate for ELISA as the target epitope can be masked by native conditions (Ramos-Vara 2005, del Campo, et al. 2015).

4.1.4.3. Flow cytometry

Flow cytometry (FC) is a powerful method that is widely used in verification and clinical validation of new biomarker. FC is used to assess the phenotypic and functional features of cells and quantify large number of cells by assessing their subsets percentage, distribution, activation status and other cellular functions (Baumgarth and Roederer 2000). Different assays can be carried out using FC, including immunophenotyping, intracellular staining, cell cycle, cell proliferation, apoptosis and phosphor flow assays. These applications enable FC to have more advantages over other similar assays such as ELISA and ELISPOT (enzyme linked immunospot) which measures only the total amount of secreted proteins or cytokines and detects a cell that secretes a specific protein or cytokine, respectively (Millán and Brunet 2015). In oncology, immunophenotyping is widely used in clinical laboratories for diagnosis and classification of various hematologic malignancies and leukaemias. The availability of clinically useful antibodies has enabled profiling complex immunophenotypic signatures. In addition to the advances in FC multicolour instrumentations that allow profiling co expression of different cell surface markers or cytokine receptors in a single experiment,

facilitating identification phenotypic signatures of specific subpopulations of cells (Maecker, McCoy and Nussenblatt 2012). Metrock et al. (Metrock, et al. 2017) have demonstrated the clinical importance of immunophenotyping as a non-invasive procedure for monitoring and diagnosing paediatric patients who are suspected to have acute leukaemia. The authors confirmed that immunophenotyping by FC could be more sensitive and specific for detecting the residual tumour cells after chemotherapy than bone marrow aspiration which is an invasive procedure (Metrock, et al. 2017). FC can also be used to predict the prognosis of many cancers by analysis the enrichment of T cells in the peripheral blood or tumour sections of patients. For example, in melanoma, low baseline of myeloid derived suppressor cells (MDSCs) and Tregs frequency, and high absolute eosinophil counts has been found associated with better outcome in patients who received ipilimumab (Martens, et al. 2016). Besides immunophenotyping, intracellular staining is an essential method used to assess functional pathways operating in the cytoplasm of a specific cell population. In immunology, intracellular staining is mainly used to detect expression of functional cytokines which are produced by a specific cell population of immune cells during an immune response. For instance, in organ transplantation, intracellular staining has been used to assess the production of interleukin 2 (IL 2), IL 10 and interferon gamma (IFN γ) as candidate biomarkers to predict the outcome of tissue grafting (Benítez and Najafian 2008). Akoglu et al. have found that IL 2 secreting $CD8^+$ T cells were significantly enriched at the site of rejected liver tissue in patients who underwent liver transplantation, confirming that IL 2 production by $CD8^+$ T cells could be as a biomarker for acute tissue rejection after organ transplantation (Akoglu, et al. 2009). Millán et al. have also identified that the percentage of IFN γ secreting $CD4^+$ and IFN γ/IL 2 secreting $CD8^+$ T cells was significantly increased in the peripheral blood of patients who experienced tissue rejection after receiving long term immunosuppression treatments for stable liver transplantation. The authors have suggested that soluble concentration and the intracellular expression of IL 2 and IFN γ could be used as predictive biomarkers of the risk of T cell mediated tissue rejection (Millán, et al. 2010). However, despite the advances in the application of FC, this technique has several limitations. First, complexity of optimisation the panel of fluorochrome conjugated antibodies which requires a careful multicolour compensation between different fluorochromes to

prevent signal overlapping between antibodies. Second, lack of method standards and external quality controls which are still needed for improving analysis and reporting of FC data and minimising subjectivity in the analysis of FC data (Millán and Brunet 2015).

4.1.4.4. Bioinformatics

The biological data generated from high throughput omics approaches such as proteomics for the development of clinical biomarkers have become significantly complex and exceeded the potential of conventional data analysis (Luscombe, Greenbaum and Gerstein 2001). Bioinformatics has developed in the past years and played an essential role in biomarker discovery and validation. It is a multidisciplinary field combining biology, computational science, mathematics and statistics to deduce knowledge from biological omics data (Mount and Pandey 2005). Database resources, computational algorithms and statistical models are usually involved in the framework of bioinformatics to efficiently manage, process and analyse the complex biological data (Huang, Sherman and Lempicki 2008). For instance, proteomics approaches integrated with bioinformatics tools have markedly promoted the discovery of new novel serum biomarkers to detect cancer (Ueda, et al. 2010).

4.1.4.4.1. Artificial neural networks (ANN) as a tool of bioinformatics

ANNs are a form of machine learning approaches and have been initially developed to determine a panel of biomarkers based on their high predictive performance (Lancashire, Rees and Ball 2008a). The ANNs are often used to model complex biological datasets such as gene and protein expression datasets based on predictive performance and predictive test error. The architecture of ANNs is derived from the workings of human brain, which consists of an artificial network of computational elements (neurons) that are mathematically interconnected to perform the learning process (Zafeiris, Rutella and Ball 2018). The principle of ANNs is to create multiple models based on a stepwise approach to predict a distinct panel of biomarkers that can differentiate between various samples within a dataset (Lancashire, Rees and Ball 2008b). The schematic workflow of the ANNs stepwise modelling approach is shown in figure 4.5. Due to the complexity of proteomics data, ANNs have been used

to analyse proteomics data in order to identify and verify a panel of novel biomarkers based on their predictive performance (Swan, et al. 2013).

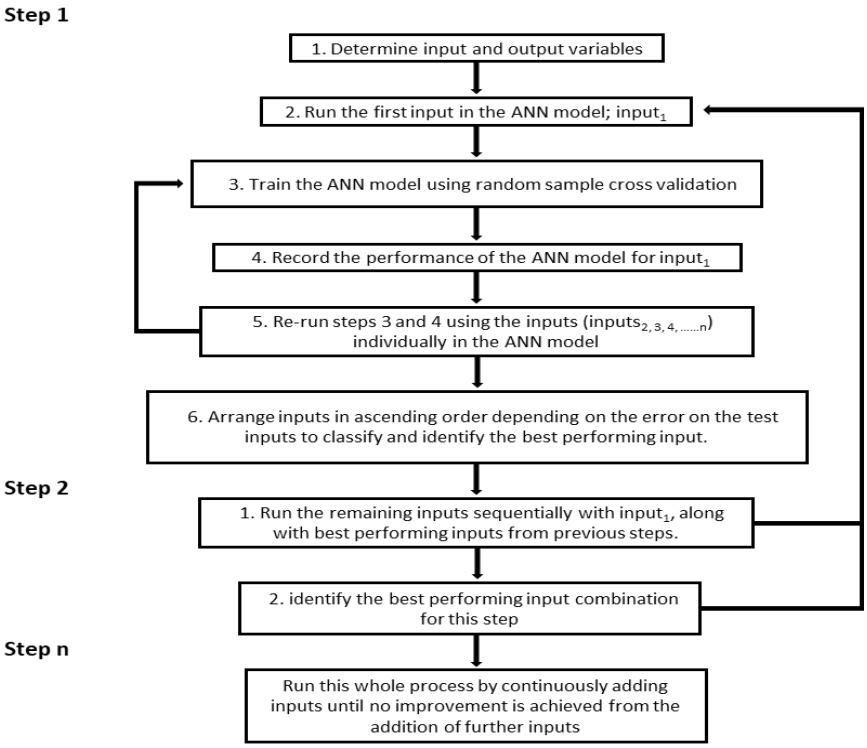


Figure 4.5. The schematic workflow of the ANNs stepwise modelling approach. This figure was adaptive from (Lancashire, Rees and Ball 2008b).

4.1.5. Aims of this study

The aim of this chapter is to generate proteomic profiles to enable discovery of biomarkers (differentially expressed proteins) of induced Tregs, which can differentiate them from natural Tregs and other subpopulations of $CD4^+$ T cells.

The objectives of this chapter are:

1. Confirm purified distinct populations of induced and natural Tregs and other control populations.
2. Generate a list of all proteins and peptides identified in all Treg populations which can be used as a Spectral Library (ion library).
3. Generate quantitative SWATH MS profiles for subpopulations of Tregs.
4. Process data to identify differential biomarkers of iTregs.
5. Pathway analysis of differentially expressed proteins.
6. Verification and validation of identified biomarkers using bioinformatics tools and flow cytometry.

4.2. Materials and Methods

4.2.1. Proteomic profiling of mouse $CD4^+$ T cells using mass spectrometry

Five different subpopulations of $CD4^+$ T cells were purified and sorted to create a library of subcellular proteins of $CD4^+$ T cells, including naïve $CD4^+CD25$ $Foxp3^-$ T cells, naïve $CD4^+CD25^+Foxp3^+$ nTregs, activated (stimulated) $CD4^+CD25^+Foxp3^+$ nTregs, activated (stimulated) $CD4^+CD25^+Foxp3^-$ T cells and activated (stimulated) $CD4^+CD25^{++}Foxp3^+$ iTregs. Five separate replicates (samples) were processed from each of $CD4^+$ subpopulations, each replicate contained $2x10^6$ cells. Cytoplasmic and membrane proteins were isolated from each sample using the Mem PER™ Plus membrane proteins extraction kit as mentioned in chapter 3 section (3.2.2.3). Then proteins were precipitated using cold acetone (HPLC grade) as mentioned in section (3.2.2.4). Then protein samples were processed for MS analysis as mentioned in section (3.2.2.5) by the proteomics facility in the John van Geest Cancer Research Centre at Nottingham Trent University. Cytoplasmic and membrane protein samples of 25 replicates were processed and digested simultaneously to minimise experimental variations between samples. Table 4.1 shows a summary of the subpopulations of $CD4^+$ T cells and their culture conditions which were prepared for proteomic profiling.

Table 4.1: A summary of details of the CD4+ T cells subsets that were processed for proteomic profiling. Samples were processed to separate cytosolic and membrane fractions.

$CD4^+$ subsets	Culture condition	No. of cells	No. of replicates
Naïve $CD4^+CD25^-$ $Foxp3^-$		$2x10^6$	5
Naïve $CD4^+CD25^+$ $Foxp3^+$ nTregs		$2x10^6$	5
Activated $CD4^+CD25^+$ $Foxp3^+$ nTregs	+ CD3/CD28 beads + IL 2 (30 U/mL)	$2x10^6$	5
Activated $CD4^+CD25^-$ $Foxp3^-$	+ CD3/CD28 beads + IL 2 (30 U/mL)	$2x10^6$	5
Activated $CD4^+CD25^{++}Foxp3^+$ iTregs	+ CD3/CD28 beads + IL 2 (30 U/mL) + TGF β1 (5 ng/mL)	$2x10^6$	5

4.2.2. Proteomic data generation, processing and identification of differentially expressed proteins

4 µL of each sample was injected onto an Eksigent ekpert nanoLC 400 system (Eksigent 425 pump and autosampler) via a trapping column (YMC Triart-C_{18} trap column 300 µm ID, 0.3 × 5 mm) using mobile Phase A; 0.1% formic acid, B; Acetonitrile with 0.1% formic acid; at 10 µL/min mobile phase A for 2 min before a valve switch and gradient elution onto the analytical column (YMC Triart C_{18} 150 × 0.3mm ID, 3 µm). The LC was hyphenated to a SCIEX TripleTOF 6600 mass spectrometer via the Duospray Source using a 50 µm electrode. The instrument was set in positive mode +5500V.

All samples were analysed both by SWATH (Data Independent Acquisition, to generate quantitative data) mode and by IDA (Information or Data Dependent Acquisition, DDA, for the generation of a list of identified proteins and peptides to create a spectral or ion library). The following linear gradients were used: for IDA, mobile phase B increasing from 2% to 30% over 68 min; 40% B at 72 min followed by column wash at 80% B and re-equilibration (87 min total run time). For SWATH, 3-30% B over 38 min; 40% B at 43 min followed by wash and re-equilibration as before (57 min total run time). IDA acquisition mode was used with a top 30 ion fragmentation (TOFMS *m/z* 400-1250; product ion 100-1500) followed by 15 sec exclusion using rolling collision energy, 50 ms accumulation time; 1.8 s cycle. SWATH acquisition used a 100 variable window method (previously optimised in the laboratory for complex protein sample types) using an accumulation time of 25 ms giving a 2.6 s total cycle time (over the range *m/z* 400-1250). IDA data was searched all together and by experimental group using ProteinPilot 5.0.2 (SCIEX) to generate a list of proteins and peptides for a local spectral library, with the following parameters; iodoacetamide alkylation, thorough search with emphasis on biological modifications (Swissprot human database Oct 2018).

4.2.2.1. Generation of a local spectral library for SWATH data analysis

The generated ".group" file from ProteinPilot search of the IDA runs contained all protein and peptide information from the samples. This data was imported into PeakView 2.2 (SCIEX) using the SWATH microapp plugin (SCIEX) and aligned to the

SWATH data files using a selection of endogenous peptides with high intensity and unambiguous retention time.

SWATH data was imported into the SCIEX OneOmics cloud processing software based on software developed from (Lambert, et al. 2013) and was extracted against a spectral library created from aligning both the locally generated library and a comprehensive mouse 3T3 cell external library (obtained from Dr Nick Morrice, SCIEX, UK) in OneOmics with the parameters 6 peptides per protein, 6 transitions per peptide, XIC width 30 ppm, 5 min retention time window. This method has been previously published by the proteomics facility in the John van Geest Cancer Research Centre (Aldiss, et al. 2019). See figure 4.6 for schematic overview of the proteomics data generation and processing.

4.2.3. Selection of novel membrane and cytoplasmic biomarkers of natural and induced Tregs using artificial neural networks (ANN)

SWATH-MS protein peak area (intensity) data of confident 99 membrane and 344 cytoplasmic proteins across the samples were subjected to ANN analysis, a robust bioinformatics tool (Abdel-Fatah, et al. 2016, Furini, et al. 2018, Wagner, et al. 2019), to statistically identify and select novel markers that can individually or in combination distinguish between nTregs and iTregs.

Herein, a bespoke 3 layered backpropagation ANN, coupled with a stepwise search and cross validation procedure was applied to evaluate protein peak area data for the confidently identified membrane and cytoplasmic proteins (Lancashire, et al. 2008a, Lancashire, et al. 2010).

Membrane and cytoplasmic protein data were run on ANN using two different stepwise analysis runs. First, multiple stepwise analysis which included 20 loops, each loop consisted of 10 steps. Second, single stepwise analysis which included 20 loops, each loop consisted of one step. The ANN analysis was carried out under the supervision of Professor Graham Ball using in house software developed for ANN (Stepwise OpenCL version 1.3.0). Data was exported from Excel as a .csv file in the appropriate format for importing into the Stepwise software for processing with the following parameters: 2 hidden nodes, momentum 0.5, 60%:20%:20% split for training:testing: validation, running time 50 and 3000 epochs.

After ANN analysis, the results of stepwise analysis of both runs were collected. For multistep analysis, a panel of markers with the lowest average test error (ATE) from each step of 20 loops were selected. Then markers were sorted based on the time of repetition in the 20 loops. For single stepwise analysis, as it included only one step for each loop, then only one marker with the lowest ATE was selected among other markers in each step by ANN. Then the resulting 20 markers were sorted based on the time of repetition in the 20 loops.

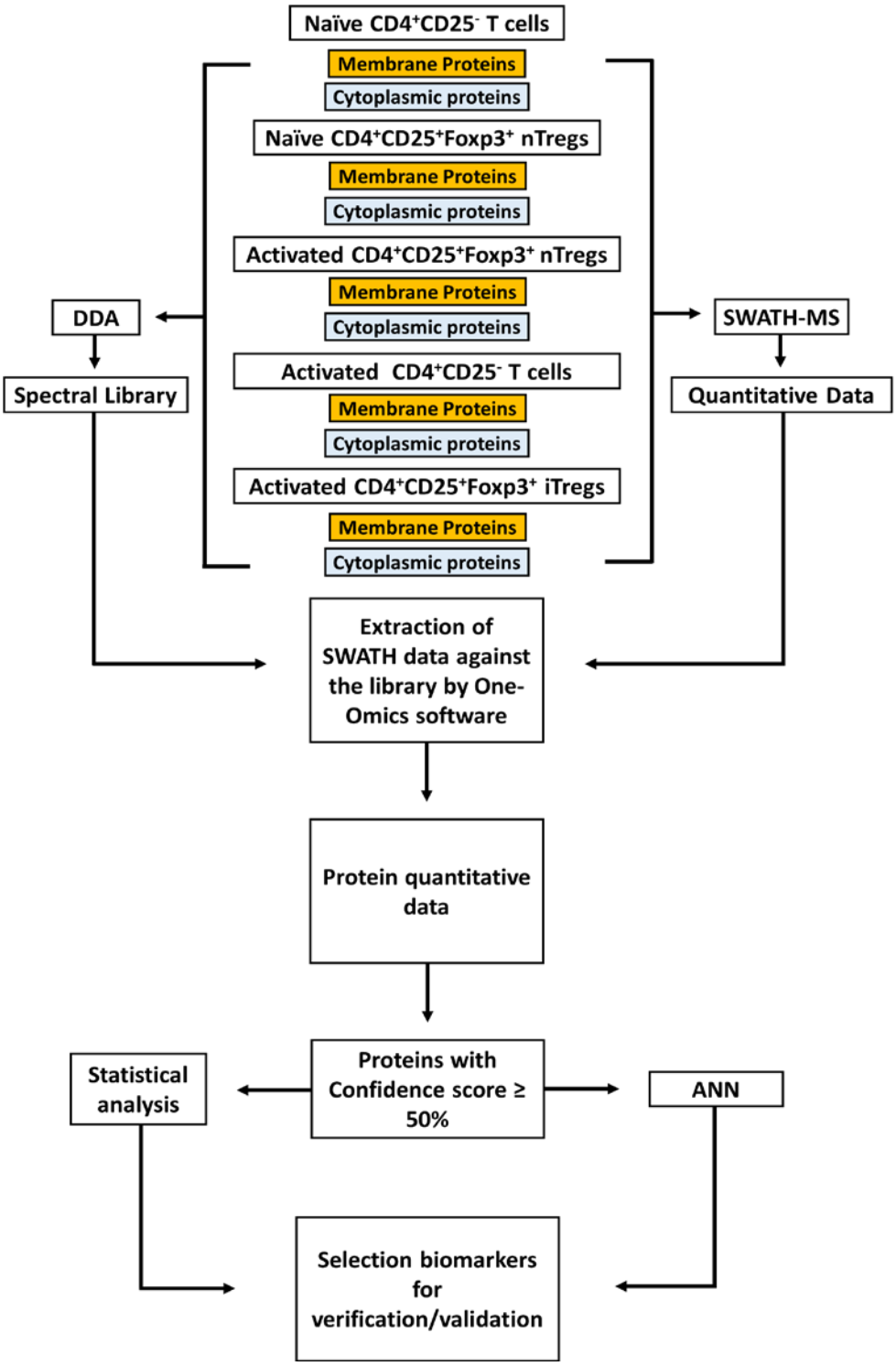


Figure 4.6. Schematic representation showing the overall method that was followed for identification, selection and verification of proteins.

4.2.4. Verification/validation of novel biomarkers of nTregs and iTregs using flow cytometry and in silico tools

4.2.4.1. Conjugated antibodies

Conjugated anti mouse monoclonal antibodies, including anti CD25, CD4, Foxp3, CPOX and MAVS antibodies were purchased for flow cytometry as shown in table 4.2.

Table 4.2: List of conjugated monoclonal antibodies that were used for flow cytometry analysis.

Monoclonal Abs	Fluorochrome	Clone	Company	Cat No.
CD25	PE	PC61.5	Thermo Fisher	12 0251 82
CD4	eFluor 450	GK1.5	ThermoFisher	48 0041 82
Foxp3	eFluor 660	FJK 16s	ThermoFisher	50 5773 82
MAVS	PE	E 6	Santa Cruz Biotechnology	sc 365334
CPOX (HEM6)	PE	B 9	Santa Cruz Biotechnology	sc 393388

4.2.4.2. Unconjugated antibodies

Conjugated antibodies against EPHX1 (HYEP), PLP2 and ITIH4 were not commercially available. Therefore, unconjugated polyclonal antibodies tested for immunofluorescence (IF) application specifically were purchased as IF antibodies are compatible to be conjugated for flow cytometry. Unconjugated monoclonal antibodies were not commercially available against those three proteins.

Table 4.3: List of unconjugated antibodies compatible for conjugation used for flow cytometry.

Polyclonal Abs	Concentration	Host/Species	Reactivity	Company	Cat No.
EPHX1 (HYEP)	1 mg/mL	Rabbit	Human Mouse Rat	Biorbyt	Orb48363
PLP2	0.5 mg/mL	Rabbit	Human Mouse Rat	Biorbyt	Orb312802
ITIH4	0.5 mg/mL	Rabbit	Human Mouse Rat	Biorbyt	Orb1128

4.2.4.3. Conjugation of unconjugated antibodies with Phycoerythrin (PE) fluorochrome

Lightning link® R Phycoerythrin conjugation kit (Expedeon, Cat No. 703 0010) was used to conjugate or label unconjugated antibodies with PE fluorochrome. The kit included vials of lightening link (LL) mix, LL modifier reagent, and LL quencher reagent. Each vial of LL mix was designed to conjugate 60 µL volume of an antibody when the concentration of antibody is 1 mg/mL to give a 1:1 (antibody:R PE) conjugation ratio. For setting up conjugation reaction, 6 µL of LL modifier reagent were added into 60 µL of antibody and mixed gently. Then the antibody sample with LL modifier was immediately added into the vial of LL mix which included lyophilised PE fluorochrome, and resuspended gently by pipetting the solution up and down twice. Then, the vial containing the conjugation reaction was incubated for 6 h at room temperature (20-25^{o}C) in the dark on standing position. After incubation, 6 µL of LL quencher reagent were added into the vial containing conjugation reaction and mixed very gently. Then the vial was incubated again for an hour at room temperature in the dark on standing position. Then PE conjugated antibody was ready to use and kept at 4^{o}C.

The concentration of PLP2 and ITIH4 antibodies was 0.5 mg/mL which was lower than the ratio limit of the conjugation kit. Also, the solution of both antibodies included higher concentration of BSA (0.5%) which was five time higher than the recommended concentration (<0.1%) for conjugation reaction. Therefore, antibodies were concentrated, and the BSA was removed from their solution using AbSelect™ BSA removal kit (Expedeon, Cat No. 820 0100). This kit included a vial of BSA removal buffer and a vial of re suspension buffer. Then BSA removal buffer was heated in warm water at 37^{o}C for 30 min to dissolve the contents. Once dissolved, the buffer was placed at room temperature ~20^{o}C for 10 min to prevent the formation of crystal. Then the buffer was centrifuged at 13000 x g for 1 min. After centrifuging, 160 µL of supernatants of BSA removal buffer were added into the solution of antibody (200 µL) and mixed gently. The solution then was incubated for 10 min at room temperature. After that, the solution was centrifuged at 13000 x g for 5 min. The supernatants were carefully discarded. The pellet of antibody was resuspended in 100 µL of re suspension buffer. After that, 60 µL of the concentrated

antibody were taken for conjugation with PE fluorochrome following the protocol mentioned above.

4.2.4.4. Pathway analysis of differentially expressed and identified proteins

Proteins uniquely expressed in each category of T cell population from the IDA runs were identified using the Venn diagram. The uniport IDs of the unique proteins were exported into MetaCore™ (Clarivate Analytics, accessed Oct 2019) for further analysis. MetaCore is a functional analysis integrated software suite for the analysis of high-throughput genomics, metabolomics and proteomics data using curated database of human and mouse gene and protein expression information from the peer reviewed publications. Significant pathway enrichments in each group was calculated by comparing the unique proteins identified in each group to the mouse database. Any pathways below a corrected p-value of 0.05 was considered as significant. An enrichment report was generated using each of the unique protein list and bar graph was generated in GraphPad prism. Similarly, Pathway enrichment analysis was performed using the differentially expressed proteins in iTregs compared to nTregs and nTreg depleted CD4 T cells along with its expression values. Significant pathways are plotted using a bar graph and important alterations in major pathways are highlighted as in the expression context.

4.2.4.5. Analysis of public datasets of markers of interest in cancer

To check the expression of markers on murine immune cells, each gene corresponding to the identified protein was surveyed using BloodSpot (http://servers.binf.ku.dk/bloodspot/ accessed Oct 2019), an online database of healthy and malignant haematopoiesis comprising large amounts of gene expression data of human and murine immune system. Although the markers were identified using a murine *in vitro* model, its translational validity in humans was explored using publicly available datasets. This study used OncoLnc (http://www.oncolnc.org/ accessed Oct 2019), a web-based tool to analyse The Cancer Genome Atlas (TCGA). OncoLnc designed to derive survival relationships of a query gene with gene expression data and survival information of 8,647 patients from 21 cancer types performed by TCGA. Genes were queried singly (FOXP3, ITIH4, PLP2, MAVS, HEM6 (CPOX), HYEP (EPHX1)) among 21 cancer types. Poor prognosis

using survival information of FOXP3 data among all the cancer types was considered as a surrogate marker for T reg enriched tumour. Kidney Renal Clear Cell Carcinoma (KIRC) was identified as the top cancer with significant association of FOXP3 expression and poor survival outcome. All the human equivalents of the shortlisted genes were subsequently queried in KIRC to see whether any of these genes explain similar survival outcome. The plots were generated using OncoLnc and the Significance of difference in survival probability between high and low group of each genes were calculated using Log rank p value, p-Value less than 0.05 was considered as significant.

4.3. Results

4.3.1. Subcellular proteomic profiling of mouse CD4+ T cells using using LC-MS-MS.

Membrane and cytoplasmic proteins derived from five different subsets of CD4+ T cells were profiled using LC-MSMS proteomics. The MS analysis was carried out in two different steps. First, a local library of proteins was created based on the total number of proteins identified from all subpopulations of CD4+ T cells using an information/data dependent acquisition approach (IDA/DDA). Second, samples were analysed by SWATH-MS, and the quantitative data extracted against this library combined with a comprehensive externally generated mouse library to give a list of proteins with relative quantitation across all samples.

4.3.1.1. Information Dependent Acquisition (IDA) proteomic analysis of sorted mouse CD4+ cells (qualitative analysis)

IDA acquisition was carried out on the 5 replicate samples in each sorted cell type for both membrane and cytosolic fractions and each set of biological replicates searched together. Proteins were identified and enumerated in Table 4.4.

Table 4.4: Number of proteins identified in each T cell subpopulation in both membrane and cytosolic fractions.

Subsets of CD4+ T cells	Membrane Fraction*	Cytoplasmic Fraction*	Total distinct protein groups
Naïve $CD4^+CD25^-$	656	677	**1096**
Naïve $CD4^+CD25^+Foxp3^+$ (nTregs)	431	310	**634**
Activated $CD4^+CD25^-$	621	582	**1000**
Activated $CD4^+CD25^+Foxp3^+$ (nTregs)	658	553	**991**
$CD4^+CD25^{++}Foxp3^+$ (iTregs)	611	681	**1090**

***Reversed sequences removed; results displayed at 1 % protein False Discovery Rate (FDR) cut off.**

Qualitative analysis of the data showed that the fractionation into membrane/cytosolic fraction was not absolute with many proteins present in both samples for each cell type (figure 4.7). The total cellular protein (membrane and cytosolic combined) is presented in figure 4.8 across all 5 sorted cell types.

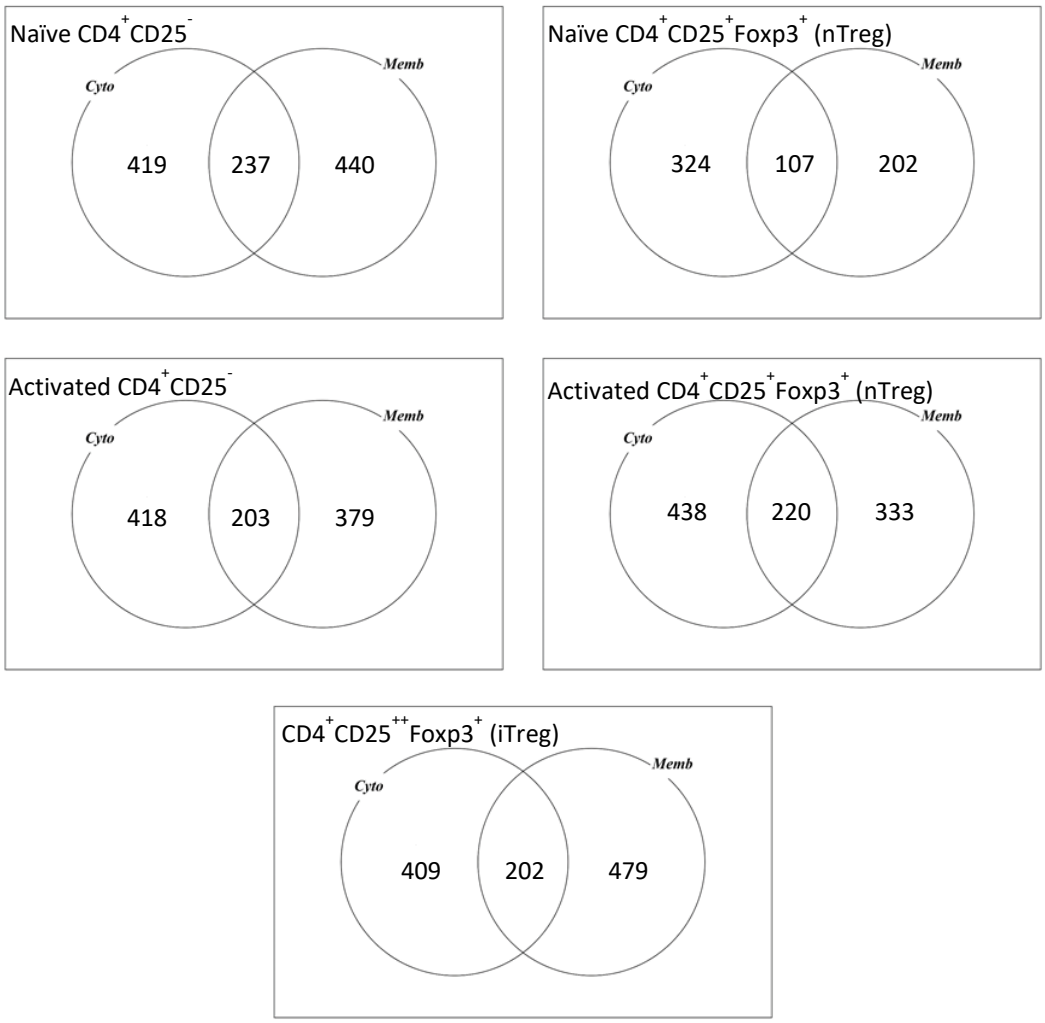


Figure 4.7. Venn diagrams showing the overlap of proteins identified in membrane versus cytosolic fractions by IDA acquisition and protein search in each group of 5 replicate samples from the sorted cell types.

Protein data was also analysed for each sorted cell type by cross referencing the proteins for each cell type with the known protein localisations from the literature based on the Uniprot database (http://www. Uniprot.org, accessed April 2019) this is shown in table 4.5. The IDs of shared, unique and differentially expressed proteins of the five CD4+ T cell subsets were mapped using MetaCore software to identify

significantly enriched pathways based on p value (< 0.05) and FDR (<0.05) (figure 4.9 and 4.10).

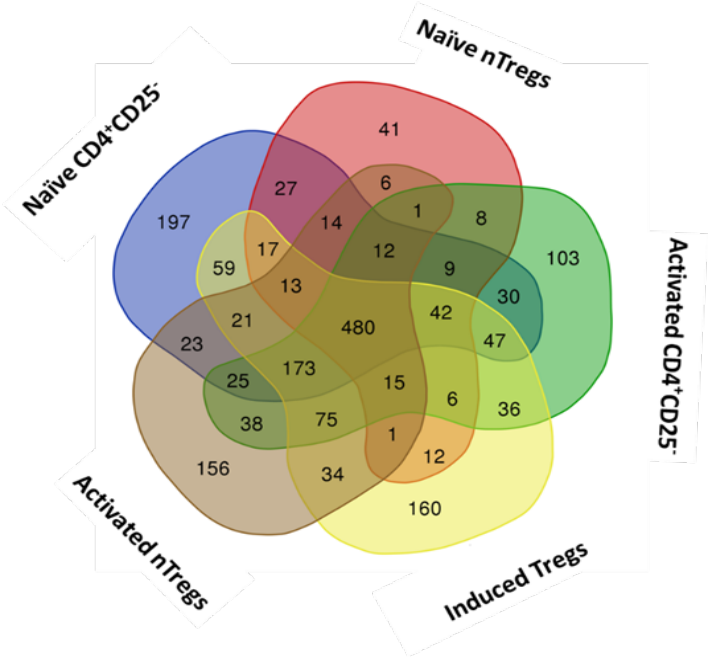


Figure 4.8. Venn diagram comparing the total number of proteins identified in five different subsets of CD4⁺ T cells.

Table 4.5: Characterisation of the known subcellular location (as defined by Uniprot) of the identified proteins from the sorted CD4⁺ cells from the qualitative IDA analysis.

Subsets of CD4⁺ T cells	Known Membrane Location	Known Cytoplasmic Location	Proteins with unknown location
Naïve CD4⁺CD25⁻	466	708	15
Naïve CD4⁺CD25⁺Foxp3⁺ (nTregs)	300	404	0
Activated CD4⁺CD25⁻	409	687	4
Activated CD4⁺CD25⁺Foxp3⁺ (nTregs)	438	643	6
CD4⁺CD25⁺⁺Foxp3⁺ (iTregs)	437	745	9

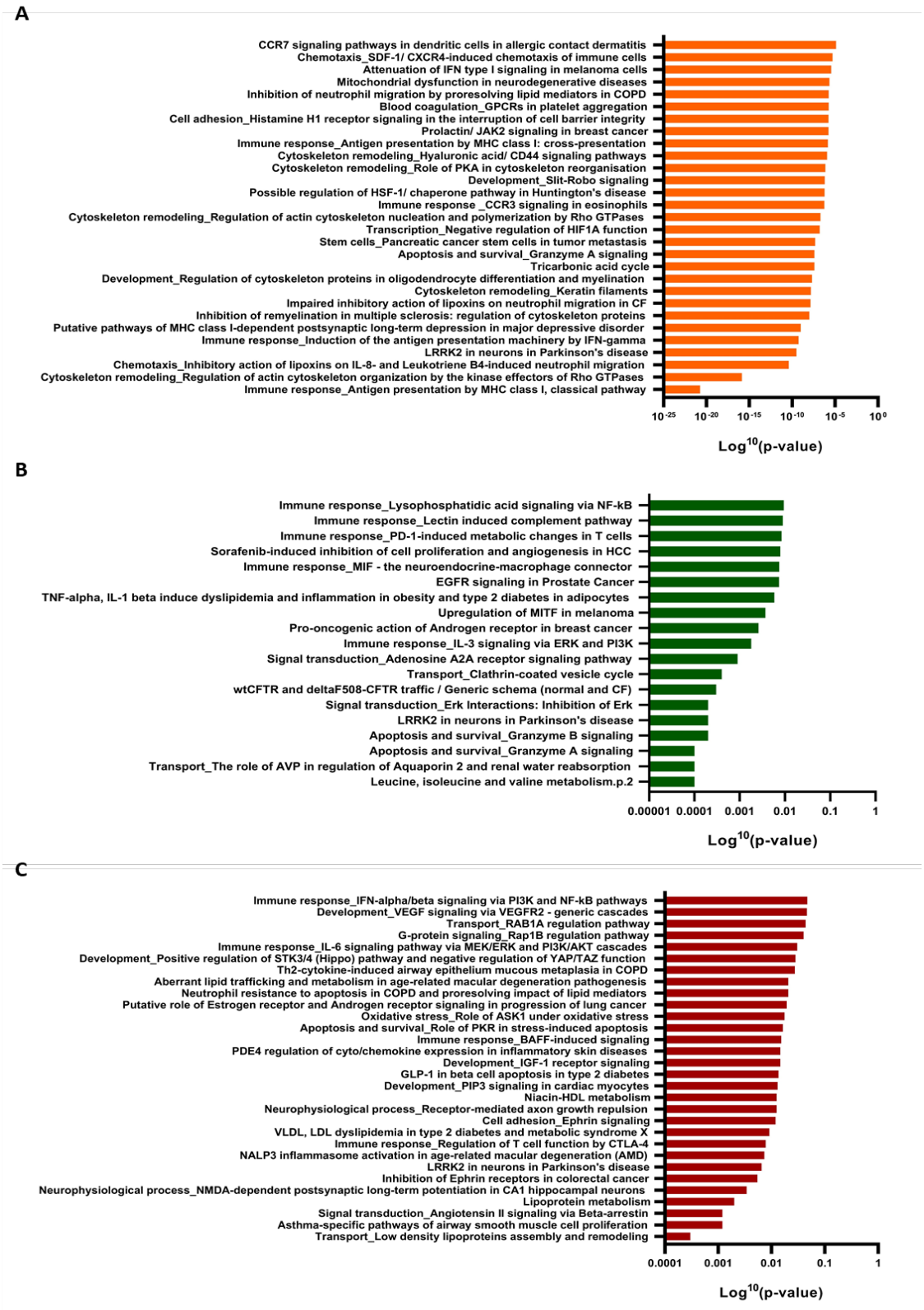


Figure 4.9. Significantly enriched pathways derived through mapping the IDs of proteins identified in the DDA-library using Metacore analytics based on p-value < 0.05 and FDR < 0.05. **(A)** Significantly enriched pathways that are shared between five different subsets of $CD4^+$ T cells. **(B)** Significantly enriched pathways derived from the IDs of the unique proteins of naïve $CD4^+CD25^-$ T cells. **(C)** Significantly enriched pathways derived from the IDs of the unique proteins of activated $CD4^+CD25^-$ T cells.

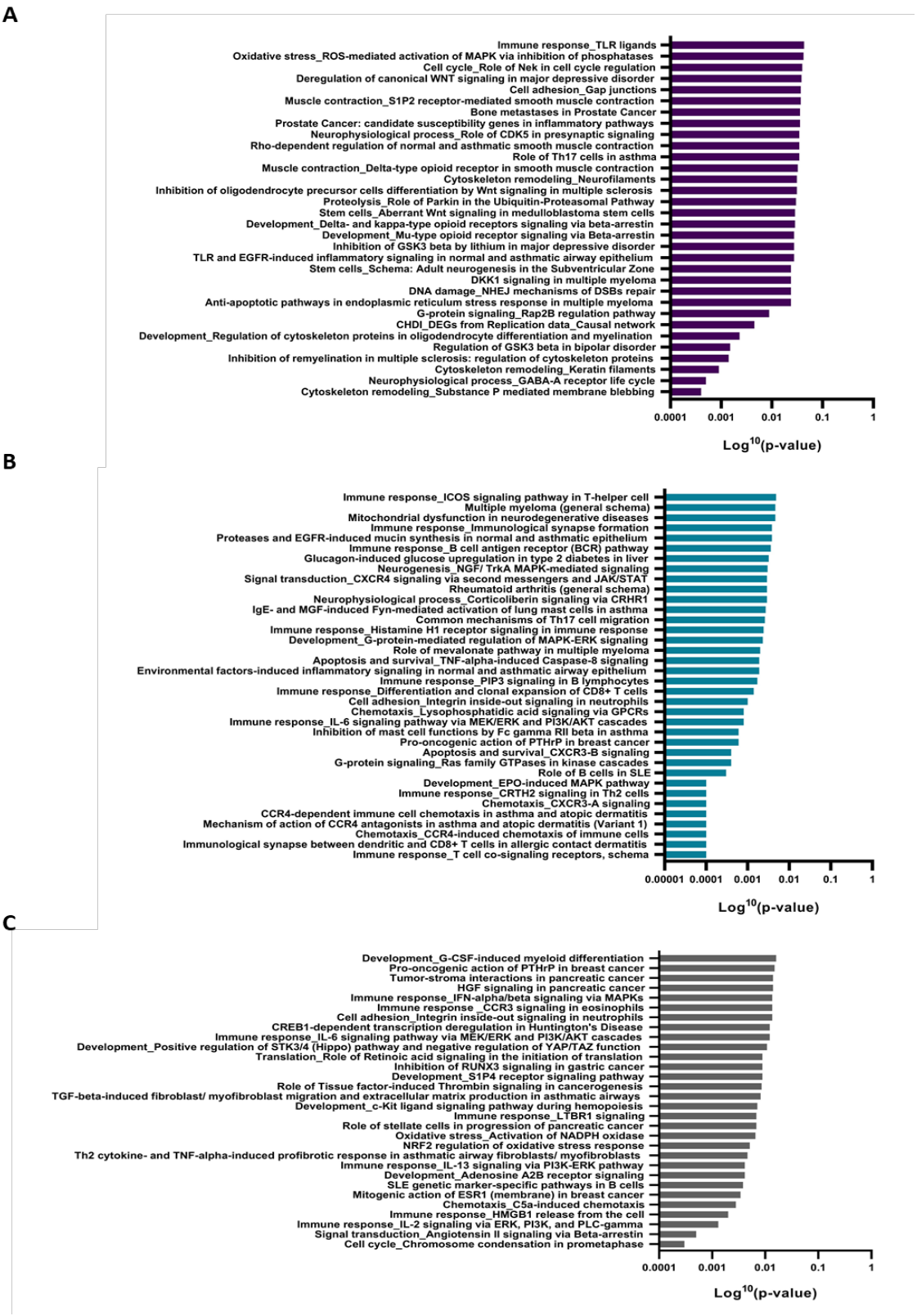


Figure 4.10. Significantly enriched pathways derived through mapping the IDs of proteins identified in the DDA-library using Metacore analytics based on p-value < 0.05 and FDR < 0.05. **(A)** Significantly enriched pathways derived from the IDs of the unique proteins of naïve $CD4^+CD25^+$ nTregs. **(B)** Significantly enriched pathways derived from the IDs of the unique proteins of activated $CD4^+CD25^+$ nTregs. **(C)** Significantly enriched pathways derived from the IDs of the unique proteins of iTregs.

4.3.1.2. Generation of a local spectral library (ion library) for SWATH MS profiling

A local (generated from the real samples) spectral library was generated from searching the mass spectrometer acquisition file of all fractions of all sorted cells together into a single result (.group) file (Table 4.6). 1414 proteins from 5441 peptides were identified from the Swissprot Human database in all samples.

Table 4.6 Summary of proteins, peptides and spectra identified in all the cell fractions and sorted cell types analysed by IDA mass spectrometry and searched in ProteinPilot.

Summary of Identification Yields

Identification Yield at FDR Threshold				
Data Level	***FDR Type***	**FDR**	**ID Yield**
Protein	***Local***	*1%*	*1056*
		5%	*1429*
		10%	*1546*
	Global	**1%**	**1414**
		5%	*1846*
		10%	*2050*
Distinct peptide	***Local***	*1%*	*3753*
		5%	*5425*
		10%	*6169*
	Global	**1%**	**5441**
		5%	*7657*
		10%	*9145*
Spectral	***Local***	*1%*	*41703*
		5%	*57653*
		10%	*69601*
	Global	**1%**	**59683**
		5%	*90162*
		10%	*115479*

*Global FDR at 1% values accepted.

The .group results file was then converted to the spectral library format (.txt) as previously described (section 4.2.2.1) after removal of all proteins with shared peptides.

4.3.1.3. Improvement of the spectral library (ion library)

While reasonably good, 1414 proteins and 5441 peptides in a spectral library will allow fewer than 1414 proteins to be quantified from SWATH data. In order to improve the quantitation an external library was obtained from Dr Nick Morrice

(SCIEX) containing a large pre-generated (from a TripleTOF 6600 instrument) mouse 3T3 cell spectral library in .txt format. This library contained 8,446 proteins and it was reasonably speculated that many of these mouse proteins would also be present in the cell lysate protein fractions from this study.

4.3.1.4. SWATH-MS quantitative profiling

As the subcellular profiling was carried out based on the number of cells ($2x10^6$) in each replicate of the five subsets of CD4+ T cells, the reproducibility in SWATH-MS profiling data of the five subsets of CD4+ T cells was firstly assessed based on the percentage of reproducible transitions of peptides detected in each sample. The results showed that the percentage of reproducible transitions of peptides was in range between 33% and 55% among 50 samples that were analysed on MS (figure 4.6). The variability of SWATH MS data was relatively low in the most of samples (40/50) as the percentage of reproducible transitions of peptides to fragment ions was in range of 44 -55%, whereas only 10 samples showed higher variability with 33-43% of reproducible transitions of peptides (figure 4.11). A high % of reproducible transitions would not be expected here as the cells have been fractionated into different sub-types likely with significant differences in protein profiles, what matters is the variation between all samples. While there are a few samples that vary more, having a reduced % reproducible transitions, on discussion

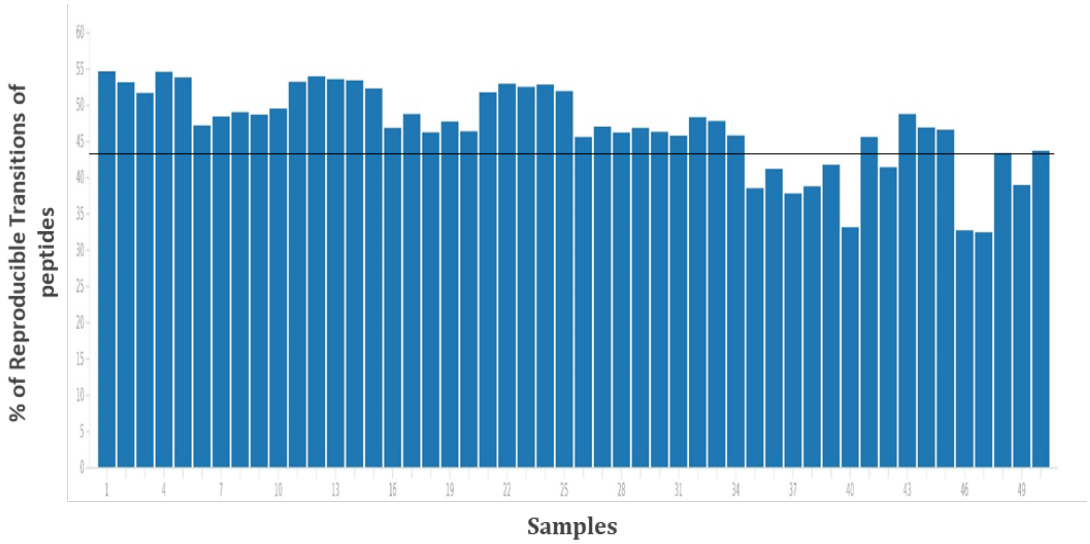


Figure 4.11. The percentage of reproducible transitions of peptides in each of 50 samples of the five different subsets of $CD4^+$ T cells that were prepared for SWATH-MS profiling. Cytoplasmic and membrane proteins samples derived from each replicate of the five subsets of $CD4^+$ T cells. Five replicates for each subset were prepared.

with the OneOmics software application specialists at SCIEX, the advice was to keep all data in as the algorithm would take this into account when processing.

After assessing the variability, the SWATH-MS data were then normalised to minimise the differences due to experimental artefacts and protein/peptide load between samples and improve the data quality using SCIEX OneOmics software. The data normalisation was performed internally in the software suite using MLR (Most Likely Ratio) normalisation strategy which has been found to be the best strategy for SWATH-MS data normalisation (Lambert, et al. 2013). Figure 4.12 shows the $\log_{10}$ transformed intensity ratios of peaks derived from peptide fragments within each sample before and after normalisation.

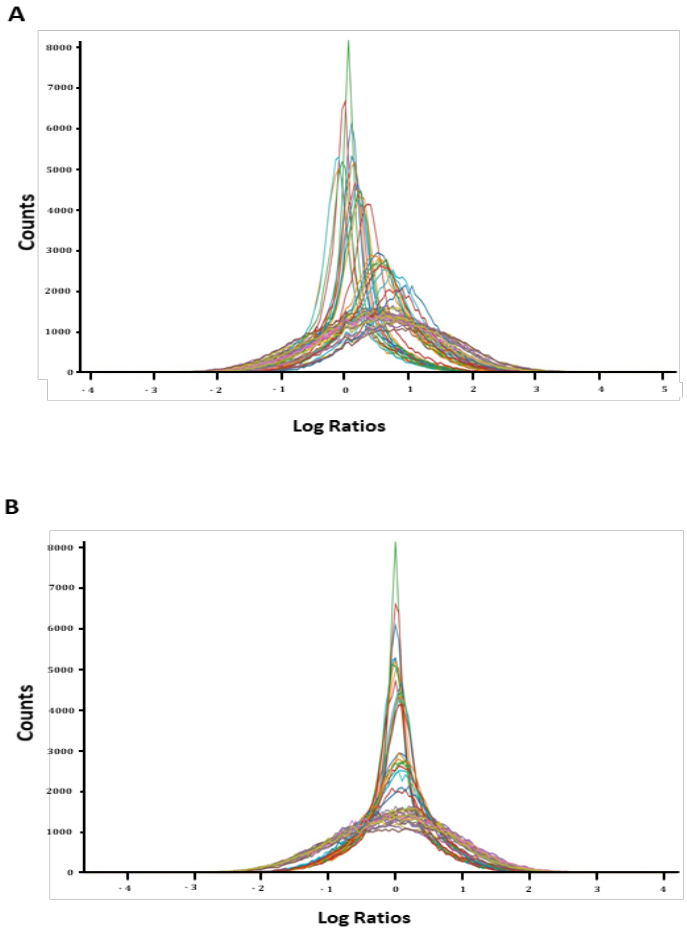


Figure 4.12. The $\log_{10}$ intensity peaks of peptide fragments of samples before and after normalisation. **(A)** Pre-normalisation. **(B)** Post-normalisation.

4.3.1.5. SWATH-MS quantitative profiling

Using the combined locally generated and external mouse 3T3 cell spectral library, a total of 3910 proteins were detected and selected for quantitation by SWATH MS analysis based on their expression across all five subsets of $CD4^+$ T cells. The expression (peak area, intensity) of each protein was compared between the five subsets of $CD4^+$ T cells to identify up and down regulated proteins for each subset. Using the Uniprot database and website, among the 3910 proteins, 1236 (31.6%) proteins were identified as membrane proteins, whereas 2294 (58.7%) proteins were classified as cytoplasmic and nuclear proteins. The remaining 380 (9.7%) proteins were identified with unknown (not mapped) subcellular locations (figure 4.13).

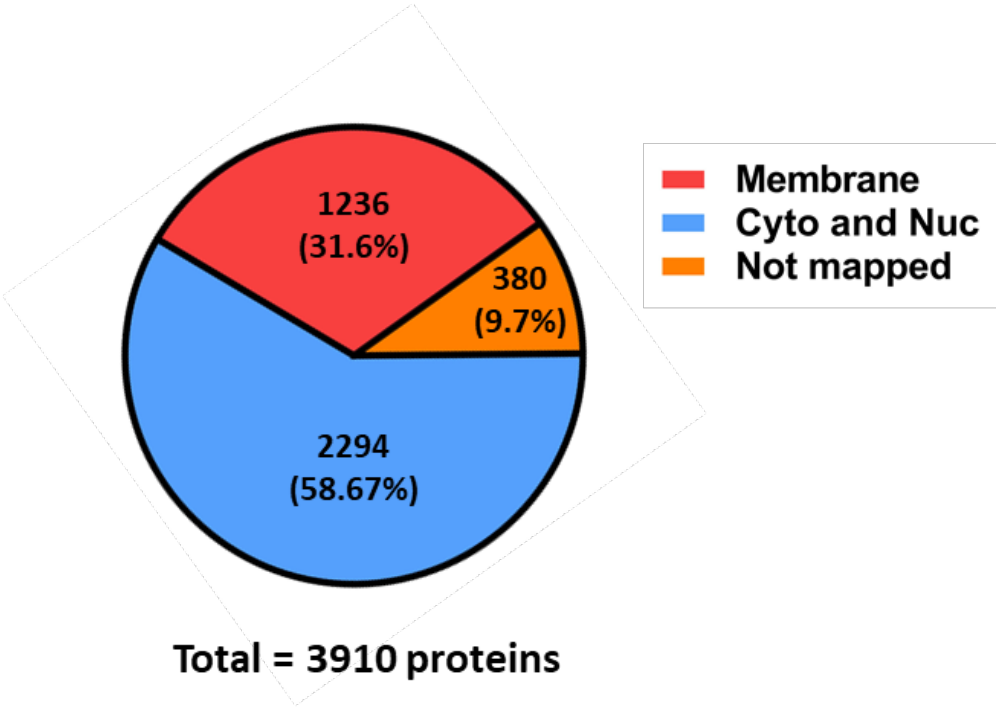


Figure 4.13. Pie chart showing the number and percentage of SWATH-MS-identified proteins based on their subcellular locations. Cyto: Cytoplasmic, Nuc: Nuclear.

To assess the reproducibility of SWATH MS profiling normalised and log transformed peak areas (intensity) of 3910 proteins were subjected to a principle component analysis (PCA). Most of the variation in the data was captured using the first two components (PC1 and PC2). For membrane proteins, the percentage of variability in the data was 45% and 9.4% at PC1 and PC2, respectively (figure 4.14 A). For cytoplasmic proteins, the percentage of variability in the data was 50.1% and 7.4% at PC1 and PC2, respectively (figure 4.14 B). There was minimum intra replicate variability observed within each group indicating good reproducibility between the samples, while five biologically different groups were well separated. Moreover, the analysis showed that groups representing naïve subsets of $CD4^+$ T cells were well separated from groups which represented activated subsets of $CD4^+$ T cells confirming the trend of reproducibility between the samples of both membrane and cytoplasmic proteins. The results also showed that there was an increase in the variability within activated groups, which allowed the group of iTregs to be nicely separated from other activated groups in both membrane and cytoplasmic protein samples.

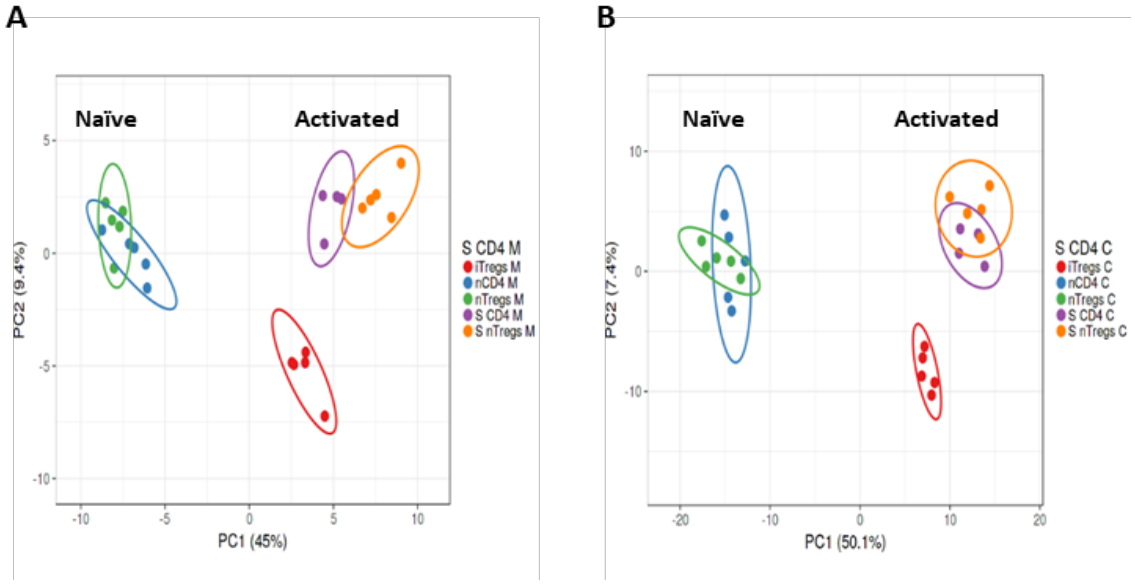


Figure 4.14. PCA analysis showing the percentage of variability between five different subsets of $CD4^+$ T cells at the first two principle components (PC1 and PC2). **(A)** PCA analysis for the samples of membrane proteins. **(B)** PCA analysis for the samples of cytoplasmic proteins. Group highlighted with green (naïve $CD4^+CD25^-$ T cells), blue (naïve $CD4^+CD25^+$ nTregs), violet (stimulated/activated $CD4^+CD25^-$ T cells), orange (stimulated/activated $CD4^+CD25^+$ nTregs) and red (iTregs).

The SWATH MS identified proteins were then sorted based on their confidence score of identification from OneOmics (this confidence score is based on a proprietary algorithm from SCIEX and has no equivalent in any classical statistical test). Proteins with confidence score ≥ 50% were classified as "possible proteins with real expression changes between experimental groups" and selected for further analysis. The total number of confident proteins was 443 (11.3%) from the total 3910 identified proteins. Of 443 proteins, 99 (22.3%) proteins were classified as membrane proteins, whereas the remaining 344 (77.7%) proteins were classified as cytoplasmic and nuclear proteins (figure 4.15).

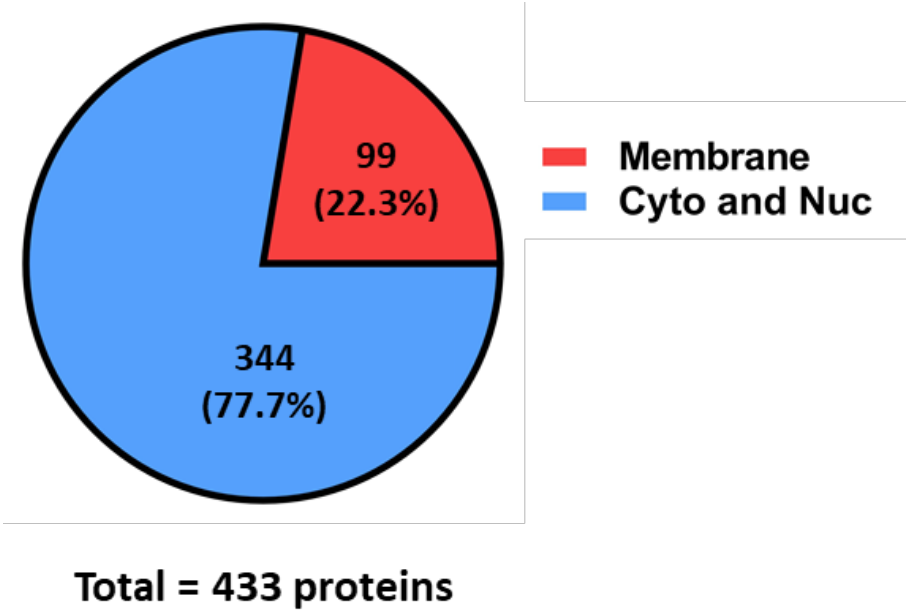


Figure 4.15. Pie chart showing the number and percentage of SWATH-MS-identified "confident proteins" based on their subcellular locations.

The SWATH quantitative data (peak area/intensity) of confidently differentially expressed proteins were analysed based on two different approaches: OneOmics-based approach and ANN-based approach.

4.3.1.6. OneOmics-based approach

The OneOmics-based results showed the fold change and confidence score of identified proteins based on a pairwise comparison of the 5 different

subpopulations of $CD4^+$ T cells against each other. The overall results of OneOmics algorithm can be found in the appendix (tables 9A, 10A, 11A, 12A).

The expression of identified proteins was compared between the population of iTregs and naïve nTregs as a control population to identify distinct proteins that are exclusively upregulated in iTregs. The results showed that 239 membrane and cytoplasmic proteins with confidence score ≥ 50% were found highly expressed in iTregs compared to nTregs at fold-change 1.5 cut-off value (results shown in the appendix, table 10A). Table 4.7 shows a list of the top 32 of 239 membrane and cytoplasmic proteins that are differentially expressed in iTregs compared to naïve nTregs.

Table 4.7. List of the top 32 cytoplasmic and membrane proteins that are differentially expressed in iTregs compared to naive nTregs.

Protein ID	Protein Name	Fold-change	OneOmics Confidence score*
Q8BHX1	HAUS1	6.0	71%
A2A4U6	ZFP334	5.0	55%
Q80XI6	M3K11	4.7	57%
Q3KNJ2	NHEJ1	4.5	55%
E9Q1M6	ANKHD1	3.9	75%
Q3UMA3	HGFS	3.9	57%
A0A0R4J0F6	GAK	3.8	56%
Q8BMC4	NOP9	3.8	59%
Q9QYG0	NDRG2	3.8	64%
Q8BV13	CSN7B	3.7	61%
Q9R1Q7	PLP2	3.7	86%
D3YXP6	PMVK	3.6	50%
Q811B1	XYLT1	3.5	72%
Q8VEJ9	VPS4A	3.5	53%
Q9D1C1	UBE2C	3.4	60%
Q0HA38	TT21B	3.4	64%
A2A4U6	Zfp334	3.3	57%
Q8BXQ2	PIGT	3.3	63%
Q9CX56	PSMD8	3.2	83%
Q920E5	FPPS	3.2	78%
E9PWW6	Zc3h7a	3.2	60%
E9PVX6	KI67	3.2	77%
K3W4Q8	Basigin	3.2	78%
Q8BKX1	BAIP2	3.2	69%
P17751	TPIS	3.2	87%
P57787	MOT4	3.1	55%
Q80VJ2	SRA1	3.1	53%
E9PX68	Slc4a1ap	3.1	57%
P70333	HNRH2	3.1	51%
Q99N92	RM27	3.1	51%
Q05920	PYC	3.1	80%
Q8VCF0	MAVS	3.1	55%

* Confidence score is based on Sciex OneOmics proprietary algorithm. <50% unlikely to be a real fold change.

The results also showed that 79 membrane and cytoplasmic proteins with confidence score ≥ 50% were found upregulated in naïve nTregs compared to iTregs at 1.5-fold-change cut-off value (data shown in the appendix, table 10A). Table 4.8 shows a list of the top 32 among 79 cytoplasmic and membrane proteins that are found to be upregulated in naïve nTregs compared to the population of iTregs.

Table 4.8. List of the top 32 cytoplasmic and membrane proteins that are differentially expressed in nTregs compared to iTregs.

Uniprot protein ID	Protein Name	Fold-change	OneOmics Confidence score
Q3KNJ2	NHEJ1	4.5	59%
Q792Z1	Q792Z1	4.5	77%
P58501	PAXB1	4.2	61%
Q8VE19	MIO	4.0	76%
Q8BGS2	BOLA2	4.0	61%
Q8BHS6	ARMX3	3.9	57%
Q64525	H2B2B	3.8	57%
Q8R422	CD109	3.5	83%
B0R091	B0R091	3.5	56%
E9QA45	E9QA45	3.5	89%
Q8BR76	MKS3	3.5	57%
Q920Q2	REV1	3.5	57%
Q8CC88	VWA8	3.4	51%
O89079	COPE	3.4	63%
Q9CZP3	Q9CZP3	3.4	65%
Q9D379	HYEP	3.3	76%
Q0PD20	Q0PD20	3.3	56%
Q8K2M0	RM38	3.3	52%
A0A0R4J0J3	A0A0R4J0J3	3.3	50%
O35609	SCAM3	3.3	50%
P16330	CN37	3.1	53%
Q9CRD2	EMC2	3.0	51%
Q9JIZ9	PLS3	2.9	57%
Q64327	MEA1	2.9	51%
Q9Z1R9	Q9Z1R9	2.9	85%
Q9R207	NBN	2.9	50%
O35598	ADA10	2.8	62%
Q3UD82	PARP8	2.8	53%
E9Q0G1	E9Q0G1	2.8	55%
A2A513	A2A513	2.7	91%
E9Q0A7	E9Q0A7	2.6	50%
Q6NS45	CCD66	2.6	53%

The expression of identified cytoplasmic and membrane proteins was also compared between the population of iTregs and stimulated $CD4^{+}CD25^{-}$ T cells as a second control population. The results showed that 56 cytoplasmic and membrane proteins were found differentially expressed in iTregs compared to stimulated $CD4^{+}CD25^{-}$ T cells population at 1.5-fold-change cut-off value (data shown in the

appendix, table 9A). Table 4.9 shows the top 32 among 56 cytoplasmic and membrane proteins that are upregulated in iTregs compared to stimulated $CD4^{+}CD25^{-}$ T cells.

Table 4.9. List of the top 32 cytoplasmic and membrane proteins that are differentially expressed in iTregs compared to activated $CD4^{+}CD25^{-}$ T cells.

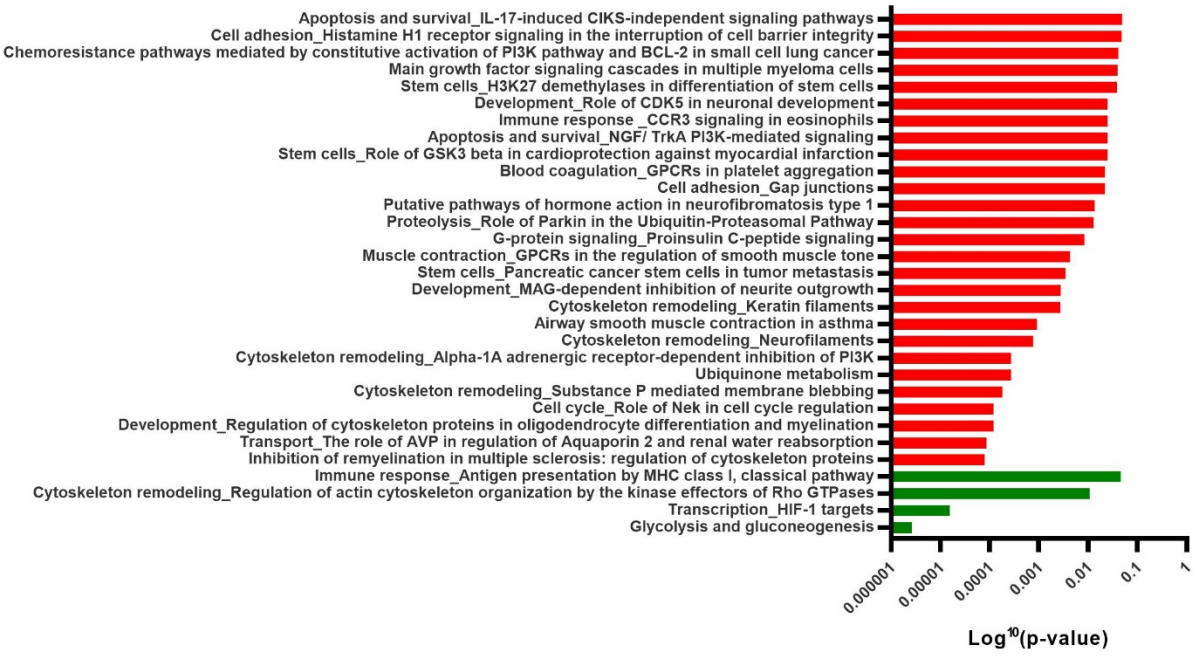
Uniprot Protein ID	Protein Name	Fold-change	OneOmics Confidence score
A2AH22	AMRA1	5.2	66%
A0A0R4J0F6	GAK	4.7	53%
Q61545	EWS	4.4	64%
A0A0R4J187	XRCC6	4.0	51%
Q3UMA3	HGFRTKS	4.0	57%
Q80TA6	MTMRC	3.7	52%
A6PW84	P3H1	3.7	57%
Q80XJ3	TTC28	3.3	59%
P14094	AT1B1	3.3	74%
P02798	MT2	3.2	51%
Q8R2Y8	PTH2	3.2	60%
D5MCW4	CUTA	3.2	63%
Q8VDD8	WASH1	3.2	59%
Q91W63	NRK1	3.2	65%
Q3UL36	ARGL1	3.0	59%
Q8C522	ENDD1	2.9	65%
Q9ERE7	MESD	2.8	58%
A2AQ07	TBB1	2.8	52%
Q80U19	DAAM2	2.8	56%
Q9EQ08	HNS	2.7	52%
Q8BX09	RBBP5	2.7	62%
Q61127	NAB2	2.6	52%
O35215	DOPD	2.6	64%
P97864	CASP7	2.5	69%
Q80XI6	M3K11	2.4	77%
Q8BVG4	DPP9	2.4	51%
P84078	ARF1	2.3	50%
Q8BI72	CARF	2.3	54%
E9Q6R7	UTRN	2.2	66%
P02089	HBB2	2.2	82%
Q9CPX4	Ferritin	2.2	59%
Q8CAK1	CAF17	2.1	51%

The results also revealed that 98 cytoplasmic and membrane proteins were found expressed at high levels in stimulated $CD4^{+}CD25^{-}$ T cells compared to iTregs at 1.5-fold-change cut-off value (data shown in the appendix, 9A). Table 4.10 shows the top 32 of 98 cytoplasmic and membrane proteins that are upregulated in stimulated $CD4^{+}CD25^{-}$ T cells compared to iTregs population.

Table 4.10. List of the top 32 cytoplasmic and membrane proteins that are differentially expressed in activated $CD4^{+}CD25^{-}$ T cells compared to iTregs.

Uniprot Protein ID	Protein Name	Fold-change	OneOmics Confidence score
Q80UF4	SDCG8	6.643856	0.758175
Q9D0S9	HINT2	6.485372	0.547028
Q45VK7	DYHC2	5.343186	0.615735
Q9CX56	PSMD8	4.357873	0.663257
Q3KNJ2	NHEJ1	4.157255	0.600902
A0A0R4J0J3	ADNHQ4	4.099722	0.582634
Q9JLV1	BAG3	3.828483	0.615726
P21981	TGM2	3.769342	0.842697
G3X920	ARMC8	3.663181	0.729015
Q792Z1	MCG140784	3.621907	0.623682
Q8BYH8	CHD9	3.614892	0.586976
B1ASU9	TLK2	3.478489	0.627116
P43277	H13	3.454123	0.701737
Q923D4	SF3B5	3.391455	0.626311
A2APB8	TPX2	3.378938	0.741857
Q8CAH8	GDPK	3.357312	0.590024
P27661	H2AX	3.303716	0.604916
Q91YR5	MET13	3.292886	0.551793
Q9D824	FIP1	3.283845	0.53197
Q8BH15	CNO10	3.249924	0.537365
Q80UF4	SDCG8	3.197587	0.64159
Q8BGS2	BOLA2	3.192564	0.828886
P40336	VP26A	3.137329	0.591784
P32037	GTR3	3.117137	0.87764
B1AW21	Myotubularin	3.091522	0.554631
Q91ZU1	ASB6	3.076995	0.631676
P07356	ANXA2	3.070642	0.753384
E9Q0G1	DRP	3.053311	0.600026
Q8CGP7	H2A1K	3.02566	0.597722
O35409	FOLH1	3.012569	0.881223
Q64735	CR1L	3.004373	0.619048
Q9ERH4	NUSAP	2.915646	0.635705

The differentially expressed proteins were further analysed using MetaCore pathway enrichment tool. The most differentially regulated pathways in iTregs in comparison to nTregs and cultured CD25 depleted CD4 T Cells are depicted in figure 4.16. the significant pathways altered in each of the comparisons are represented in a pathway context using MetaCore (Figure 4.17-4.19).



B

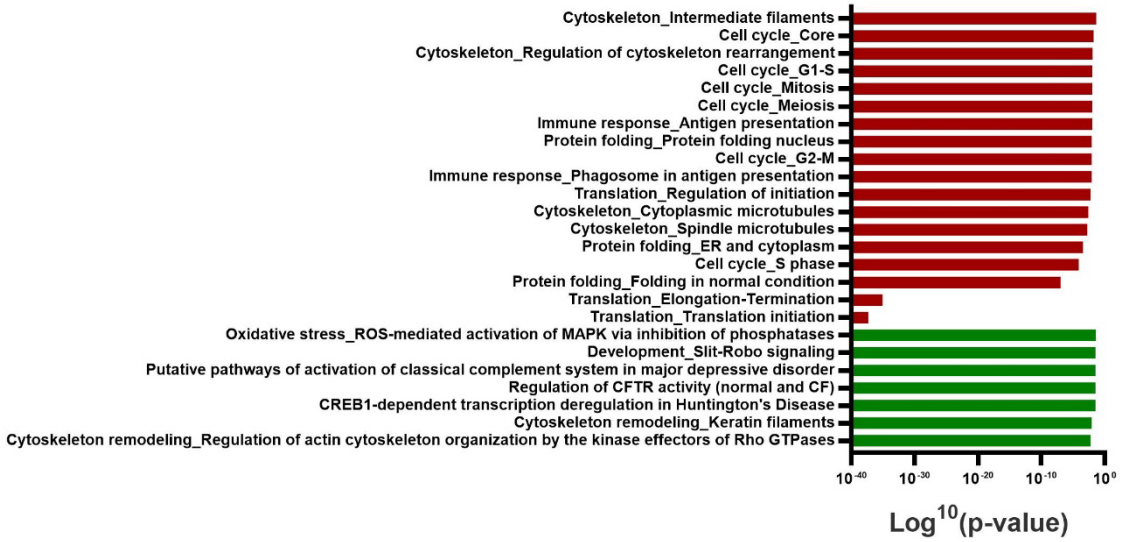


Figure 4.16. Significantly enriched pathways in iTregs compared to stimulated CD4 ($CD25^{-}$ $FOXP3^{-}$) and nTregs. **(A).** Bar graph showing significance of enrichment of pathways in iTregs compared to CD4 (CD25- FOXP3-) T cells cultured in the same way as iTregs for 5 days but without TGFβ. Y axis describe the pathways and the X axis represent corrected p value in log^{10} scale. Red colour indicating the upregulated and the green indicating down regulated pathways in iTregs. **(B).** Bar graph showing significance of enrichment of pathways in iTregs compared to pure population of nTregs ($CD25^{+}$ $FOXP3^{+}$). Y axis describe the pathways and the X-axis represent corrected p value in log^{10} scale. Dark Red colour indicating the upregulated and the green indicating down regulated pathways in iTregs comparison to nTregs.

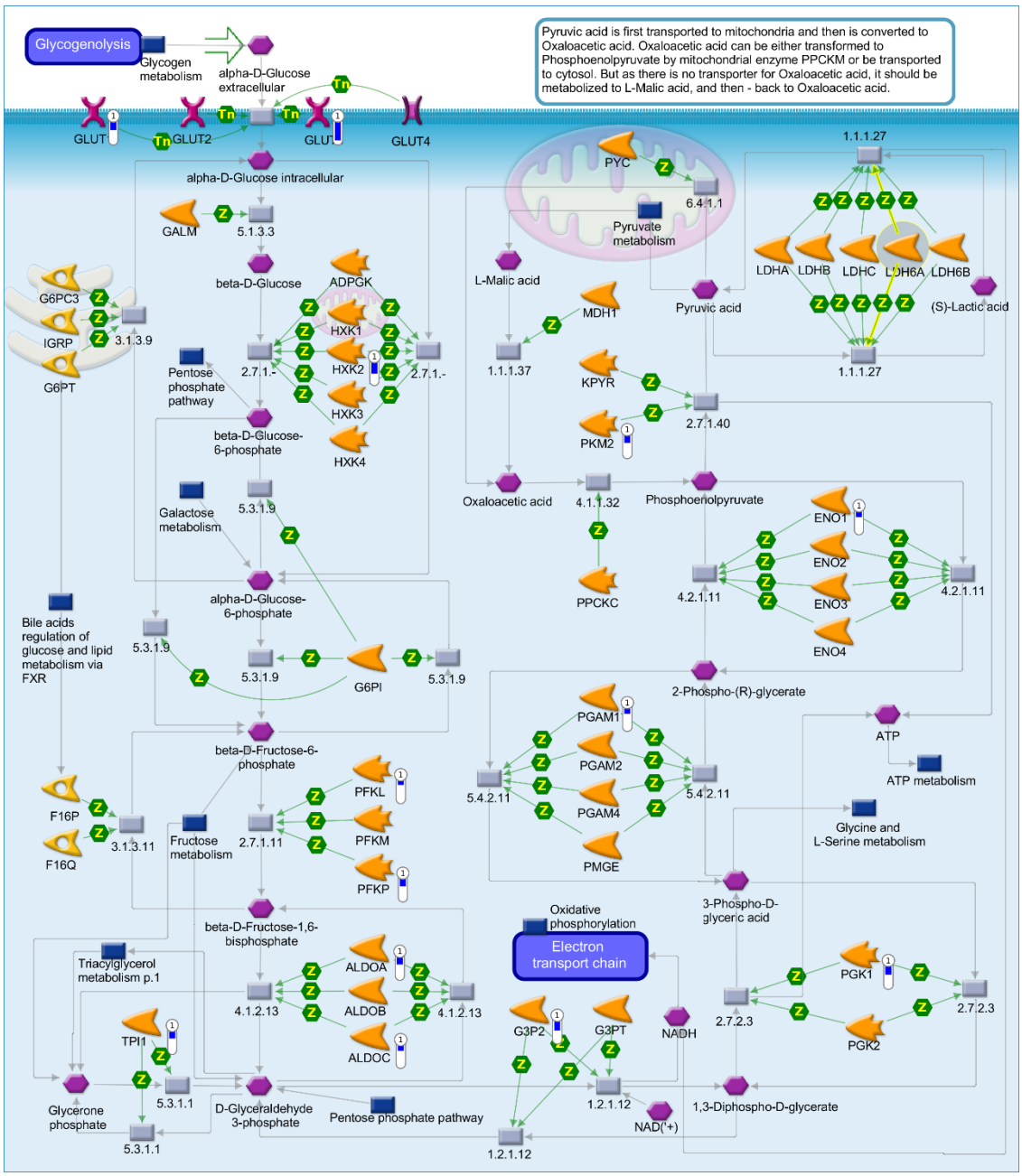


Figure 4.17. Pathway maps of glycolysis and gluconeogenesis with significant downregulated proteins were highlighted using a thermometer icon. Differentially expressed proteins shortlisted by pairwise comparison of quantitative mass spectrometry data generated from iTregs and cultured nTreg depleted CD4 T-cells were analysed using MetaCore functional enrichment tool. Glycolysis and gluconeogenesis are one of the most significantly down pathways in iTregs compared to CD4 T cells, significant proteins upregulated in iTregs compared to CD4 T cells are overlaid in the whole pathway context.

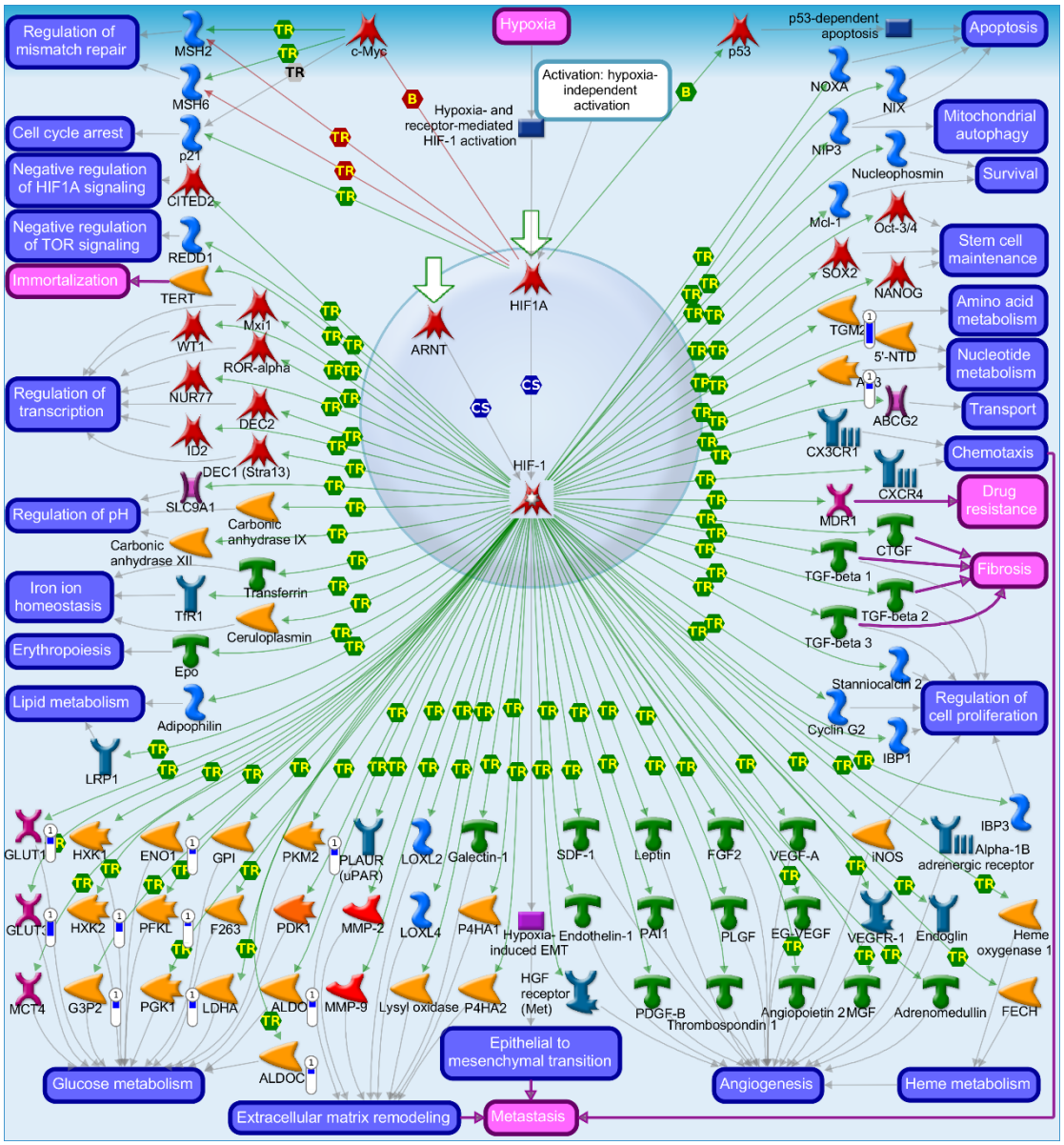


Figure 4.18. Several HIF-1 targets were found to be down regulated in iTregs in comparison to cultured nTreg depleted CD4 T-cells. Pathway maps of HIF-1 targets proteins altered in iTregs compared to nTregs are highlighted using thermometer signs (red colour). Differentially expressed proteins shortlisted by pairwise comparison of quantitative mass spectrometry data generated from iTregs and cultured CD4 T-cells were analysed using MetaCore functional enrichment tool. Transcription of HIF-1 targets are one of the most significantly downregulated pathways in iTregs compared to CD4 T- cells. Significant proteins upregulated in iTregs compared to CD4 T cells are overlaid in the whole pathway context.

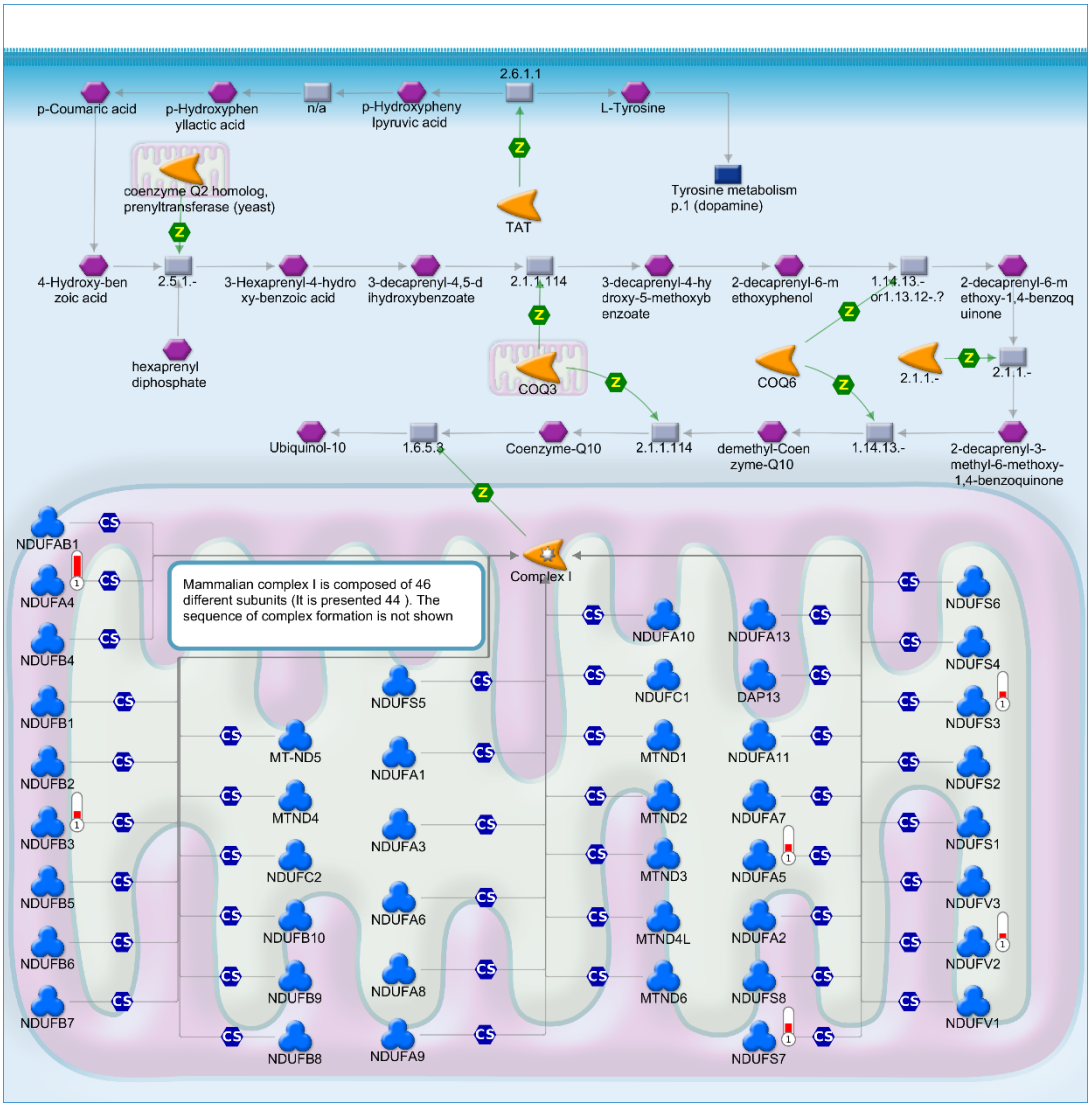


Figure 4.19. Ubiquinone metabolism with significant up regulated in iTregs compated to cultured CD4 T cells. The significant proteins were highlighted using a thermometer icon (red colour). Differentially expressed proteins shortlisted by pairwise comparison of quantitative mass spectrometry data generated from iTregs and cultured nTreg depleted CD4-T cells were analysed using MetaCore functional enrichment tool. Ubiquinone metabolism is one of the most significantly altered pathway in iTregs compared to CD4-T cells, significant proteins downregulated in iTregs compared to CD4 -T cells are overlaid in the whole pathway context.

4.3.1.7. ANN-based approach

The quantitative protein data (peak area) of 99 membrane proteins and 344 cytoplasmic proteins were analysed using ANN modelling stepwise approach to identify a predictive panel of proteins (biomarkers) that can differentiate iTregs among other $CD4^+$ T cells subpopulations. Two different methods of ANN analysis were used to analyse proteomic quantitative data, including single stepwise and multiple stepwise analysis.

The results of single stepwise analysis showed that 7 distinct proteins among 443 cytoplasmic and membrane proteins with the lowest average test error (ATE) were mathematically selected by the ANN as predictive biomarkers that differentiate iTregs from other $CD4^+$ T cells subpopulations. The distinct proteins are MAVS, PLP2, ITIH4, HEM6, CUL4A, P85A, XYLT1. Table 4.1 shows the summary results of novel membrane and cytoplasmic markers of iTregs selected by the ANN single stepwise analysis.

Table 4.11. A summary of the results of novel biomarkers of iTregs selected by the ANN single stepwise analysis.

Protein	Subcellular location	Number of selection*	Range of ATE
MAVS	Mitochondrial membrane	15	0.032 – 0.046
PLP2	Membrane	6	0.048 – 0.060
ITIH4	Membrane	5	0.050 – 0.058
HEM6	Mitochondrial membrane	5	0.030 – 0.046
CUL4A	Cytoplasmic, Nuclear	4	0.049 – 0.059
P85A	Membrane	3	0.046 – 0.060
XYLT1	Membrane, Intracellular	2	0.053 – 0.054

* Number of selection represents how many times a biomarker has been selected by the ANN in the overall single stepwise analysis.

The results of single stepwise analysis revealed that 6 distinct proteins among 443 cytoplasmic and membrane proteins were also selected by the ANN as predictive biomarkers that distinguish naïve nTregs from other $CD4^+$ T cells subpopulations. These proteins include HYEP, FERRITIN, SERPINB6A, NDKA, VWA8 and SAMH1. The summary results of novel membrane and cytoplasmic biomarkers of naïve nTregs are detailed in table 4.12.

Table 4.12. List of 6 novel biomarkers selected by the ANN single stepwise analysis to differentiate naive nTregs from other $CD4^{+}$T cells subsets.

Protein	Subcellular location	Number of selection	Range of ATE
HYEP	Membrane	12	0.048 – 0.060
FERRITIN	Cytoplasmic	10	0.040 – 0.050
SERPINB6A	Membrane	8	0.042 – 0.054
NDKA	Cytoplasmic	5	0.043 – 0.049
VWA8	Cytoplasmic, mitochondrial	3	0.041 – 0.053
SAMH1	Cytoplasmic, Nuclear	2	0.045 – 0.055

For multiple stepwise analysis, the results showed that 12 distinct proteins were selected by the ANN based on their ATE among 443 membrane and cytoplasmic proteins. These distinct proteins include MAVS, PLP2, ITIH4, CUL4A, HEM6, P85A, GSHR, SYAM, XYLT1, TFR1, CO4B and OGFR. Table 4.13 shows the summary results of novel membrane and cytoplasmic markers of iTregs selected by the ANN multiple stepwise analysis.

Table 4.13. A summary of the results of novel biomarkers of iTregs selected by the ANN multiple stepwise analysis.

Protein	Subcellular location	Number of selection	Range of ATE
MAVS	Mitochondrial membrane	17	0.018 – 0.050
PLP2	Membrane	12	0.017 – 0.039
ITIH4	Membrane	11	0.018 – 0.053
CUL4A	Cytoplasmic, Nuclear	11	0.030 – 0.058
HEM6	Mitochondrial membrane	10	0.013 – 0.046
P85A	Membrane	8	0.011 – 0.056
GSHR	Cytoplasmic, mitochondrial	5	0.006 – 0.025
SYAM	Cytoplasmic, mitochondrial	5	0.010 – 0.035
XYLT1	Membrane, Intracellular	4	0.017 – 0.029
TFR1	Membrane	3	0.015 – 0.036
CO4B	Extracellular	2	0.020 – 0.044
OGFR	Cytoplasmic, Nuclear	2	0.020 – 0.024

The results of multiple stepwise analysis also showed that 10 proteins were selected by the ANN as predictive distinct biomarkers that characterise naïve nTregs from

other subsets of $CD4^+$ T cells. The results of these distinct biomarkers are summarised in table 4.14.

Table 4.14. List of distinct proteins selected by the ANN multiple stepwise analysis as predictive biomarkers of naive nTregs

Protein	Subcellular location	Number of selection	Range of ATE
HYEP	Membrane	22	0.010 – 0.056
SERPINB6A	Membrane	17	0.011 – 0.057
SPA3F	Extracellular	15	0.018 – 0.052
VWA8	Cytoplasmic, mitochondrial	12	0.012 – 0.029
NKDA	Cytoplasmic	10	0.012 – 0.031
FERRITIN	Cytoplasmic	6	0.013 – 0.034
TCTP	Cytoplasmic	2	0.017 – 0.061
MIC26	Membrane, Intracellular	2	0.013 – 0.035
COF1	Membrane	2	0.015 – 0.033
ENOA	Membrane	2	0.020 – 0.057

4.3.1.8. Differentially expressed proteins in the membrane fraction of $CD4^+$ T cells

The most significant results of fold change in the expression of membrane proteins identified by SWATH MS profiling were shown in figure 4.20. The results showed that ITIH4 (inter alpha trypsin inhibitor heavy chain 4), XYLT1 (xylosyltransferase 1) and PLP2 (proteolipid protein 2) proteins were highly up regulated in iTregs compared to other subsets of $CD4^+$ T cells in which these proteins were found significantly downregulated. The results also showed that P85A (Phosphatidylinositol 3 kinase regulatory subunit alpha) and CUL4A (proteins were significantly overexpressed in iTregs compared to other $CD4^+$ T cell subsets which showed significant low expression for P85A and CUL4A proteins (figure 4.20). All the differentially expressed membrane proteins from all the pairwise comparisons are represented as a heatmap (Figure 4.21).

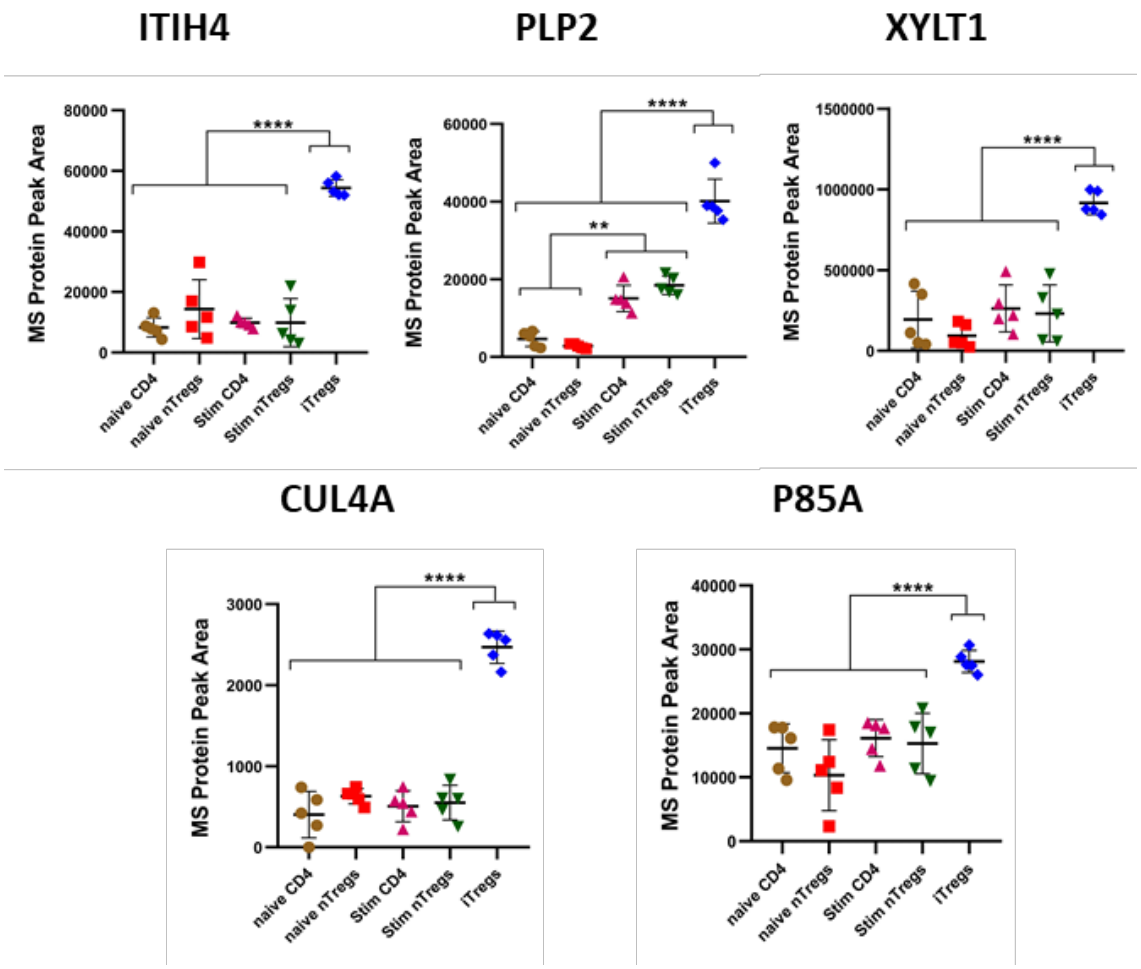


Figure 4.20. Scatter plots showing the protein peak area intensity (fold-change) of proteins identified significantly up-regulated in iTregs compared to other subsets of $CD4^+$ T cells. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, **; p-value < 0.0019). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

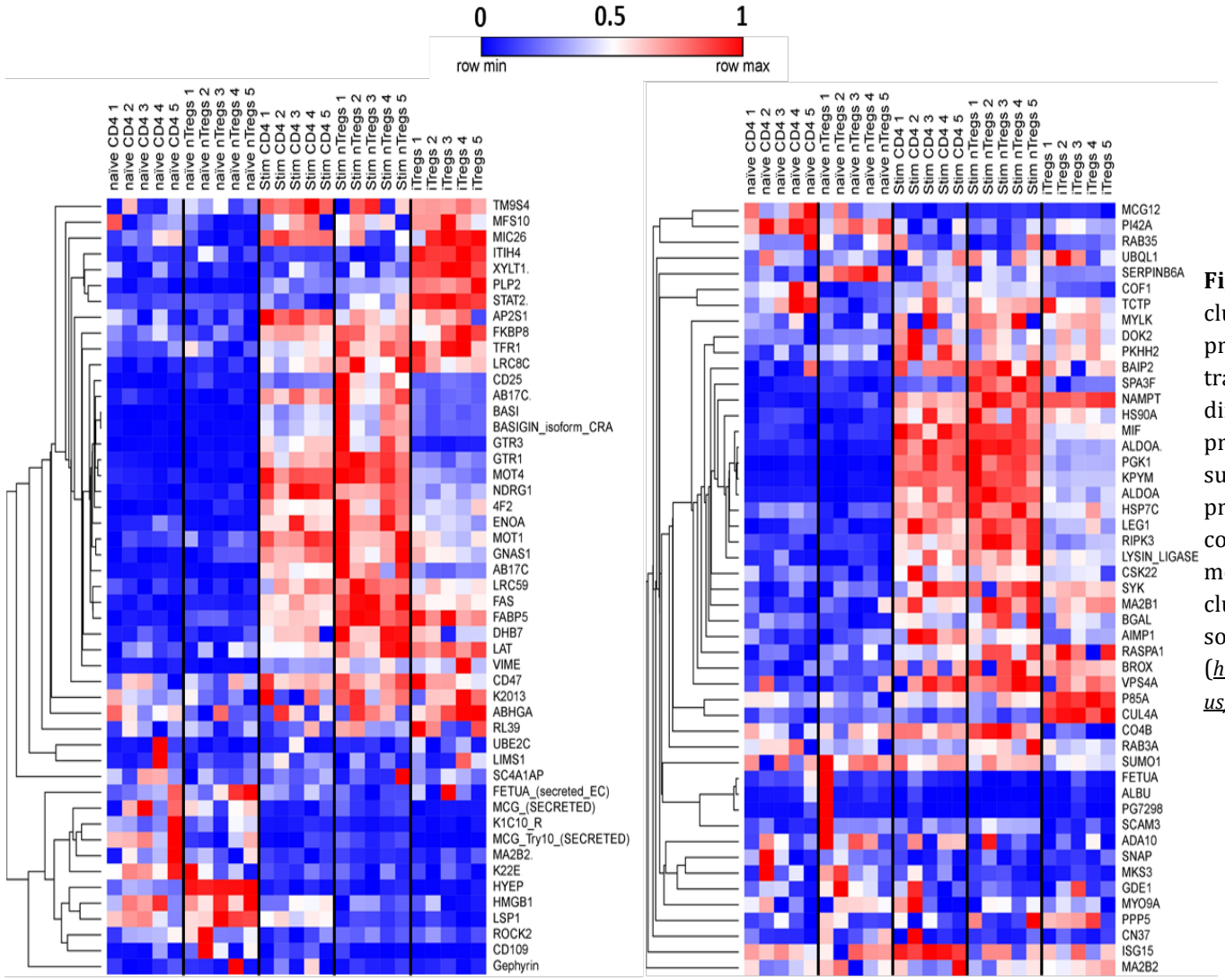


Figure 4.21. Heatmap and hierarchical clustering showing relative fold-change protein expression values (log_{10} transformed peaks intensity) of the 99 differentially expressed membrane proteins identified in the five different subsets of $CD4^+$ T cells using SWATH-MS proteomics. One minus pearson correlation metric with single linkage method was used for hierarchical clustering using MORPHEUS online software (*https://software.broadinstitute.org/morpheus/*. (Accessed on Sep 2019).

The results also showed that the expression of GTR1 (solute carrier family 2, facilitated glucose transporter member 1), MOT4 (monocarboxylate transporter 4), NDRG1 (N myc downstream regulated gene 1 protein), ALDOA (fructose biphosphate aldolase A), PGK1 (phosphoglycerate kinase 1), KPYM (pyruvate kinase PKM), LEG1 (Gelectin 1) and RIPK3 (receptor interacting serine/threonine protein kinase 3) proteins was found significantly up regulated in activated nTregs and $CD4^{+}CD25$ T cells subsets compared to iTregs and other naïve subsets (figure 4.22).

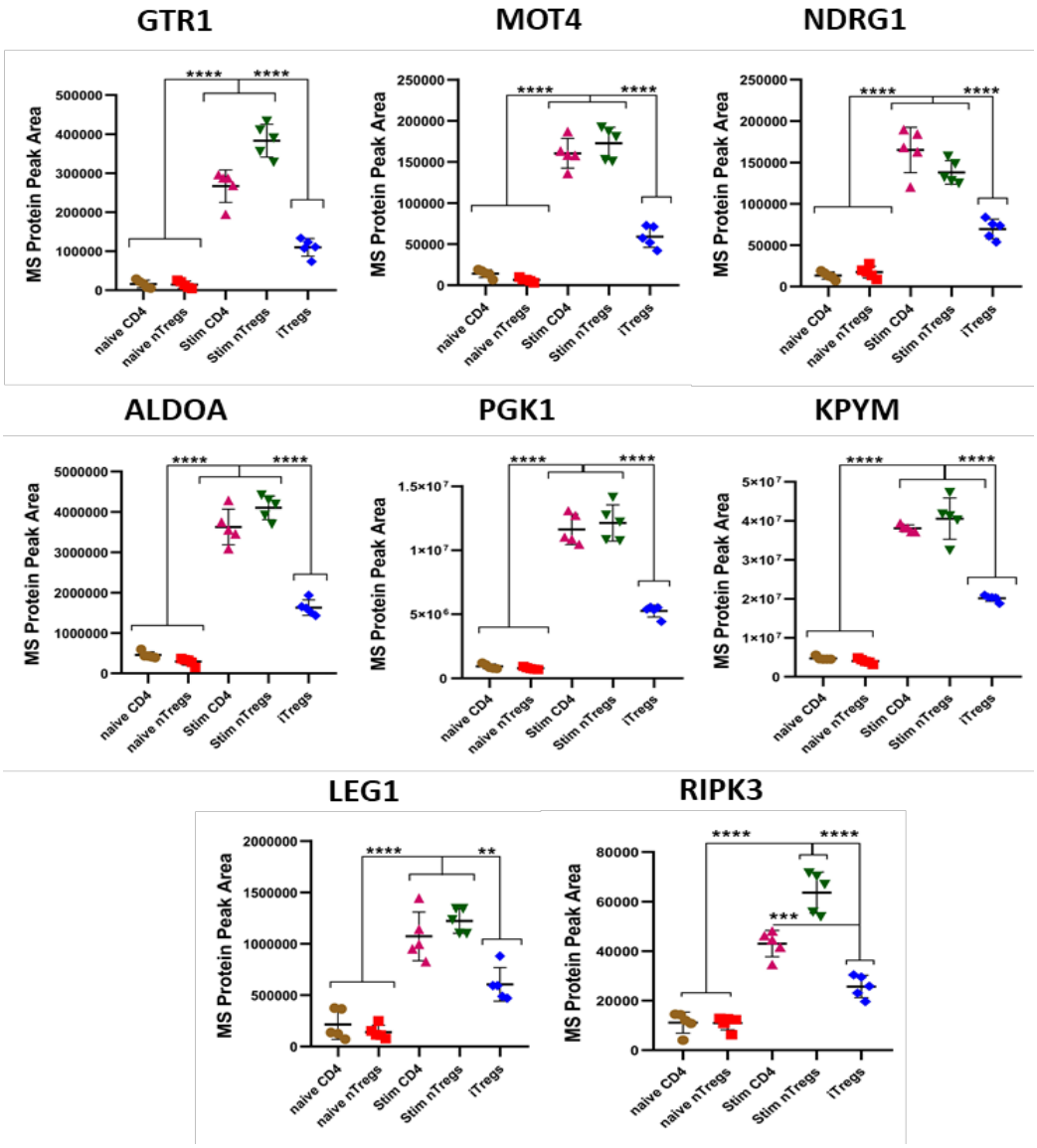


Figure 4.22. Scatter plots showing the protein peak area intensity (fold-change) of proteins identified significantly down-regulated in iTregs compared to other activated $CD4^{+}CD25^{-}$ T cells and activated nTregs subsets of $CD4^{+}$ T cells. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0007, **; p-value < 0.0022). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

The results also showed that the expression of HYEP (Epoxide hydrolase 1) protein was found significantly up regulated in naïve nTregs, while it was significantly down regulated in other subsets of CD4⁺ T cells. Another protein known as "SERPINB6A (serpin B6 domain A) was also found significantly up regulated in naïve nTregs compared to other CD4⁺ subsets (figure 4.23A).

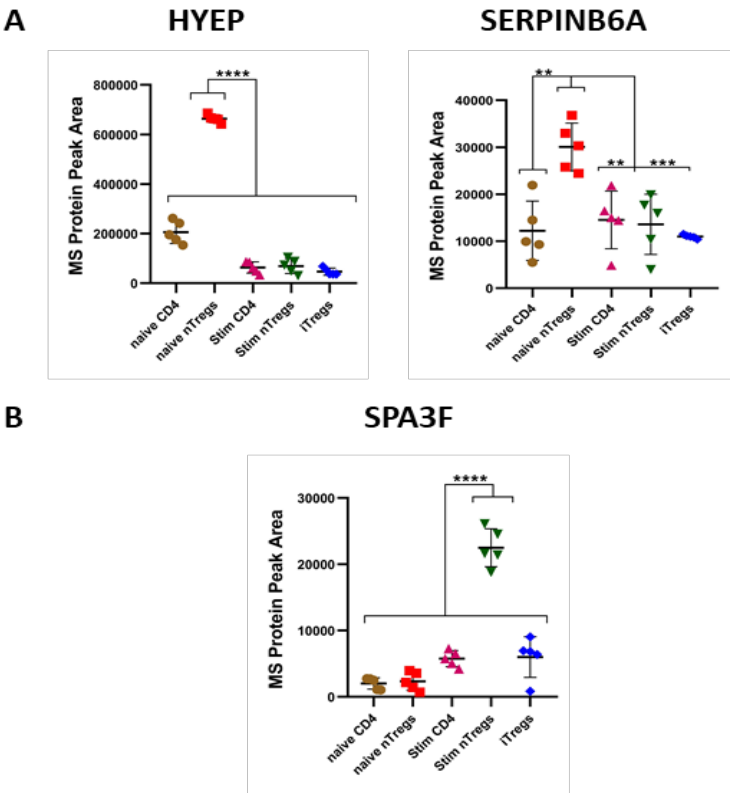


Figure 4.23. (A) Scatter plots showing the protein peak area intensity (fold-change) of proteins identified significantly up-regulated in naïve nTregs compared to other subsets of CD4⁺ T cells. **(B)** The protein peak area intensity (fold-change) of proteins identified significantly up-regulated in stimulated nTregs compared to other subsets of CD4⁺ T cells. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0005, **; p-value < 0.0016). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

The results also showed that the expression SPA3F (serine protease inhibitor A3F) protein was identified significantly up regulated in stimulated or activated nTregs, while it was significantly down regulated or expressed at very low levels in other subsets of CD4⁺ T cells (figure 4.23B).

4.3.1.9. Differentially expressed proteins in the cytoplasmic fraction of CD4+ T cells

The results of cytoplasmic and nuclear proteins quantification showed that the expression of MAVS (mitochondrial antiviral signalling), HEM6 (oxygen dependent coproporphyrinogen III oxidase, mitochondrial (CPOX)), ICS10 (integrator complex subunit 10), SYAM (alanine tRNA ligase, mitochondrial), OGFR (opioid growth receptor) and TYSY (thymidylate synthase) proteins was significantly up regulated in iTregs compared to other subsets of CD4+ T cells in which the expression of these proteins was detected at low levels (figure 4.24).

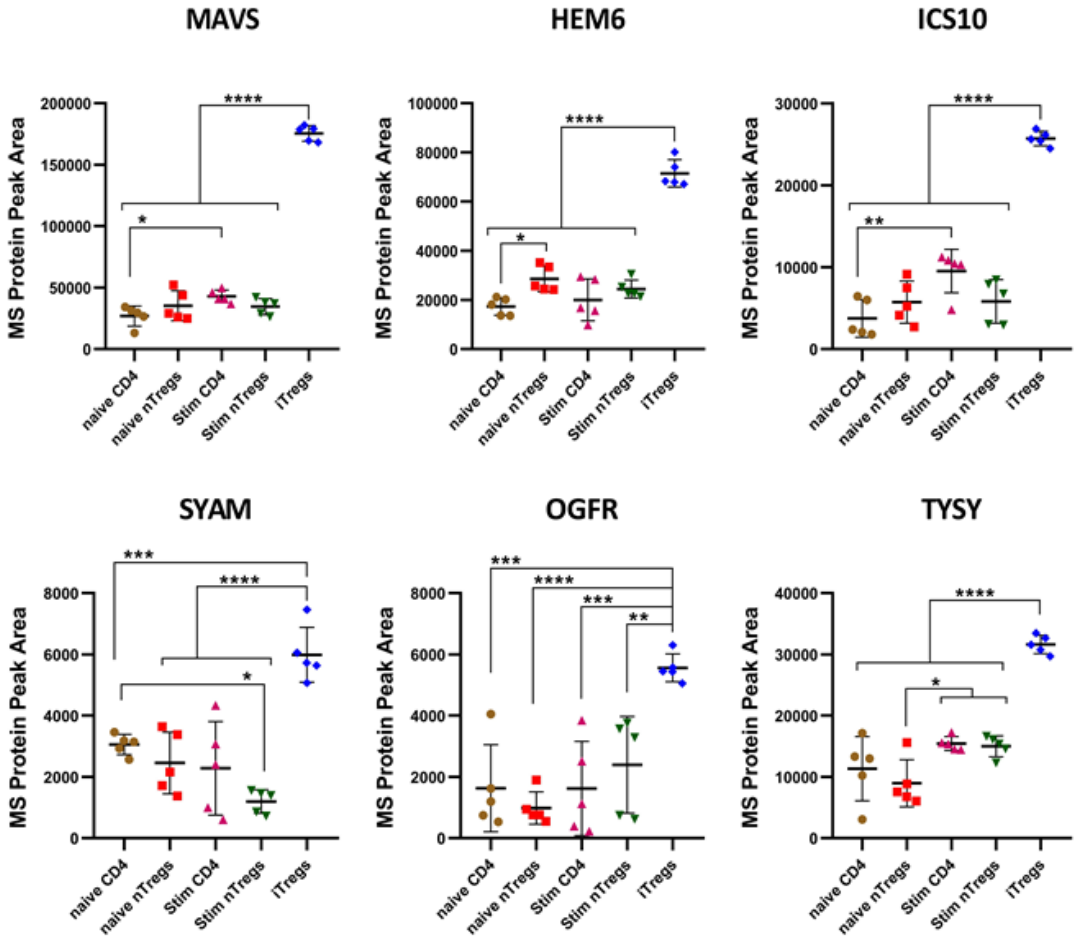


Figure 4.24. Scatter plots showing the protein peak area intensity (fold-change) of cytoplasmic and nuclear proteins identified significantly up-regulated in iTregs compared to other subsets of CD4+ T cells. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0007, **; p-value < 0.0012, *; p-value < 0.0325). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

The results also showed that among cytoplasmic and nuclear proteins identified, 12 proteins were found significantly up regulated in all stimulated or activated subsets of $CD4^+$ T cells, except iTregs in which the expression of proteins was significantly low compared to stimulated nTregs and $CD4^+CD25$ T cells although iTregs were activated at the same level of nTregs and $CD4^+CD25$ T cells. These proteins include TPIS (triosephosphate isomerase), G3P (glyceraldehyde 3 phosphate dehydrogenase), SERC (phosphoserine aminotransferase), SYYC (tyrosine tRNA ligase, cytoplasmic), RP10 (ribosomal protein 10), PGAM1 (phosphoglycerate mutase 1), PGM1 (phosphoglucomutase 1), GALK1 (galactokinase 1), LDHA (L lactate dehydrogenase A chain), RL34 (60S ribosomal protein L34), TPP2 (tripeptidyl peptidase 2) and SYSC (serine tRNA ligase, cytoplasmic) (figure 4.25).

The results showed that the expression of SAMH1 (Deoxynucleoside triphosphate triphosphohydrolase SAMHD1) and VWA8 (von Willebrand factor A domain containing protein 8) proteins was found significantly up regulated in naïve nTregs compared to other four subsets of $CD4^+$ T cells, especially stimulated nTregs in which the expression of SMAH1 and VWA8 proteins was significantly diminished after TCR activation (figure 4.26A). For cytoplasmic protein signature of activated nTregs, the results showed that the expression of CCS (Copper chaperone for superoxide dismutase) protein was significantly increased in activated nTregs, while its expression was detected at very low levels in naïve nTregs without activation (figure 4.26B).

The results also showed that the expression of DYR (Dihydrofolate reductase) protein was identified significantly up regulated in activated $CD4^+CD25$ T cells, while its expression was found significantly down regulated in other subsets of $CD4^+$ T cells (figure 4.26C).

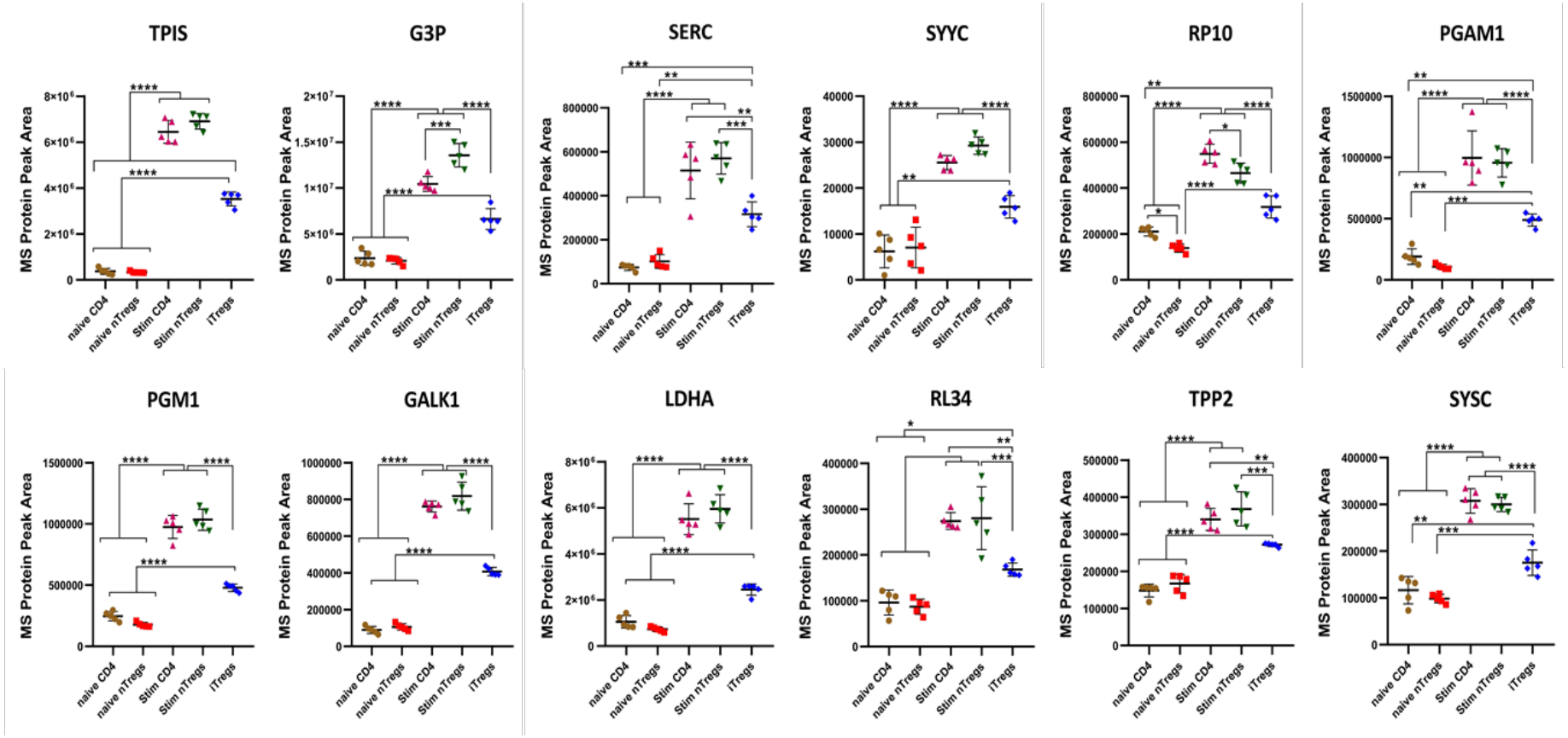


Figure 4.25. Scatter plots showing the protein peak area intensity (fold-change) of cytoplasmic and nuclear proteins identified significantly up-regulated in activated nTregs and $CD4^+CD25^-$ T cells, but down-regulated in iTregs. Stim $CD4^+CD25^-$ T cells, stim nTregs and iTregs were received the same TCR activation signals. Naïve $CD4^+CD25^-$ T cells and nTregs were not activated. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0007, **; p-value < 0.0019, *; p-value < 0.0425). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

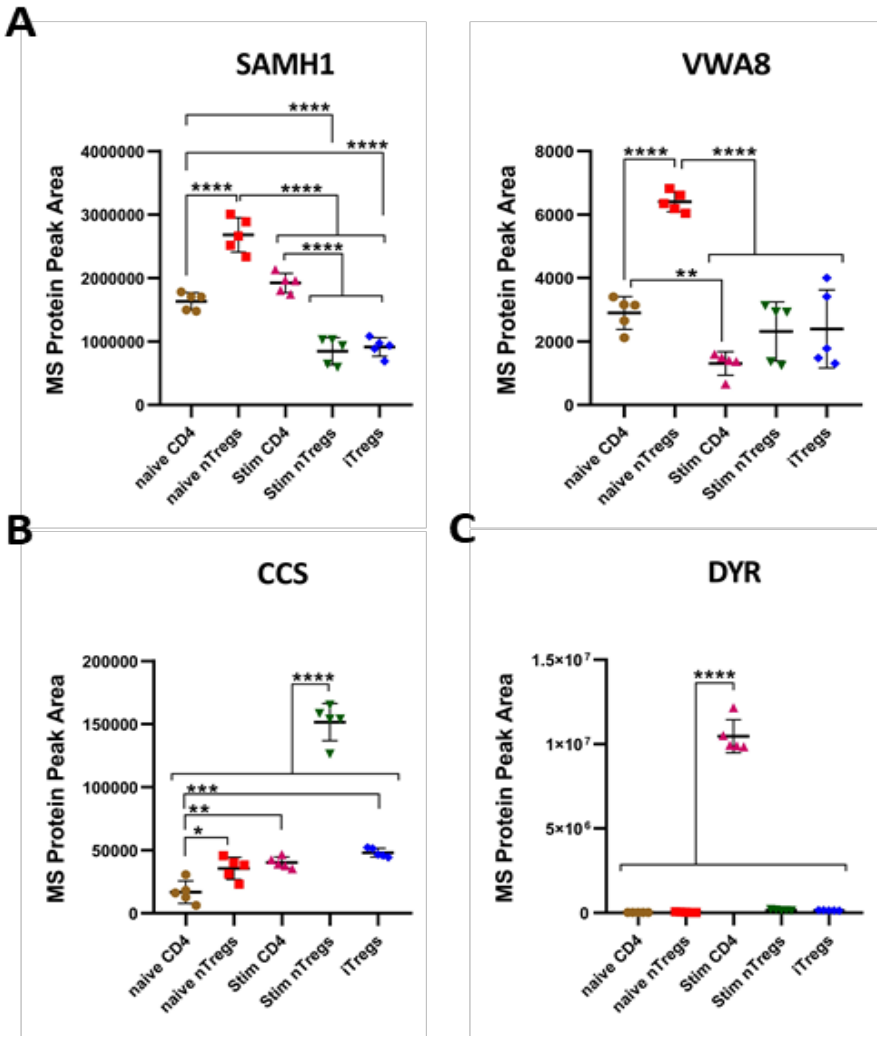


Figure 4.26. Scatter plots showing the protein peak area intensity (fold-change) of cytoplasmic and nuclear proteins identified significantly up-regulated in **(A)** naïve nTregs, **(B)** activated nTregs and **(C)** activated $CD4^{+}CD25^{-}$ T cells, compared to other subsets of $CD4^{+}$ T cells. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0004, **; p-value < 0.0030, *; p-value < 0.0444). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

To distinguish naïve from activated subpopulations of CD4+ T cells, the results showed that six distinct proteins were identified overexpressed mainly in both naïve subsets of CD4+ T cells; naïve CD4+CD25 T cells and nTregs, whereas these six proteins were identified expressed at low levels in all activated subsets of CD4+ T cells (figure 4.27). These proteins included ACOT2 (Acyl coenzyme A thioesterase 2, mitochondrial), MDHC (Malate dehydrogenase, cytoplasmic), SPRE (Sepiapterin reductase), PGM2 (Phosphoglucomutase 2), TRNT1 (CCA tRNA nucleotidyltransferase 1, mitochondrial), and GSHR (Glutathione reductase, mitochondrial).

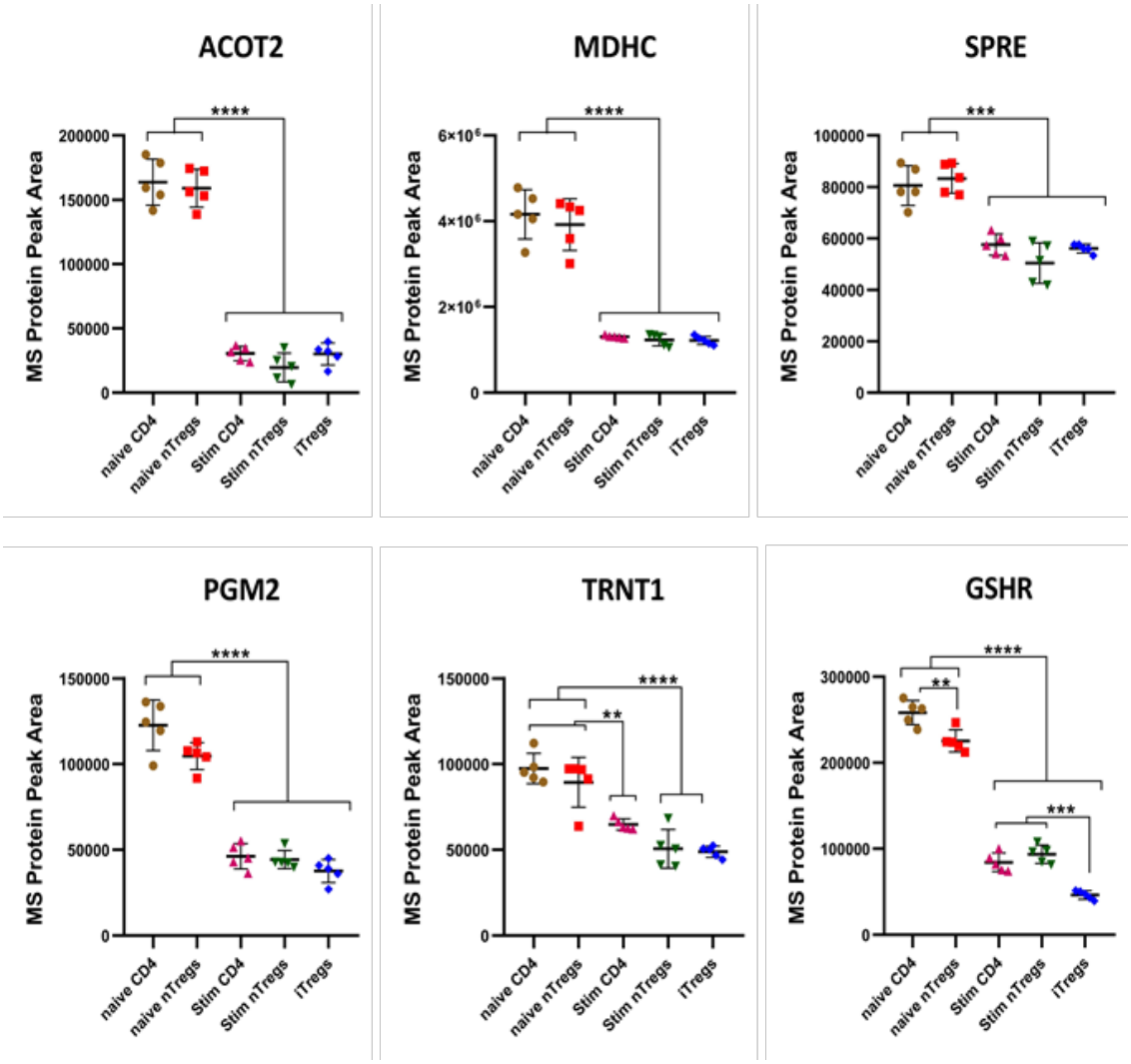


Figure 4.27. Scatter plots showing the protein peak area intensity (fold-change) of cytoplasmic and nuclear proteins identified significantly up-regulated in naïve subsets of CD4+ T cells, compared to stimulated subsets. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0008, **; p-value < 0.0026, *; p-value < 0.0387). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

The results also showed that 11 distinct proteins were identified overexpressed particularly in activated subsets of CD4+ T cells including activated CD4+CD25 T cells, activated nTregs and iTregs compared to naïve subsets; naïve CD4+CD25 T cells and nTregs, in which these proteins were detected at low levels of expression (figure 4.28). these proteins were SYMC (Methionine tRNA ligase, cytoplasmic), SYVS (Valine tRNA ligase), IPYR (Inorganic pyrophosphatase), STATHMIN, IMPA2 (Inositol monophosphatase 2), EF1G (Elongation factor 1 gamma), MCM2 (DNA replication licensing factor MCM2), MCM3 (DNA replication licensing factor MCM3), MCM5 (DNA replication licensing factor MCM5), MCM6, DNA replication licensing factor MCM6, MCM7 (DNA replication licensing factor MCM7).

The overall results of the expression fold change of 344 cytoplasmic and nuclear proteins identified and compared between five different subsets of $CD4^+$ T cells using SWATH MS proteomics are shown in figures 4.29, 4.30, 4.31 and 4.32).

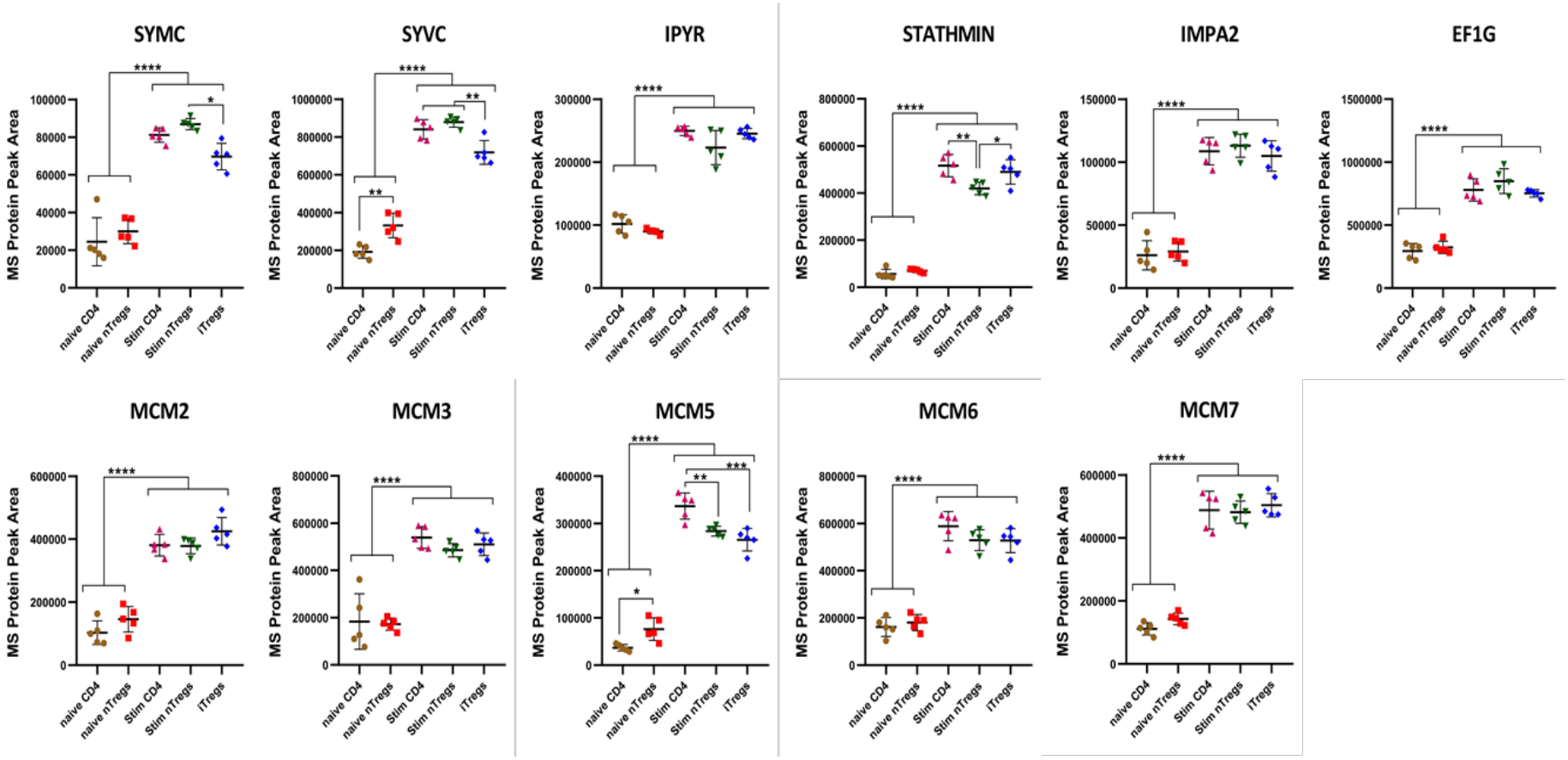


Figure 4.28. Scatter plots showing the protein peak area intensity (fold-change) of cytoplasmic and nuclear proteins identified significantly up-regulated in activated subsets of CD4+ T cells, compared to naïve subsets. Ordinary one-way ANOVA test was used for statistical analysis using GraphPad Prism 8.0.1 software. (****; p-value < 0.0001, ***; p-value < 0.0008, **; p-value < 0.0032, *; p-value < 0.0404). "Stim" means activated. Five independent experiments were carried out (n=5), each experiment compared five different cell populations. The statistical analysis was carried out on a limited data set (n=5) and hence is not powered adequately. However, the results shown here confirmed the reproducibility and robustness of the data. Bars errors represent the standard error (SE) of the mean.

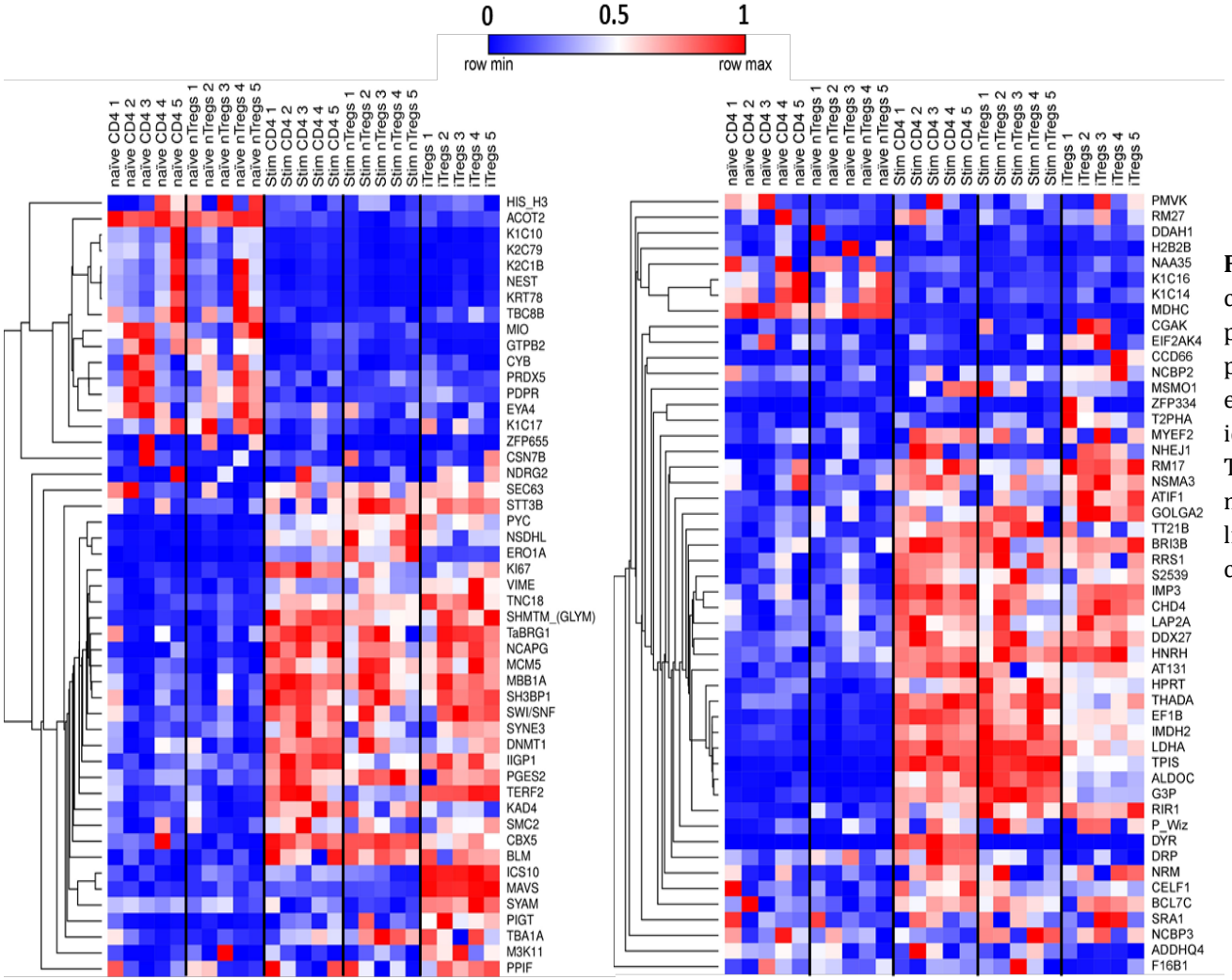


Figure 4.29. Heatmap and hierarchical clustering showing relative fold-change protein expression values (log_{10} transformed peaks intensity) of the 344 differentially expressed cytoplasmic and nuclear proteins identified in the five different subsets of $CD4^+$ T cells using SWATH-MS proteomics. One minus pearson correlation metric with single linkage method was used for hierarchical clustering using MORPHEUS online software.

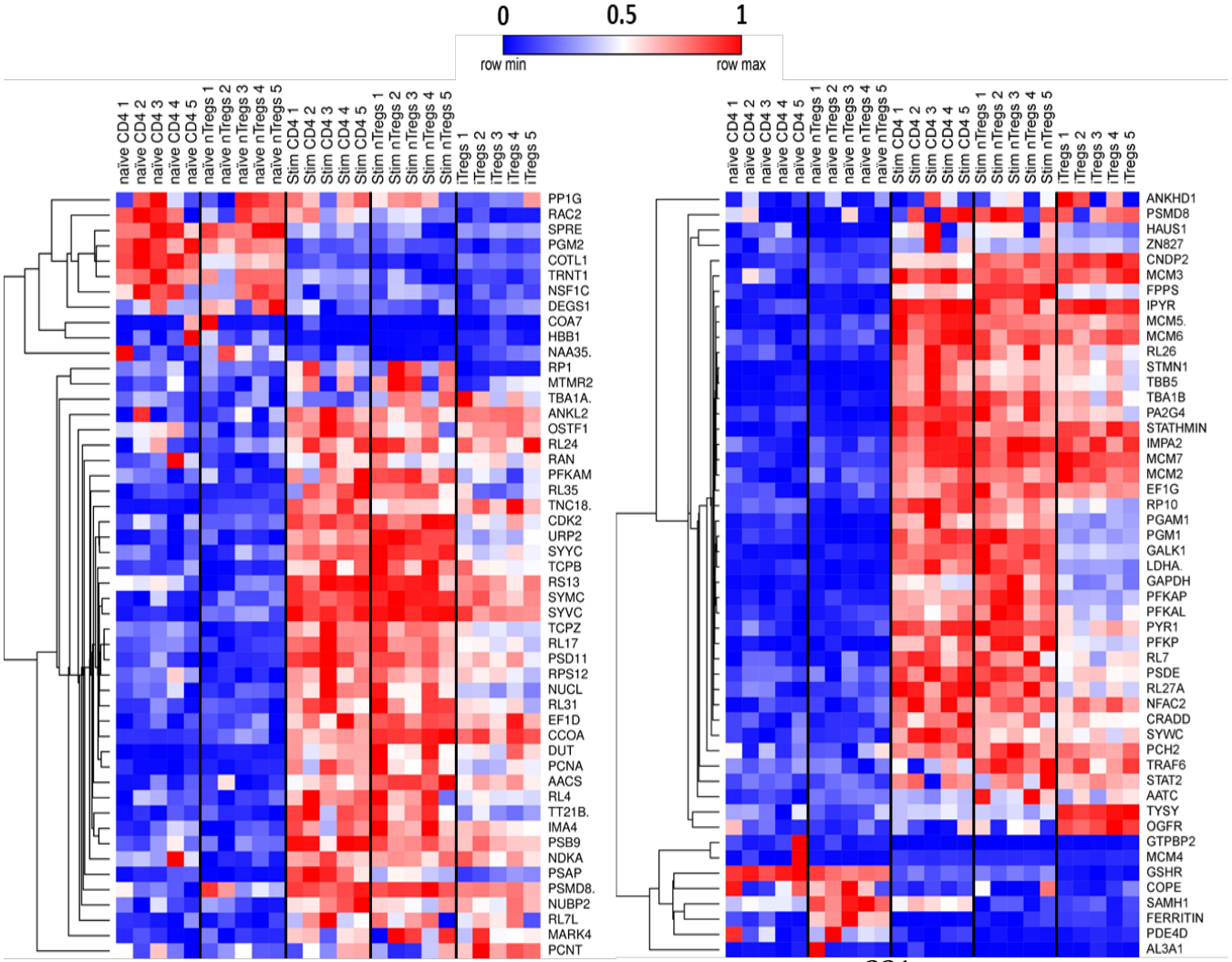


Figure 4.30. Heatmap and hierarchical clustering continued from figure 4.29.

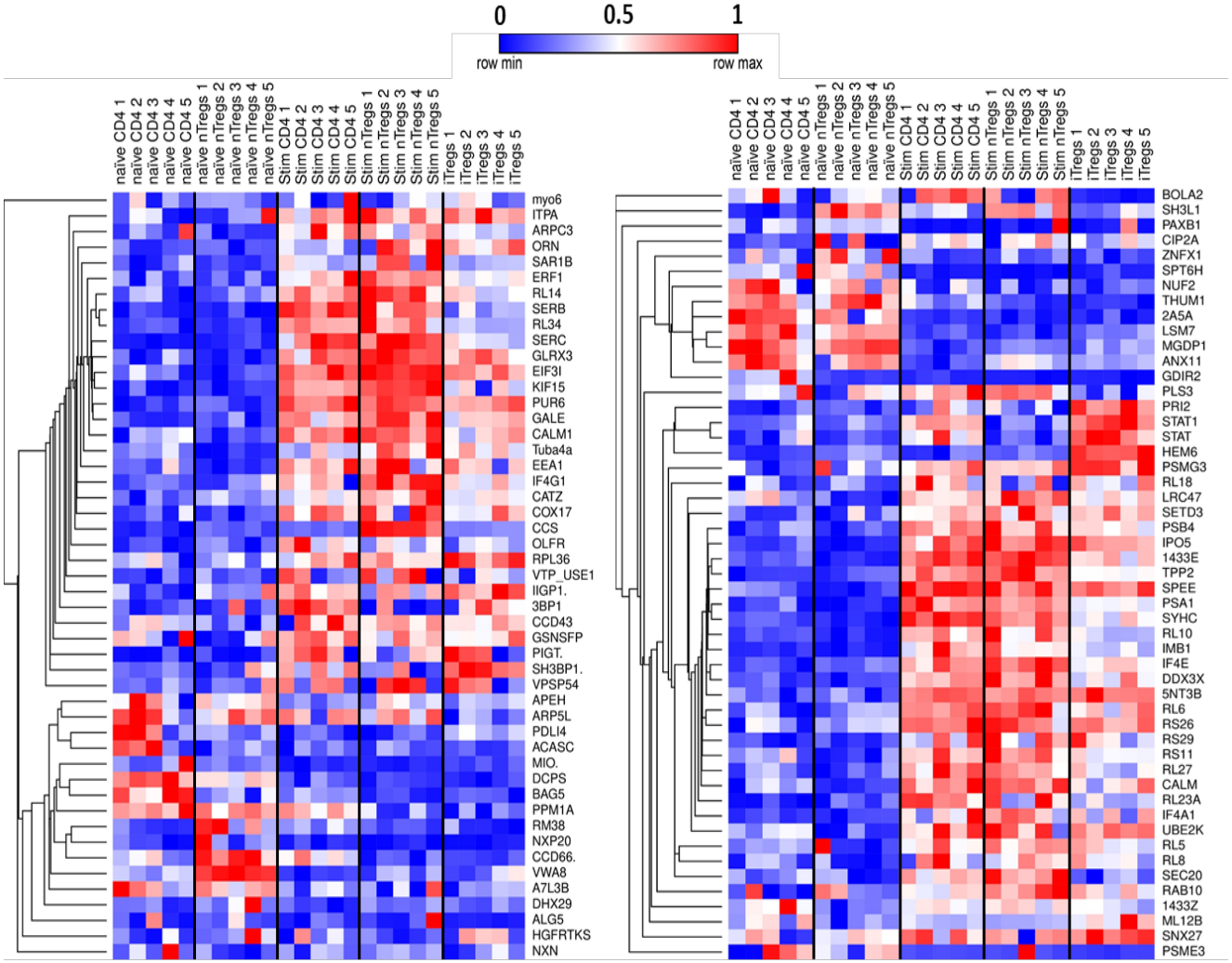


Figure 4.31. Heatmap and hierarchical clustering continued from figure 4.29.

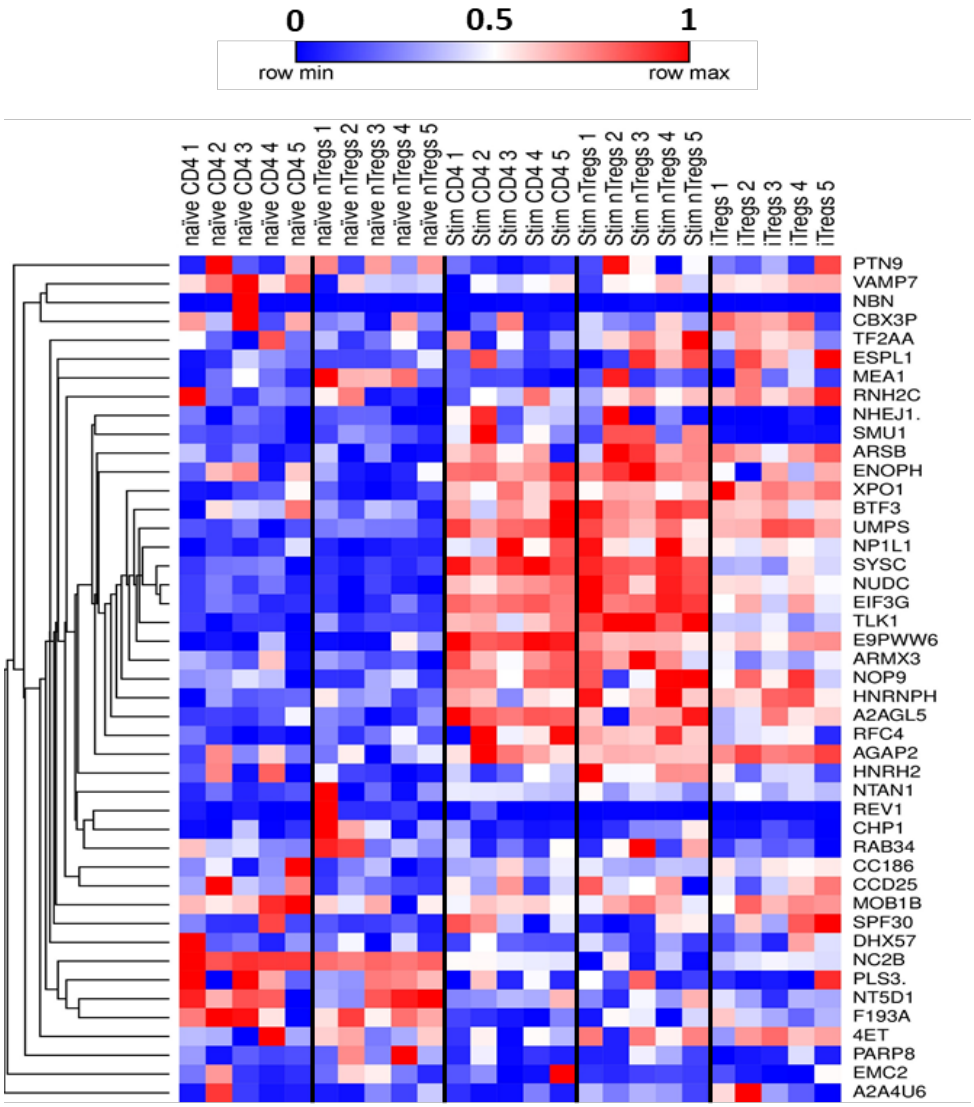


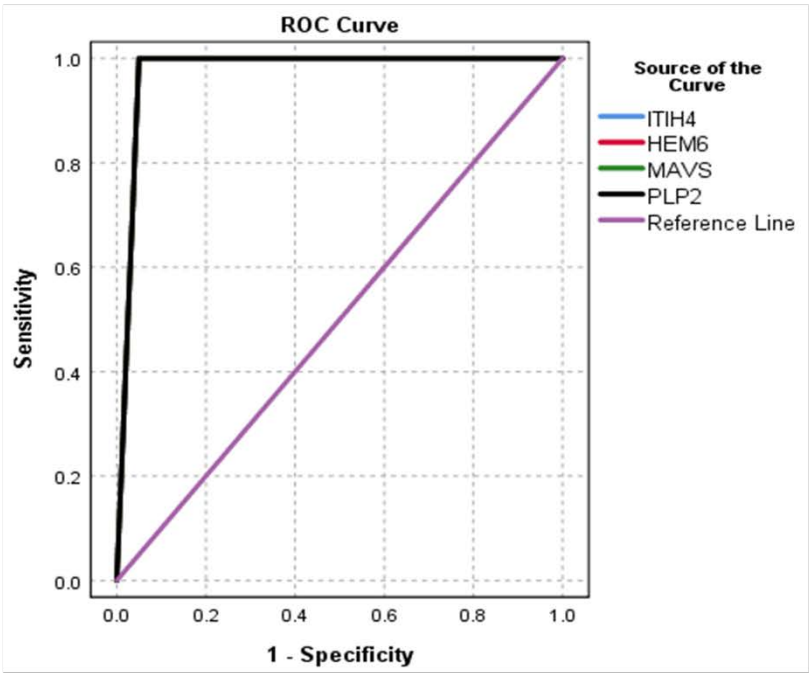
Figure 4.32. Heatmap and hierarchical clustering continued from figure 4.29.

4.3.1.10. Selection a panel of novels biomarkers of iTregs and nTregs for further verification and validation

The selection criteria of novel biomarkers for further verification and/ or validation was based on the subcellular location of differentially expressed proteins, selection by the ANN approach in a correlation with the results of statistical analysis (as shown in the sections 4.3.1.8 and 4.3.1.9). The selection priority was for differentially expressed membrane proteins of iTregs and nTregs. Accordingly, MAVS, HEM6, PLP2 and ITIH4 proteins were selected as novel markers of iTregs for further verification and validation based on the results of the ANN approach and statistical analysis. HYEP protein was also selected as a novel biomarker of naïve nTregs for further verification and validation using flow cytometry.

Prior to verification and validation, the results of ANN approach and statistical analysis were further tested using a receiver operating characterisation (ROC) analysis to assess the sensitivity and specificity of selected proteins and to assess the performance level of the ANN stepwise modelling approach in differentiating iTregs and naïve nTregs among other $CD4^+$ T cells subpopulation. The results of ROC analysis showed that the combination of MAVS, HEM6, PLP2 and ITIH4 proteins could be used as novel biomarkers to distinguish iTregs from other $CD4^+$ T cells subsets (figure 4.33). The results of ROC analysis also revealed that the panel of HYEP and SERPINB6A proteins could be used as novel biomarkers to differentiate naïve nTregs from other subpopulations of $CD4^+$ T cells (figure 4.34). The results of ROC analysis confirm the efficiency of the ANN modelling approach in identifying biomarkers that can distinguish a group of samples among various groups.

A



B

Area Under the ROC Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
PLP2	1.000	.000	.000	1.000	1.000
ITIH4	1.000	.000	.000	1.000	1.000
HEM6	1.000	.000	.000	1.000	1.000
MAVS	1.000	.000	.000	1.000	1.000

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

C

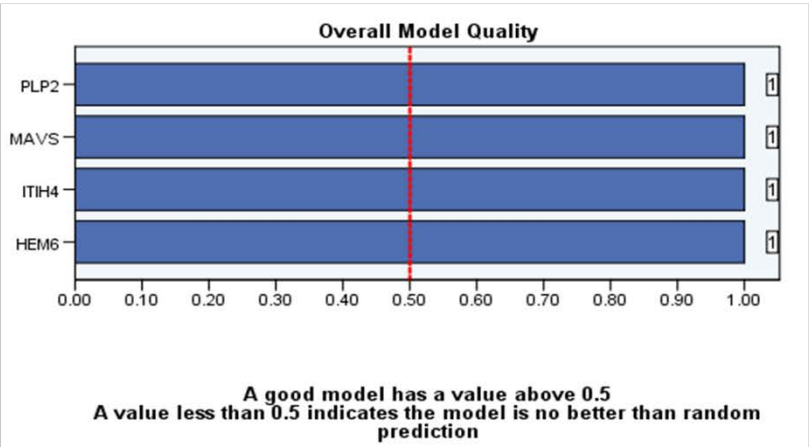
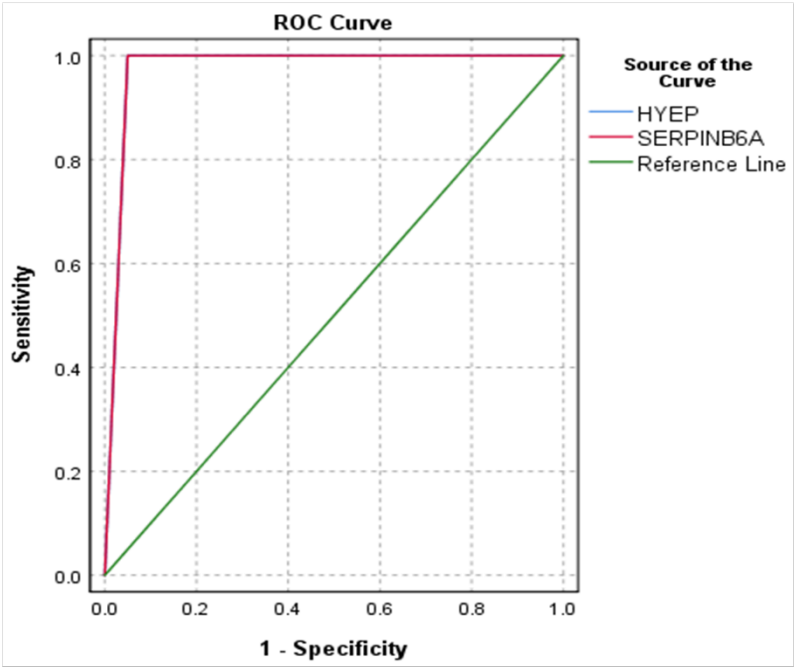


Figure 4.33. (A) ROC curve showing the sensitivity and specificity of MAVS, HEM6, ITIH4 and PLP2 as a predictive panel of novel biomarkers that are differentially expressed in iTregs compared to other CD4+ T cells subpopulations. Extracted SWATH-MS quantitative data of protein were used for ROC analysis. **(B)** Table illustrating a summary of the area under the curve (AUC) results for the selected proteins. **(C)** Bar graph showing the overall model quality and performance of the proteins selected by the ANN modelling approach. IBM SPSS Statistics 26 software was used for ROC curve analysis.

A



B

Area Under the ROC Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
HYEP	1.000	.000	.000	1.000	1.000
SERPINB6A	1.000	.000	.000	1.000	1.000

a. Under the nonparametric assumption

b. Null hypothesis: true area = 0.5

C

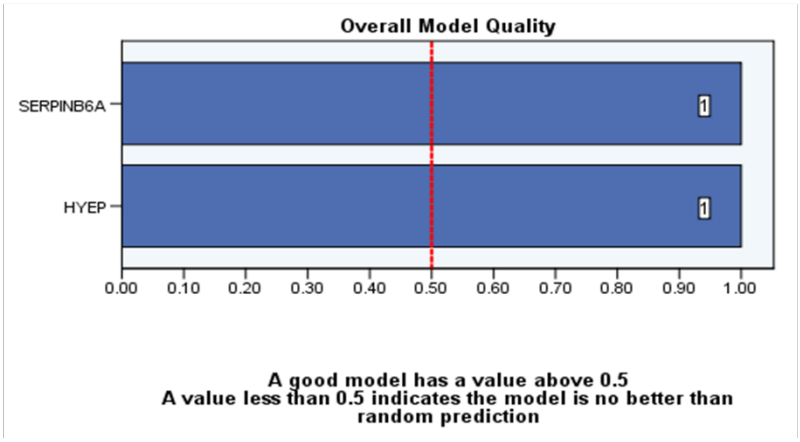


Figure 4.34. (A) ROC curve showing the sensitivity and specificity of HYEP and SERPINB6A as a predictive panel of novel biomarkers that are differentially expressed in naïve nTregs compared to other CD4+ T cells subpopulations. Extracted SWATH-MS quantitative data of protein were used for ROC analysis. **(B)** Table illustrating a summary of the area under the curve (AUC) results for the selected proteins. **(C)** Bar graph showing the overall model quality and performance of the proteins selected by the ANN modelling approach. IBM SPSS Statistics 26 software was used for ROC curve analysis.

4.3.1.11. Verification and/or validation of the selected proteins using flow cytometry and *in silico* tools

Different commercial conjugated and unconjugated antibodies were used to validate the selected proteins using flow cytometry as mentioned in section 4.2.4. The four different subpopulations of $CD4^+$ T cells were stained with anti- HEM6, MAVS, PLP2, ITIH4 and HYEP antibodies to verify the expression of the selected proteins in the population of iTregs and nTregs, in addition to the other subpopulations as negative controls (naïve $CD4^+CD25^-$ and stimulated $CD4^+CD25^-$ T cells). Unfortunately, the results of flow cytometry did not show any staining for the expression of the selected biomarkers in all of the subpopulations (figure 4.35). Possible reasons are discussed in section 4.4

BloodSpot *in silico* analysis of the shortlisted biomarkers highlighted its expression pattern in haematopoietic cell lineages. The analysis of the gene expression data found that there is a ubiquitous expression of MAVS and ITIH4 in most of the cell types in the murine immune system (Figure 4.36B and 4.38). The expression of PLP2, HEM6 (CPOX), and HYEP (EPHX1) were found to be minimum in most of the murine immune cells (Figure 4.36A, 4.37 A&B).

Web-based analysis of The Cancer Genome Atlas (TCGA) using OncoLnc in 8,647 patients of 21 cancer types indicated Kidney Renal Clear Cell Carcinoma (KIRC) as one of the top cancers with FOXP3 expression correlated with poor prognosis indicating a significant role of Tregs in KIRC (Figure 4.39A). Therefore, all the human equivalents of the shortlisted genes (ITIH4, PLP2, MAVS, HEM6 (CPOX), and HYEP (EPHX1)) were subsequently investigated in KIRC to see whether any of these genes explain similar survival outcome. Three out of five genes were found to be highly correlated with poor survival prognosis indicating possible role of induced Treg population in KIRC microenvironment (Figure 4.39 B, C, and D). This need further validation and confirmation using a human specific antibody against the above-mentioned genes.

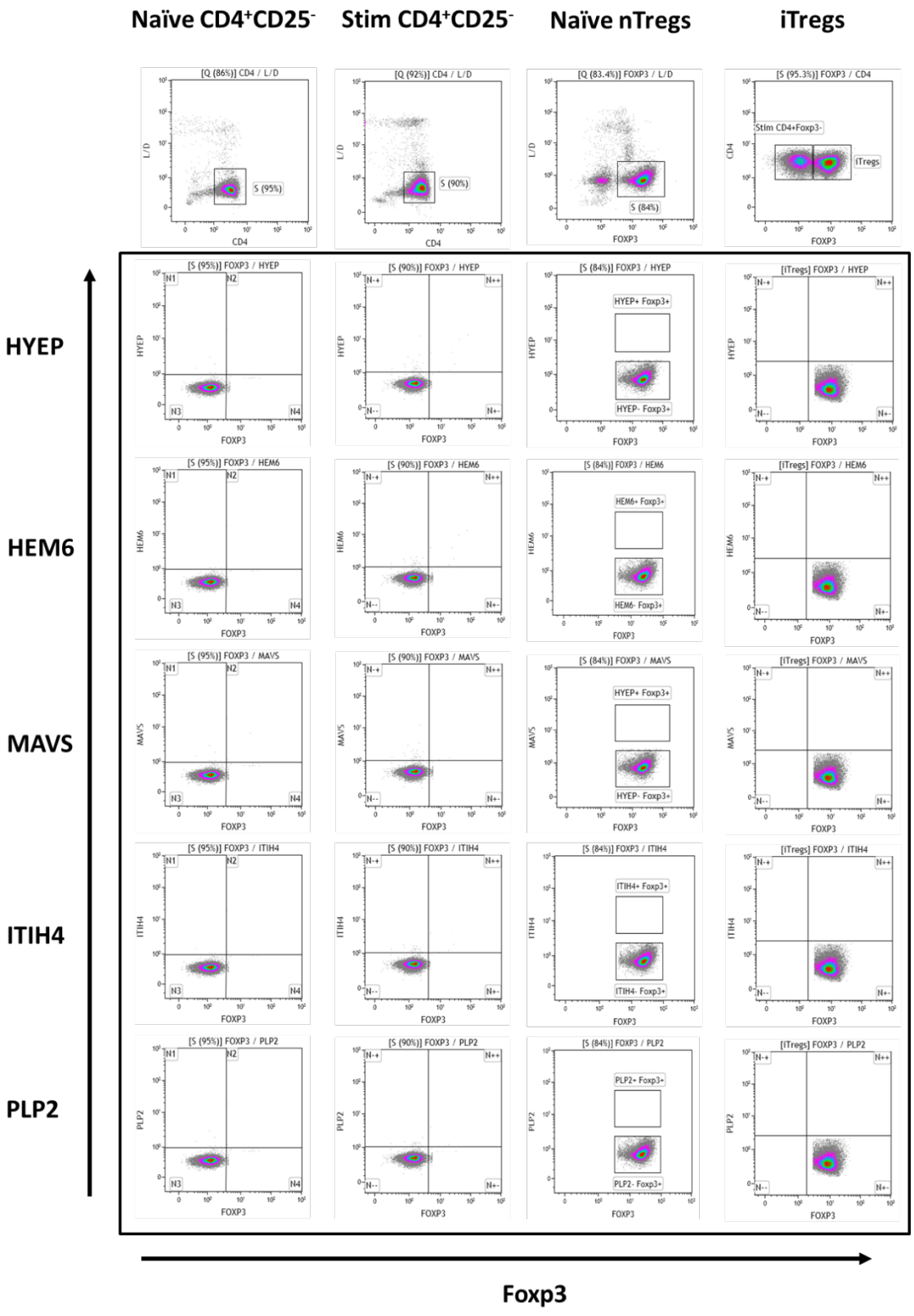


Figure 4.35. Flow cytometric density plots showing the results of verification of the selected proteins. Naïve nTregs were negative for the expression of HYEP, whereas iTregs showed negativity for the expression of HEM6, MAVS, ITIH4 and PLP2. Naïve and stimulated CD4+CD25- T cells as a control were also negative for the expression of all selected proteins. Kalusa software 1.3 was used for flow cytometry analysis.

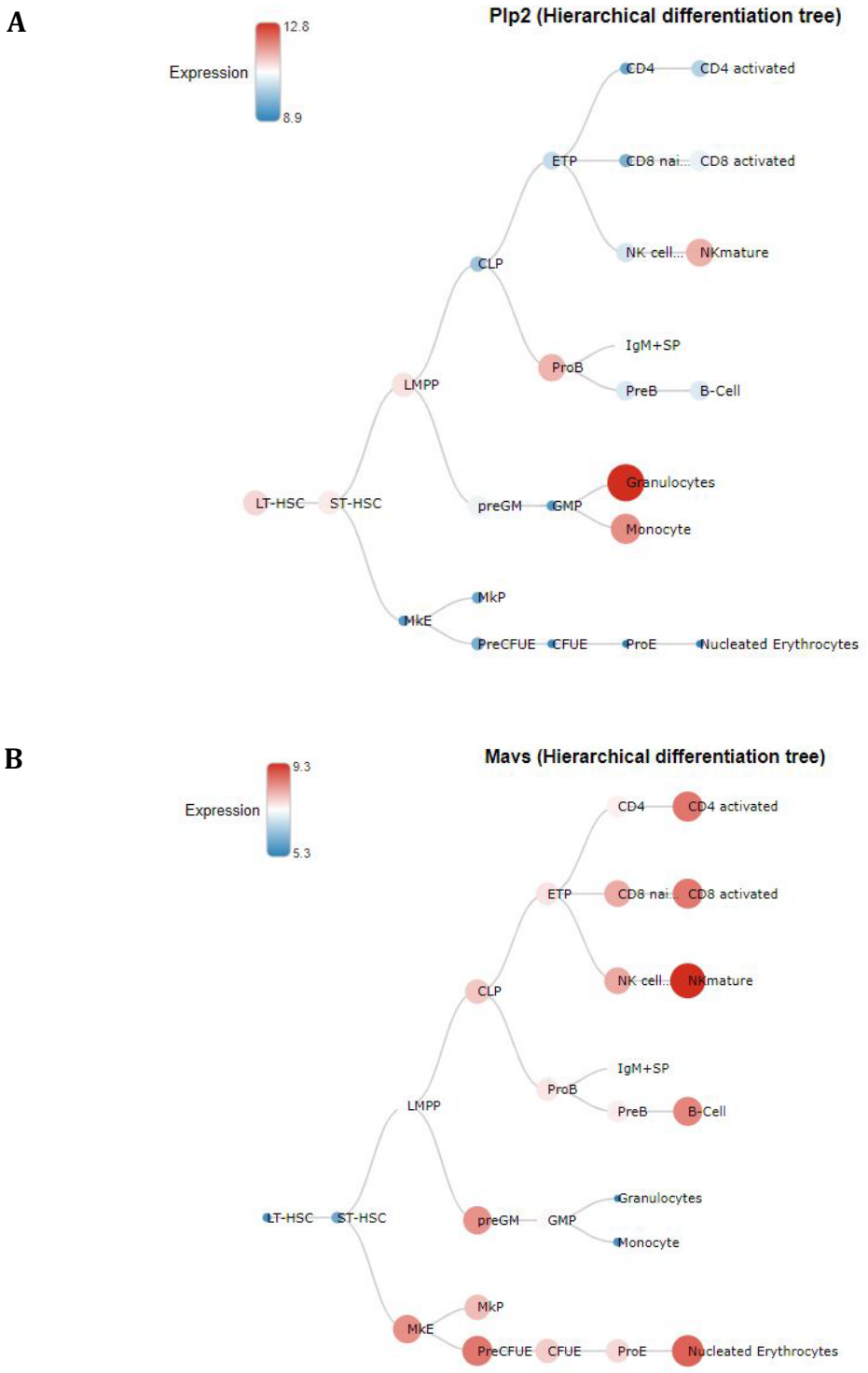
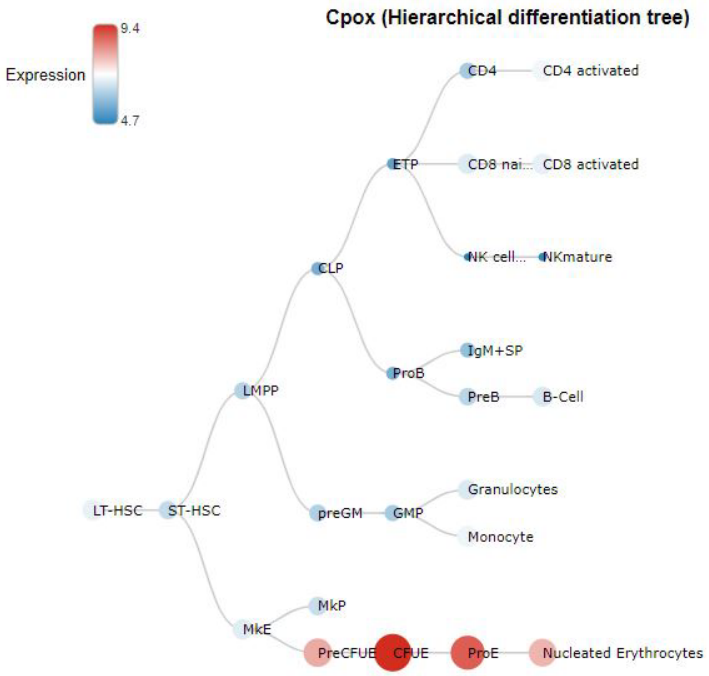


Figure 4.36. Hierarchical differentiation tree of markers PLP2 and MAVS generated using normal murine haematopoiesis data set through BloodSpot online portal. The expression pattern of each gene from haematopoietic stem cell through its progenitors up to the effector cells are displayed. Red colour indicates maximum expression and blue colour represent minimum expression.

A



B

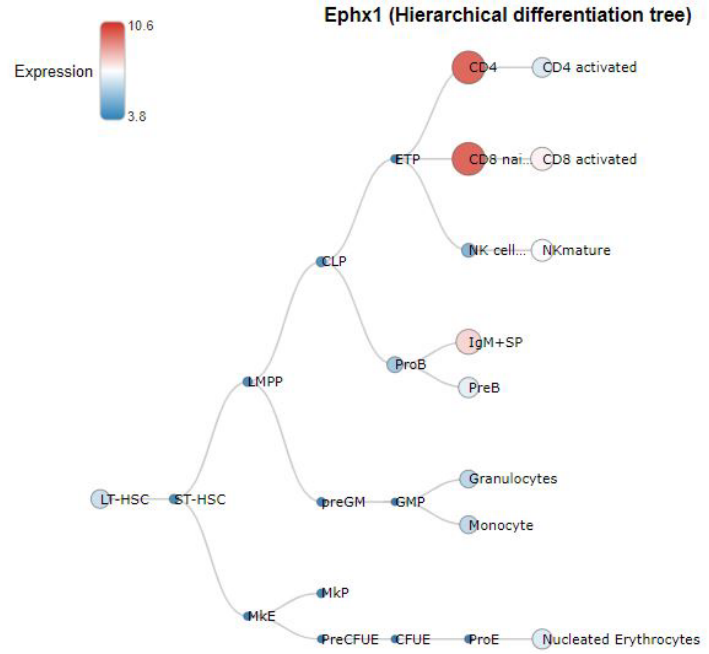


Figure 4.37. Hierarchical differentiation tree of markers HEM6 (CPOX) and HYEP (EPHX1) generated using normal murine haematopoiesis data set through BloodSpot online portal. The expression pattern of each gene from haematopoietic stem cell through its progenitors up to the effector cells are displayed. Red colour indicates maximum expression and blue colour represent minimum expression.

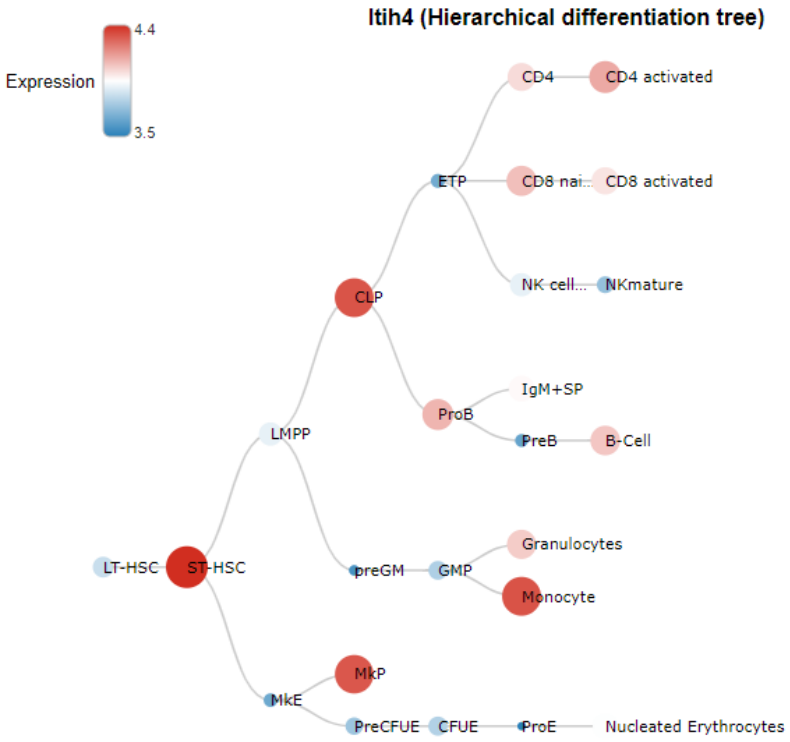


Figure 4.38. Hierarchical differentiation tree of marker ITIH4 generated using normal murine haematopoiesis data set through BloodSpot online portal. The expression pattern of each gene from haematopoietic stem cell through its progenitors up to the effector cells are displayed. Red colour indicates maximum expression and blue colour represent minimum expression.

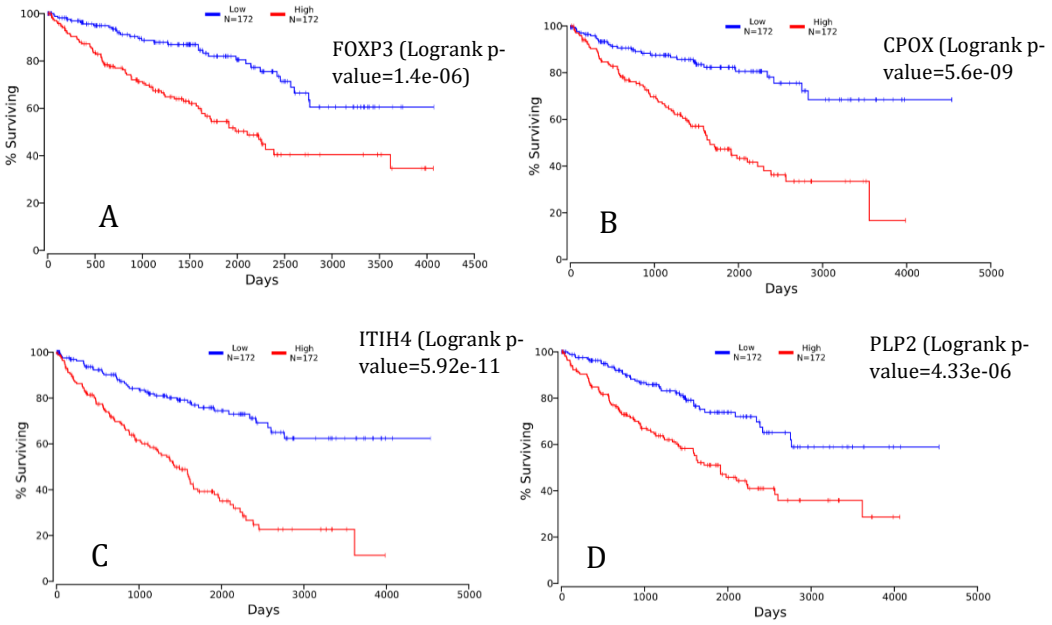


Figure 4.39. Three out of five shortlisted markers associated with poor survival in cancer patients. (A-D). Survival analysis was performed using OncoLnc online survival analysis tool (TCGA). First, Tumour which showed high association of FOXP3 expression with poor survival outcome was identified as kidney renal clear cell carcinoma (KIRC) among 21 cancer types. The association of shortlisted biomarkers were then assessed in KIRC using upper and lower quartiles (33%). X- axis represent % survival and Y-axis represent number of days survived. Blue line indicates patient with low expression of corresponding gene and the red line indicates the high expression. Significance of difference in survival probability between high and low group of each genes were calculated using Log rank p value, p-Value less than 0.05 was considered as significant.

4.4. Discussion

The presence of Tregs within the tumour microenvironment is found significantly correlated with poor prognosis in different types of cancer (Tanaka and Sakaguchi 2017, Shang, et al. 2015b). Although the origin of tumour infiltrating Tregs is still controversial, a study has been reported that tumour infiltrating Tregs are mainly developed or induced from naïve $CD4^+$ T cells that are recruited into the site of tumour (Su, et al. 2017a). Tumour infiltrating Tregs contribute to the progression of tumour growth via immune evasion by inhibiting anti-tumour immune response (Takeuchi and Nishikawa 2016). Systemic depletion of Tregs has contributed to the development of autoimmune diseases in humans, thereby suggesting the requirement for approaches to selectively targeted tumour infiltrating Tregs to enhance anti-tumour immunity (Togashi, et al. 2019). Such approaches require sensitive and specific cell surface drug targets/biomarkers that can distinguish between naturally occurring Tregs (nTregs) and induced or peripherally developed Tregs (iTregs). To date, phenotypic signature of Tregs is defined by the expression of CD25 and Foxp3 proteins which are expressed in both nTregs and iTregs. Therefore, neither CD25 nor Foxp3 can be used for differentiating nTregs from iTregs (Lin, et al. 2013). Many studies have been conducted to identify distinct biomarkers that can differentiate nTregs from iTregs. Neuropilin-1 (Nrp-1) was firstly identified as a surface marker expressed constitutively on nTregs (Bruder, et al. 2004). However, Milpied et al. have demonstrated that Nrp-1 is not a specific marker for nTregs as a population of $CD4^+Foxp3^-$ T cells were found to express Nrp-1. Also, the authors found that Nrp-1 expression was acquired by T cells after activation, suggesting that Nrp-1 could be used as an activation marker for T cells (Milpied, et al. 2009). Soon after, Thornton et al. have found that Helios is exclusively expressed on nTregs and can be used as a nuclear marker to differentiate nTregs from iTregs (Thornton, et al. 2010). However, Himmel et al. have found that not all nTregs express Helios as both $Helios^+$ and $Helios^-$ cells were found within nTregs. The authors also found that both $Helios^+$ and $Helios^-$ nTregs have a similar suppressive activity, indicating that Helios cannot be used as a distinct marker for nTregs (Himmel, et al. 2013).

In the last three years, transcriptomic profiling of Tregs was the key topic for deeply understanding the biology of Tregs to identify specific molecular pathways which control the development and differentiation of Tregs. Cuadrado et al. have performed proteomics and transcriptomics on human naïve and effector nTregs and $CD4^+CD25^+$ T cells. The authors have identified the mechanistic pathways that define the identity of both naïve and effector nTregs with the difference between them at the RNA and protein level. The authors have found that there was no correlation between the transcriptomic and proteomic data for the expression of identified proteins, confirming the importance of proteomics analysis for the identification of proteins that are involved in the biological pathways (Cuadrado, et al. 2018). Similar approach has been conducted by Mohammad et al. the authors analysed the difference between T helper 17 (Th17) and iTregs using transcriptomic and proteomic profiling approach. The authors have demonstrated that the changes in the protein expression were not associated with changes in the transcriptional level, highlighting the importance of proteomics analysis since proteins are the functional product of the gene and easily targetable (Mohammad, et al. 2018). Another recent study conducted by Schmidt et al. has identified 37 novel molecules that regulate the induction of Foxp3 during the differentiation of iTregs *in vitro* (Schmidt, et al. 2018). The results of all studies discussed above were only analysed mechanistic pathways at the transcriptional level. Also, none of previous studies have analysed or compared the differences between naïve $CD4^+CD25^-$ T cells, naïve $CD4^+CD25^+Foxp3^+$ nTregs, activated nTregs, activated $CD4^+CD25^-$ T cells and $CD4^+CD25^{++}Foxp3^+$ iTregs at the protein level, particularly at the level of membrane proteins since membrane proteins can be easily targeted using antibody based applications either for targeting or for viable isolation using flow cytometry.

In this study, the high throughput proteomic subcellular profiling was performed on five different purely sorted subsets of $CD4^+$ T cells including naïve $CD4^+CD25^-$ T cells, naïve $CD4^+CD25^+Foxp3^+$ nTregs, activated $CD4^+CD25^-$ T cells, activated nTregs and $CD4^+CD25^{++}Foxp3^+$ iTregs. Using SWATH-MS proteomics, 99 confident membrane proteins and 344 cytoplasmic and nuclear proteins were identified in the five different subsets of $CD4^+$ T cells. The confident proteins were defined based their OneOmics confidence score at ≥ 50% cut-off value. The fold change of the confident proteins was compared pairwise manner in five subsets using OneOmics software.

The functional association of significant proteins were elucidated using pathway analysis in MetaCore. Several pathways found to be differentially regulated in iTregs in comparison to nTregs and $CD25^-$ T cells indicating completely different molecular mechanisms operating in iTregs. Major pathways upregulated in iTregs in comparison to the nTregs are mainly associated with cell cycle progression. This is not surprising because iTregs are cultured for 5 days in a proliferative condition for the induction, whereas, the nTregs have been isolated *ex vivo* without any manipulations. For that reason, an appropriate comparison would be iTregs against nTreg depleted CD4 T cells cultured in the same way as iTregs but without TGF-beta. Pathway analysis of this comparison showed a complete different metabolic status for induced Tregs indicating a possible mechanism for converting CD4 into iTregs. The major patways down regulated in iTregs in comparison to cultured CD4s are glycolysis, HIF1 targets whereas the upregulated pathway is Ubiquinone metabolism. Studies have previously shown that inhibiting glycolysis prevented TH17 progression and supported Treg differentiation. More over same study also revealed that HIF1-α downregulation possibly by abolishing mTOR patway was critical in the development of Treg differentiation (Shi, et al. 2011). Our studies showed that presence of TGF-beta in the proliferating condition is good enough to convert naïve CD4 T cells in to iTreg possibly through the same mechanism.

An ANN based pattern recognition approach was employed to analyse the intensity of the expression of proteins in all of the subpopulations simultaneously in one run. The main advantage of using the ANN approach is due to its ability to recognise the pattern of variability, dimensionality and non-linearity of complex proteomics datasets and train the data by throughput stepwise modelling cycles with high performance to identify a panel of predictive biomarkers (Lancashire, Lemetre and Ball 2009).

For the population of iTregs, the results showed that 5 different novel membrane proteins were significantly up regulated in iTregs including ITIH4, PLP2, XYLT1, CUL4B and P85A. The overexpression of these proteins might be mediated via TGF-β1 signalling and could be linked to the suppressive activity of iTregs, while these proteins were detected at low levels in activated $CD4^+CD25^-$ T cells and other subsets of $CD4^+$ T cells analysed in this study. Activated $CD4^+CD25^-$ T cells and iTregs

were received the same strength of TCR activation signals via CD3/CD28 co-stimulatory signalling *in vitro* in the presence of IL-2. However, the only difference was that iTregs were generated in the presence of TGF-β1.

The results also showed that the expression of MAVS, HEM6 and SYAM proteins, as intracellular mitochondrial membrane proteins, was found significantly up regulated in iTregs compared to other subsets of $CD4^+$ T cells. These proteins also could be correlated to the suppressive activity of iTregs which was acquired due to exposure to TGF-β1 after TCR activation. The results suggest that these novel intracellular proteins could be used for defining the phenotypic features of iTregs. To date, the expression of CD73 and CTLA-4 ,as intracellular membrane proteins, has been used to detect to both nTregs and iTregs (Zhao, Liao and Kang 2017), however the expression of MAVS, HEM6 and SYAM proteins was specifically up regulated in iTregs. Therefore, these results suggest that these proteins could be a suitable target for detecting iTregs from other $CD4^+$ T cells subsets.

The results also showed that 20 distinct proteins, including GTR1, MOT4, NDRG1, ALDOA, PGK1, KPYM, LEG1, RIPK3, TPIS, G3P, SERC, SYYC, RP10, PGAM1, PGM1, GALK1, LDHA, RL34, TPP2, SYSC were found significantly overexpressed in activated $CD4^+CD25^-$ T cells and nTregs after TCR activation, while they were down expressed in iTregs. However, the expression of these proteins was detected at very lower levels in naïve subsets. When compared to naïve subsets, the overexpression of these proteins could be due to TCR activation, suggesting that these proteins are activation markers for $CD4^+$ T cells. However, the expression of these proteins was significantly decreased in iTregs subset even after TCR activation, suggesting that the expression of 20 different proteins could be affected by TFG-β1 cytokine during the differentiation of iTregs *in vitro*.

The phenotypic features of naïve nTregs are characterised by the expression of CD25 as a membrane protein and $Foxp3^+$ as a nuclear protein. However, CD25 protein is also upregulated upon TCR activation and considered as an activation marker for T cells (Bajnok, et al. 2017). Therefore, the expression of CD25 is not a good marker to distinguish naïve nTregs from other activated $CD4^+$ T cells. Moreover, the expression of Foxp3 is not a distinct marker for naïve nTregs as it is also expressed in both activated nTregs and iTregs. However, the results of this

study showed that the expression of HYEP, a membrane protein, was significantly up regulated in naïve nTregs only, while it was significantly down regulated in other subsets of $CD4^+$ T cells that were analysed in this study. The results also showed that the expression of SERPINB6A, an extracellular protein, was detected significantly up regulated in naïve nTreg, compared to other $CD4^+$ T cell subsets. These results suggest that HYEP and SERPINB6A can be novel surface markers to distinguish naïve nTregs from activated nTregs and iTregs. To distinguish activated nTregs from naïve and iTregs, the results showed that the expression of SPA3F, an extracellular protein, was found significantly up regulated in activated nTregs, compared to other subsets of $CD4^+$ T cells that were analysed in this study.

The identified proteins mentioned above were subjected for further validation on the subpopulations of $CD4^+$ T cells that were sorted and generated in the developed BALB/c mouse model. However, unfortunately, the validation experiment failed to show any positive staining for the identified proteins on the subpopulations of $CD4^+$ T cells that were used for validation. The possible reasons of failing include: (1) the quality of commercially-available antibodies that were used for the validation as majority of them were unconjugated polyclonal antibodies that have not been tested for flow cytometry analysis. The multi-reactivity of polyclonal antibodies in different species such as human, rat and mouse could affect their recognition and specific binding pattern for the protein of interest (Voskuil 2014). (2) failure in the conjugation experiment; since incomplete or insufficient fluorochrome-conjugation might prevent the antibody to be recognised by the lasers of flow cytometer. (3) the selected proteins might undergo posttranslational modifications which alter the structure and stability of proteins and contribute to the production of neo-epitopes which might not be recognised by the already designed antibody. Alternative methods such as western blotting or ELISA will be tested in the future using the same antibodies to confirm the results. Once validated in humans these markers/targets can be used to inhibit iTregs in tumours or any other inflammatory diseases. Also, all of the identified markers are on the cell surface will allow the user to purify the cells without fixing and permeabilising these cells. This will allow the scientist to understand the immunobiology of iTregs in more details.

All proteins identified in this study were searched in the literature using NCBI Pubmed and Google Scholar, however there were not any published studies found where the levels of these proteins were discussed in relevance to Tregs to confirm the results of this study. It is likely therefore, the lack of information in the literature confirms the novelty of the results of this study in identifying membrane markers that can discriminate iTregs from nTregs.

Chapter 5.

Summary of Discussion

5.1. Background and overall aim

Tregs are a distinct immunosuppressive subpopulation of $CD4^+$ T cells and function to control the immune homeostasis and tolerance by inhibiting the proliferation and function of autoreactive T cells, thereby preventing autoimmunity (Fehervari and Sakaguchi 2004). Tregs are broadly grouped into two main subpopulations based on the site of development, including thymic-derived natural Tregs (nTregs) and peripheral-derived induced Tregs (iTregs). Both nTregs and iTregs are characterised by the expression of CD25 and Foxp3. (de Lafaille, Maria A Curotto and Lafaille 2009).

In cancer, Tregs are found to enhance tumour progression by inhibiting the proliferation and function of $CD8^+$ T effector cells which induce anti-tumour immune response (Elkord, et al. 2010). The enrichment of $CD4^+CD25^+Foxp3^+$ Tregs within the tumour microenvironment has been correlated with worse prognosis and reduced survival rates in various types of cancer including breast cancer, lung cancer, ovarian cancer, glioblastoma, hepatocellular carcinoma, renal carcinoma, pancreatic ductal carcinoma, non-Hodgkin's lymphomas, melanoma and other malignancies (Zhang, et al. 2017, Curiel, et al. 2004, Tao, et al. 2012, Sun, et al. 2017, Sayour, et al. 2015, Tang, et al. 2014, Yang, et al. 2014, Yang, et al. 2006, Li, et al. 2010, Gerber, et al. 2014, Shang, et al. 2015a). The accumulation of Tregs within the tumour milieu is thought to be as a major hurdle for the development of effective immunotherapy (Liu, Workman and Vignali 2016c). Therefore, several attempts have been carried out to target Tregs within the tumour nest for enhancing anti-tumour immunity. Tregs have been targeted through developing antibodies against the CD25, CTLA-4, PD-1, CCR4 and CCR8. However, targeting Tregs via the expression of these biomarkers has resulted in the development of autoimmune disease as a side effect of non-selective treatment since these specific surface receptors are also expressed by Tregs circulating in the periphery where they maintain the peripheral immune tolerance and homeostasis (Bakacs, et al. 2012, Kurose, et al. 2015, Larkin, et al. 2015). In addition, CD25 and CTLA-4 are also

expressed by activated T effector cells, and thus targeting Tregs via these receptor could inhibit both Tregs and T effector cells, thereby promoting tumour progression indirectly (Ha, et al. 2019). Therefore, to selectively inhibit or block the function of tumour-infiltrating Tregs without affecting the function of circulating Tregs and other T effector cells, biomarkers that are exclusively expressed tumour-infiltrating Tregs are still required (Liu, et al. 2016a, Togashi, et al. 2019).

The question about the root of tumour-infiltrating Tregs is still debatable. It has been hypothesised that the origin of tumour-infiltrating Tregs could be from (1) nTregs which have been recruited into the tumour microenvironment via the expression of various chemokine receptors including CCR4, CCR5, CCR10 and CXCR3, (2) naive $CD4^+CD25^-$ T cells which have been recruited into the tumour microenvironment where they are converted into induced Tregs (iTregs) in the presence of TGF-β, or (3) from both nTregs and iTregs subpopulation (Chaudhary and Elkord 2016, Deng 2018). However, a novel study by Su et al. has proven that tumour-infiltrating Tregs within the tumour microenvironment of breast cancer are mainly developed from naïve CD4+ T cells that are recruited into the tumour bed, but not from circulating Tregs (Su, et al. 2017b). Therefore, to answer the question about the origin of tumour-infiltrating Tregs, it is still essential to identify biomarkers that can differentially distinguish iTregs that are generated from naïve $CD4^+CD25^-$ T cells from nTregs and other T cells. Identifying such biomarkers would facilitate the development of selective Tregs targeting therapy without affecting the function and homeostasis of other T effector cells.

Accordingly, the aim of this study was to identify membrane biomarkers that can differentially characterise phenotypic signature of iTregs from nTregs using SWATH-MS proteomics. To achieve this aim, the study was divided into 3 major objectives: (1) development of murine model for generating $CD4^+CD25^+Foxp3^+$ iTregs from naïve $CD4^+CD25^-Foxp3^-$ T cells in the presence and absence of tumour cells *in vitro*, (2) optimisation of a quantitative proteomics method to profile iTregs and nTregs , and (3) identification and verification of novel biomarkers of iTregs and nTregs.

5.2. The generation of iTregs from naïve $CD4^+CD25^-$ T cells *in vitro*

In this study, the BALB/c mouse strain was selected to develop a murine model for studying nTregs and iTregs since the percentage of $CD4^+CD25^+Foxp3^+$ nTregs in the spleen of BALB/c was significantly higher than in the spleen of C57BL/6 mice. These results have been supported by several previous studies (Chen, et al. 2005, Vogelsang, et al. 2009b, Blankenhaus, et al. 2014b). This higher percentage of nTregs was important to obtain a sufficient cell number for further downstream experiments.

For generating iTregs, naïve $CD4^+CD25^-$ T cells were purely purified and sorted using cell sorting. The purity of cell sorting was ≥ 98% and assessed by analysing the expression of CD25 on the sorted cells using flow cytometry. The sorted cells were grown *in vitro* in the presence of CD3/28 TCR activation beads, IL-2 (30 U/mL) and TGF-β1 (5 ng/mL) over a period of 5 days in the presence of tumour cells. The induction of iTregs from naïve $CD4^+CD25^-Foxp3^-$ T cells was confirmed by the positive expression of CD25 and Foxp3 on the generated iTregs using flow cytometry and defined as $CD4^+CD25^{++}$ $Fxop3^+$ iTregs.

5.2.1. The differentiation of iTregs requires a synergistic interaction of CD3/CD28 TCR and TGF-β1 signals in a time-dependent manner

The generation of iTregs in this model was mainly based on the activation of TCR in the presence of TGF-β1 and IL-2. In the absence of CD3/28 TCR activation signals, there was no induction of iTregs even in the presence of TGF-β1 and IL-2. However, the presence of both TCR activation signals, IL-2 and TGF-β1, the induction of iTregs was significantly increased, confirming the significance of the CD3/28 TCR activation signals in mediating the development of iTregs. These results were consistent with the finding that suboptimal TCR activation signals mediate the differentiation of iTregs from naïve $CD4^+CD25^-Foxp3^-$ T cells by promoting the expression of Foxp3 via NFAT (nuclear factor of activated cells) pathway (Vaeth, et al. 2012). Moreover, the stimulation of TCR signals has also been found to induce the

expression of Foxp3 in iTregs through the activation of NF-kB pathway in which c-Rel, a member of NF-kB family, promotes the expression of Foxp3 by binding to the Foxp3 promoter (Ruan, et al. 2009). The pivotal role of the TCR activation signals in the development of iTregs has also been confirmed by the finding that TCR-deficit Tregs lost their suppressive activity and failed to inhibit T cells response *in vivo* (Vahl, et al. 2014).

However, the stimulation of CD3/28- TCR signals solely was not enough to generate iTregs from naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells *in vitro* in the presence of IL-2 in the absence of TGF-β1, whereas the induction of iTregs was significantly increased in the presence of IL-2 and TGF-β1 following TCR activation, confirming the crucial role of TGF-β1 signalling in the differentiation of iTregs. The role of TGF-β1 signalling in the development of iTregs has been extensively acknowledged by many studies. Upon binding to its receptor, TGF-β1 activate its downstream transcription factor complex " Smad2/3" which is recruited into the nucleus where Smad2/3 binds the Foxp3 enhancer element, thereby inducing the expression of Foxp3 (Zheng, et al. 2010a, Tone, et al. 2008). The induction of Foxp3 has been also found t to be mediated by TGF-β receptor and IL-2 receptor through CNS2-5-azacytidine (5-aza-C) mechanism (Freudenberg, et al. 2018).

In this model, it was found that induction of iTregs from naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells was significantly decreased following continuous TCR engagement in the absence of TGF-β1 for 24 hours, whereas the induction of iTregs was significantly increased when the TGF-β1 was immediately added following TCR engagement. This finding confirms that naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells following constant TCR engagement become irresponsive to the TGF-β1 signalling, whereby preventing the differentiation into iTregs. These results also confirm that both TCR and TGF-β1 signals synergistically act to mediate the induction of iTregs from naïve $CD4^{+}CD25^{-}Foxp3^{-}$ T cells in a time-dependent manner. These results were consistent with the finding that that CD3/28 co-stimulation could overcome TGF-β-mediated repression of the proliferation of activated T effector cells, and after persistent CD3/28 co-stimulation, effector activated T cells become more resistant to the influence of TGF-β signalling (Koehler, et al. 2007). Moreover, it has been reported that activated T cells could reduce the expression of TGF-β receptor II (TGF-βRII) on

their surface following TCR activation via CD28 engagement, thereby avoiding the effect of TGF-β signalling (Sanjabi, et al. 2017).

5.2.2. The presence of 4T1 tumour cells significantly increases the development of iTregs in the presence of TGF-β1

In this model, the induction of iTregs from naïve $CD4^+CD25^-Fox3^-$ T cells was significantly increased in the presence of 4T1 tumour cells, TGF-β1 and IL-2, compared to those induced in the presence of TGF-β1 and IL-2 only. However, the induction of iTregs was at very low percentages (2%) in the presence of IL-2 and 4T1 cells but in the absence of TGF-β1, confirming that there is a cellular synergy between 4T1 cells and TGF-β1 signalling which in turn could enhance or increase the percentage of the iTregs induction. These findings suggest that the presence of TGF-β1 may influence 4T1 cells to initiate TGF-β1 signalling to produce TGF-β1 in an autocrine signalling loop-dependent manner (Daroqui, et al. 2012, Dumont, Bakin and Arteaga 2003), since 4T1 cells have been found to express functional TGF-β receptors (McEarchern, et al. 2001). In addition, the increase in the induction of iTregs was also correlated with the number of 4T1 cells co-cultured with naïve $CD4^+CD25^-Foxp3^-$ T cells in the presence of IL-2 and TGF-β1. The results revealed that the increase in the ratio of 4T1 cells to naïve $CD4^+CD25^-Foxp3^-$ T cells in the presence of TGF-β1 and IL-2 results in the increase of iTregs induction. Even at low concentrations of TGF-β1, particularly at 0.05 ng/mL, the induction of iTregs was significantly higher (20.6%) in the presence of 4T1 cells and IL-2, where the induction of iTregs at the same concentration of TGF-β1 in the presence of IL-2 only was 4.46%. These results confirm that the presence 4T1 tumour cells under the effect of TGF-β1 signalling promotes the induction of iTregs.

5.2.3. The generated iTregs are highly immunosuppressive

A functional assay was optimised to assess whether the generated iTregs are immunosuppressive and can inhibit the proliferation of $CD8^+$ T cells. To achieve this, naïve $CD8^+$ T cells were purely sorted, stained with CFSE (5 μM) and activated with CD3/28 T cell activator beads. The $CD8^+$ T cells were then co-cultured with the generated iTregs at different ratios. The overall results showed that the generated iTregs were highly immunosuppressive and could inhibit the proliferation of $CD8^+$

T cells. The percentage of the $CD8^+$ T cells inhibition by iTregs was based on the co-culture ratio of iTregs and $CD8^+$ T cells.

5.2.4. The Foxp3-TSDR of iTregs and stimulated nTregs is hypermethylated

The methylation status of the Foxp3-TSDR at four different CpG sites was assessed in the five different subpopulations of $CD4^+$ T cells. The naïve unstimulated nTregs were considered as a control population since their TSDR is completely demethylated (Schreiber, et al. 2014). The TSDR of the generated iTregs was significantly hypermethylated compared to naïve unstimulated nTregs. Similar results have been also found in a study carried out by Schmidt et al. where the authors assessed the methylation status of 15 different CpG regions within human Foxp3 gene locus and found that iTregs were significantly hypermethylated in all CpG sites, compared to unstimulated naïve nTregs (Schmidt, et al. 2016). It has been reported that the high methylation of the Foxp3-TSDR affects the stability and suppressive activity of iTregs (Huehn, Polansky and Hamann 2009, Zheng, Josefowicz, Chaudhry, Peng, Forbush and Rudensky 2010b, Baron, et al. 2007). However, the iTregs generated in this model were highly suppressive, could inhibit the proliferation of $CD8^+$ T cells and were stable over 5 days after induction. Also, other studies have generated highly suppressive human and murine stable iTregs in which the TSDR was hypermethylated (Hippen, et al. 2011, Lee, et al. 2012, Strainic, et al. 2013, Gu, et al. 2014). These results confirm that the TSDR methylation status is not correlated with the suppressive activity of iTregs, although the demethylation status of Foxp3-TSDR is considered as a marker for the stability of Foxp3 expression. The stability of the Foxp3 expression of iTregs has been discussed to be affected by the experimental protocols used for the generation of iTregs. For instance, iTregs that were generated using CD3/28 T cell activator beads have been found to acquire stable Foxp3 expression more than iTregs which were generated using plate-bound anti-CD3 and soluble anti-CD28 antibody, confirming the importance of the quality and affinity of TCR signals for maintaining the stability of Foxp3 expression (Geiger and Tauro 2012, Gottschalk, Corse and Allison 2010, Lu, et al. 2010, Hippen, et al. 2011, Qian, et al. 2011, Gu, et al. 2014). These findings were consistent with the results of the Foxp3 stability revealed in this study.

Intriguingly, in this study, it was found that naïve nTregs lost their demethylation status and became significantly hypermethylated following TCR activation. This finding was also supported by Zhang et al.; where the authors demonstrated that nTregs lost their Foxp3 expression after TCR engagement and experienced re-methylation of the CNS1 region within Foxp3 locus (Zhang, et al. 2017). A similar finding was also found in an earlier study by Bailey-Bucktrout et al. in 2013. There the authors demonstrated that nTregs underwent Foxp3 instability and secreted pro-inflammatory cytokines following self-antigen-driven activation during inflammation in the CNS (Bailey-Bucktrout, et al. 2013). Taken together, it seems that not only iTregs, but nTregs may lose their stability *in vivo* following TCR engagement. This may justify the notion that the acquired immunosuppressive activity of Tregs is inducible based on the presence of stressful stimuli that promote prolonged acute or chronic inflammation, which can make Tregs behave with more flexibility based on the physiological condition in the milieu.

5.3. Optimisation methods for protein isolation and preparation for quantitative proteomic profiling

The percentage of the induction of iTregs generated in this study allowed to obtain sufficient cell number of iTregs to be used for optimising different methods for quantitative proteomic profiling.

Prior to approaching proteomics, it was initially planned to carry out transcriptomic profiling for five different subsets of $CD4^+$ T cells to identify distinct biomarkers that can differentiate iTregs from nTregs and other control subpopulation of $CD4^+$ T cells. However, information derived from transcriptomic profiling reflects the level of RNAs abundance, not proteins. Therefore, the interpretation of transcriptomic data is difficult to predict the phenotypic feature of a cell or tissue (Misra, et al. 2018). In addition, several studies have applied a transcriptomic profiling integrated with proteomic profiling on different subsets of T cells and confirmed that the differential changes in the expression of biomarkers at RNA level were not correlated with the changes at the protein level (Cuadrado, et al. 2018, Mohammad, et al. 2018). For this reason (and that of cost and logistics), it was decided to carry out proteomic profiling that enable discovery biomarkers to distinguish iTregs from nTregs and other control subpopulation of $CD4^+$ T cells.

In this study, two different methods were optimised for isolation and preparation protein samples for proteomic analysis, including whole cell lysate and subcellular fractionation. The main aims of the optimisation were to isolate sufficient quantity of proteins from samples containing a low cell number ($2x10^6$), to identify a sufficient number of proteins yielded from MS (mass spectrometry) analysis and to increase the identification coverage of membrane proteins of total proteins identified by MS analysis.

Initially global (total cell lysate) proteomic profiling has been carried out using the method of whole cell lysate. Here, proteins were isolated from whole cell lysates derived from samples containing $2.5x10^6$ of purified murine T cells. 25µg of proteins were then processed for MS analysis. The total number of proteins identified from each sample was in range between 1053 and 1232 at 1% FDR (false discovery rate). The percentage of the identified membrane proteins was in range between 7.5% and 8.3% from total proteins, whereas the identified plasma membrane proteins constituted about 4.2-4.9% of total membrane proteins and more than 50% of total membrane proteins. These percentages were not sufficient to reach the aims of this study and confirm that the optimisation of proteins isolation using whole cell lysate lysed by Erika's or urea buffer failed to purify sufficient percentages of membrane proteins. Purification of plasma membrane proteins using whole cell lysate method was limited due to insolubility and hydrophobicity of plasma membrane proteins and low recovery rate of lipid-anchored surface proteins (Orsburn, et al. 2011).

To overcome the limitation of the whole cell lysate methods for identifying membrane proteins, the method of subcellular fractionation was optimised in this study to increase the coverage rate of quantification of plasma membrane proteins. Subcellular fractionation has also enabled identification and quantification of plasma membrane proteins which are the key target of antibody-based therapies (Leth-Larsen, et al. 2010). Moreover, Subcellular fractionation has enabled detection of peptides that belong to low-abundant proteins, thereby increasing the coverage rate of protein quantification using MS-based proteomics (Lee, et al. 2010).

In this study, the subcellular fractionation for isolation membrane and cytoplasmic proteins was carried out prior to proteomic analysis. The percentage of total known membrane proteins identified following subcellular fractionation was significantly increased to 57% of the total proteins identified by MS analysis, while it was only

around 8% using the whole cell lysate method. For plasma membrane proteins, the main of the percentage of identification was significantly increased to 26% of total proteins identified using subcellular fractionation method, whereas it was only 4.5% using the whole cell lysate method. These results were relatively reasonable according to the low cell number ($2x10^6$) which was used for isolation of membrane proteins. It was noticed that the kit does not completely separate membrane and cytosolic proteins and both fractions share a significant (10-30%) proportion of proteins which was not unexpected. Based on these results, the optimised method of the subcellular fractionation was followed to carry out a comprehensive quantitative proteomic profiling of the five different subsets of $CD4^+$ T cells using LC-MS-MS and SWATH-MS proteomics.

5.4. Identification and verification of distinct biomarkers of nTregs and iTregs using LC-MS-MS/SWATH-MS proteomics

3910 proteins were quantified from the five subpopulations of $CD4^+$ T cells by SWATH-MS analysis. The quantified proteins were sorted based on their confidence score as mentioned in chapter 4 (section 4.3.1.1). Accordingly, 99 membrane proteins and 344 cytoplasmic and nuclear proteins with confidence score ≥ 50% were selected for further analysis. The data of protein peak area of 99 membrane proteins and 344 cytoplasmic proteins were further analysed the ANN as a powerful bioinformatics tool to identify a panel of biomarkers that can differentiate iTregs from nTregs and other subpopulations of $CD4^+$ T cells. The ANN has been used as a machine learning approach to train and analyse complex proteomic data in order to identify and verify a combination of novel biomarkers (Swan, et al. 2013).

A panel of proteins, including PLP2, IHIT4, MAVS and HEM6, was identified to be significantly upregulated in the population of iTregs compared to the other subsets of $CD4^+$ T cells in which the proteins were detected at low levels. These proteins were also selected by the ANN as a predictive panel with the best predictive performance which can differentiate iTregs from other $CD4^+$ T cells subsets included the analysis. The sensitivity and specificity of these biomarkers were further tested

using a receiver operating characteristic (ROC) analysis which confirmed that the combination of PLP2, IHIT4, MAVS and HEM6 proteins could be as novel biomarkers of iTregs. The upregulation of these proteins in iTregs might be mediated by TGF-β1 signalling and could be associated with the suppressive activity of iTregs, since all other subpopulations were grown *in vitro* in the presence of CD3/28 T cell activator beads and IL-2, except iTregs were generated in the presence of CD3/28 beads, IL-2 and TGF-β1.

According to the Uniprot database, PLP2 (proteolipid protein 2) is classified as a multi-pass membrane protein which has been found predominantly localised on plasma membrane and endoplasmic reticulum membrane (www.uniprot.org, accessed on 25 Nov 2019). PLP2, which is also known as intestinal membrane A4 protein, has been initially found to be enriched in the differentiated colonic epithelial cells (Oliva, Wu and Yang 1993). PLP2 has been found to play a role in the differentiation of colonic epithelial cells in which, upon activation, PLP2 could localise to the endoplasmic reticulum and multimerise to form ion channels (Breitwieser, et al. 1997). PLP2 has been also found to mediate the migration of human osteogenic sarcoma (HOS) cells via the interaction with CCR1 expressed on the surface of HOS cells, indicating that PLP2 might act as a chemotactic ligand for chemokines receptors such as CCR1 (Lee, et al. 2004). In melanoma, PLP2 has been found to promote the progression and metastasis of tumour cells by inducing the secretion of MMP-2 (metalloproteinase-2) which serves as invasion-related molecule (Sonoda, et al. 2010). A study by Li et al. has shown that IL-10- and TGF-β-secreting Tregs promotes the invasiveness and proliferation of endometrial stromal cell derived from endometrial tumour cells by enhancing the expression of MMP-2 (Li, et al. 2014). A recent study by Chen et al. has shown that the upregulation of PLP2 in glioma tumour cells promotes the aggressiveness of tumour via the expression of MMP-2 and is associated with poor prognosis (Chen, Hueng and Tsai 2018). Based on these results, it seems that PLP2 can be as novel biomarker of iTregs that may promote the progression of tumour cells via the promotion of MMP-2 expression. There was no published work found confirming the expression of PLP2 in Tregs or any other T cells subsets, confirming the novelty of this discovery.

ITIH4 (Inter-alpha-trypsin inhibitor heavy chain H4) was also identified in this novel study to be significantly upregulated in iTregs compared to other subsets of

$CD4^+$ T cells. ITIH4 is a member of Inter-alpha-trypsin inhibitor family (ITIH1, ITIH2, ITIH3, ITIH4 and ITIH5), a family of plasma protease inhibitors. ITIH4 has been found to interact with hyaluronic acid (HA), which is of major molecules of the extracellular matrix, via forming a covalent bound (Hamm, et al. 2008). HA is a negatively-charged glycosaminoglycan which has been detected to be predominantly enriched on endothelial cells and in the extracellular matrix. HA has been found highly enriched in the tumour stroma of breast cancer. HA has also been found to mediate the recruitment of Tregs via the expression of CD44 and participate to enhance expression of Foxp3, thereby promoting the suppressive activity of Tregs. HA-binding Tregs have been found significantly enriched in the peripheral blood of patients with breast cancer (Perfilyeva, et al. 2015). HA has been found to function as a ligand with high affinity for the CD44 and HA-CD44 binding promotes the production of TGF-β1 and IL-10 by Tregs (Bollyky, et al. 2009) ITIH4 has been also found as an anti-inflammatory protein, and its expression was found highly correlated with the expression of IL-10 and IL-2 (Kashyap, et al. 2009). These finding could justify that ITIH4 might be a novel biomarker of iTregs that could recruited to the tumour microenvironment via the expression of HA on endothelial cells or in the extracellular matrix, since there was no any published work confirming the expression of ITIH4 on Tregs.

MAVS (mitochondrial antiviral signalling protein) was also identified to be significantly upregulated in iTregs compared to other subsets of $CD4^+$ T cells that were analysed in this study. MAVS has been found to mediate host innate immune responses against viral infection through the activation of the transcription factors NF-kB (nuclear factor kappa-light-chain-enhancer of activated B cells) and IRF3 (interferon regulatory factor 3). Both NF-kB and IRF3 have been found to regulate the expression of type-1 interferons (IFNs) including IFN-α and IFN-β which suppress viral replication (Seth, et al. 2005). As a signalling adapter protein, MAVS is activated by retinoic acid-inducible gene-I (RIG-I)-like receptors (RLRs) which recognise and bind viral RNA in the cytoplasm of infected cells (Hou, et al. 2011). The activation of MAVS stimulates the cytosolic kinases IKK and TBK1 which in turn induce the activation of NF-kB and IRF3 (McWhirter and Maniatis 2005).

RLRs have been found to be expressed by T cells including Tregs and T effector cells (Suthar, et al. 2012, Anz, et al. 2010). In Treg, the question whether MAVS signalling

has an impact on the function of Tregs has been investigated by Da Costa et al. where the authors found that MAVS signalling was not essential for the proliferation and suppressive activity of Tregs in MAVS-deficient mice following West Nile virus (WNV) infection (Da Costa, et al. 2017). However, the inhibition of MAVS signalling has been found to suppress the expansion of Tregs following WNV infection, and this was correlated with uncontrolled inflammatory response (Suthar, et al. 2010). A recent study by Kawano et al. has demonstrated that the overproduction of type 1 IFNs by myeloma cells could induce the recruitment and expansion of Tregs within the tumour microenvironment via IFN-α/β receptor 1 (IFNAR1) expressed on Tregs. The authors also found that the immunosuppressive of Tregs was significantly diminished following blocking IFNAR1 on Tregs (Kawano, et al. 2018). Therefore, based on the studies discussed above, the expansion of Tregs in response to the overproduction of type 1 IFNs might be regulated by MAVS signalling which induces the production of type IFNs via the activation of NF-kB and IRF3. However, the question whether the expression of IFNAR1 on Tregs can be regulated by MAVS signalling remains unanswered. The answer of this question might help in understanding the upregulation of MAVS in iTregs that was identified in this study.

HEM6, which is known as CPOX (Oxygen-dependent coproporphyrinogen-III oxidase, mitochondrial), was also identified to be significantly upregulated in iTregs compared to other subsets of $CD4^+$ T cells. According to Uniprot online database, the CPOX (HEM6) is a mitochondrial inner membrane protein which is implicated in the biosynthesis of heme (Uniprot online database, accessed on 25 Nov 2019). The role or function of this proteins in the biology of T cells was searched, however there was no any published study investigated its role.

Even though all proteins mentioned above are uniquely upregulated in iTregs, HYEP (epoxide hydrolase-1 (EPHX-1)) was found to be downregulated in iTregs compared nTregs and all other subsets of $CD4^+$ T cells that were analysed in this study. EPHX-1 is an enzyme localised in the endoplasmic reticulum membrane and involved in the metabolism of lipids (Uniprot online database, Accessed on 25 Nov 2019). There was no any published study investigated this protein in detail, therefore the role of HYEP in differentiating nTregs from iTregs should be further explored.

Validation of the biomarkers identified in this study using antibody-based flow cytometry assay was attempted. Very few antibodies were commercially available

for the identified biomarkers, and those that were, were not originally developed for flow cytometry and were unlabelled (unconjugated). Therefore, it was attempted to label/conjugate the antibodies with PE (Phycoerythrin) fluorochrome. The labelled antibodies were further tested on the subpopulations of $CD4^+$ T cells analysed in this model. However, the results showed that there was no positive staining for the antibodies on all the subpopulations, indicating that further work is needed to develop the identified panel of biomarkers for validation and testing in a large cohort of populations.

Conclusion and future work

Successfully established a murine iTreg induced model for studying its biology at proteome level and identification of markers which could distinguish iTregs from other CD4 subsets. As far as the literature search suggested, this is the first study carried out either in humans or mouse combining membrane fractionation with quantitative mass spectrometry. The ANN based analysis have identified five markers (PLP2, MAVS, HYEP, HEM6 and ITH4) as potential biomarkers. However, attempted validation of these markers using commercially available antibodies were not successful due to the antibody failure to detect the protein in flow cytometry. This could be negotiated in the future by either using a different technique such as western blot using the same antibody or developing a completely new antibody panel specifically suited for flow cytometry. Pathway enrichment analysis using literature curated database (MetaCore) and most differentially expressed proteins in iTreg compared to its control CD4 cells suggested a potential mechanistic role as supressed metabolic glycolytic pathway and HIF-target protein. Which agrees with the existing literature using other murine models. The role of TGF beta in metabolic alteration of iTregs were reported in the literature (Priyadharshini et al 2018). However, this study has identified precise molecular targets preferentially altered at protein level was identified. This study also has identified several target proteins involved in Ubiquinone metabolism (CoenzymeQ) up regulated in iTregs suggesting potential difference in mitochondrial activity. Further works needs to validate these findings which could not be completed within the time frame of this project. One of the potential drawbacks of this project is translation into human settings. Although, biomarkers and mechanistic insights

were gained using mouse model, its potential translation depend on how well these markers represent human iTreg counter parts, which was lacking at this moment due to the unavailability of suitable antibodies and human samples. This will be the priority in the immediate future. Overall this study achieved its objective in generating quantitative proteomics data using purified subsets of CD4 T cells and identified markers to distinguish iTreg, future validations needed to fully translate these findings into humans.

References

Abdel-Fatah, T.M., et al., 2016. SPAG5 as a prognostic biomarker and chemotherapy sensitivity predictor in breast cancer: a retrospective, integrated genomic, transcriptomic, and protein analysis. *The Lancet Oncology,* 17 (7), 1004-1018.

Abdolzade-Bavil, A., et al., 2004. Convenient and versatile subcellular extraction procedure, that facilitates classical protein expression profiling and functional protein analysis. *Proteomics,* 4 (5), 1397-1405.

Abul, K., 2003. Natural versus adaptive regulatory T cells. *Nat Rev Immunol,* 3 (3), 253-257.

Adotevi, O., et al., 2010. A decrease of regulatory T cells correlates with overall survival after sunitinib-based antiangiogenic therapy in metastatic renal cancer patients. *Journal of Immunotherapy,* 33 (9), 991-998.

Aebersold, R. and Mann, M., 2016. Mass-spectrometric exploration of proteome structure and function. *Nature,* 537 (7620), 347-355.

Akiyama, T., et al., 2008. The tumor necrosis factor family receptors RANK and CD40 cooperatively establish the thymic medullary microenvironment and self-tolerance. *Immunity,* 29 (3), 423-437.

Akoglu, B., et al., 2009. Interleukin-2 in CD8 T cells correlates with Banff score during organ rejection in liver transplant recipients. *Clinical and Experimental Medicine,* 9 (4), 259.

Aldinucci, D. and Colombatti, A., 2014. The inflammatory chemokine CCL5 and cancer progression. *Mediators of Inflammation,* 2014, 292376.

Aldiss, P., et al., 2019. Interscapular and Perivascular Brown Adipose Tissue Respond Differently to a Short-Term High-Fat Diet. *Nutrients,* 11 (5), 1065.

Aleksandrova, N., et al., 2016. Comparison of the results of preimplantation genetic screening obtained by a-CGH and NGS methods from the same embryos. *Gynecological Endocrinology,* 32 (sup2), 1-4.

Ali, I., et al., 2010. Separation of biological proteins by liquid chromatography. *Saudi Pharmaceutical Journal,* 18 (2), 59-73.

An, Y., et al., 2018. Transforming growth factor-β and peripheral regulatory cells are negatively correlated with the overall survival of hepatocellular carcinoma. *World Journal of Gastroenterology,* 24 (25), 2733.

Anand, S., et al., 2017. Label-based and label-free strategies for protein quantitation. *In:* Label-based and label-free strategies for protein quantitation. *Proteome Bioinformatics.* Springer, 2017, pp. 31-43.

Anderson, M.S., et al., 2002. Projection of an immunological self shadow within the thymus by the aire protein. *Science (New York, N.Y.),* 298 (5597), 1395-1401.

Anderson, N.L., 2010. The clinical plasma proteome: a survey of clinical assays for proteins in plasma and serum. *Clinical Chemistry,* 56 (2), 177-185.

Annes, J.P., Munger, J.S. and Rifkin, D.B., 2003. Making sense of latent TGFbeta activation. *Journal of Cell Science,* 116 (Pt 2), 217-224.

Anz, D., et al., 2010. Immunostimulatory RNA blocks suppression by regulatory T cells. *Journal of Immunology (Baltimore, Md.: 1950),* 184 (2), 939-946.

Arango Duque, G. and Descoteaux, A., 2014. Macrophage cytokines: involvement in immunity and infectious diseases. *Frontiers in Immunology,* 5, 491.

Araùjo, M., Hube, L.A. and Stasyk, T., 2008. Isolation of endocitic organelles by density gradient centrifugation. *In:* Isolation of endocitic organelles by density gradient centrifugation. *2D PAGE: Sample Preparation and Fractionation.* Springer, 2008, pp. 317-331.

Arike, L. and Peil, L., 2014. Spectral counting label-free proteomics. *In:* Spectral counting label-free proteomics. *Shotgun Proteomics.* Springer, 2014, pp. 213-222.

Arrivault, S., et al., 2009. Use of reverse-phase liquid chromatography, linked to tandem mass spectrometry, to profile the Calvin cycle and other metabolic intermediates in Arabidopsis rosettes at different carbon dioxide concentrations. *The Plant Journal,* 59 (5), 826-839.

Atarashi, K., et al., 2011. Induction of colonic regulatory T cells by indigenous Clostridium species. *Science (New York, N.Y.),* 331 (6015), 337-341.

Attia, P., et al., 2005. Inability of a fusion protein of IL-2 and diphtheria toxin (Denileukin Diftitox, DAB389IL-2, ONTAK) to eliminate regulatory T lymphocytes in patients with melanoma. *Journal of Immunotherapy (Hagerstown, Md.: 1997),* 28 (6), 582.

Aydin, S., 2015. A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA. *Peptides,* 72, 4-15.

Baba, T., Nakamoto, Y. and Mukaida, N., 2009. Crucial contribution of thymic Sirp alpha+ conventional dendritic cells to central tolerance against blood-borne antigens in a CCR2-dependent manner. *Journal of Immunology (Baltimore, Md.: 1950),* 183 (5), 3053-3063.
Baecher-Allan, C., Wolf, E. and Hafler, D.A., 2006. MHC class II expression identifies functionally distinct human regulatory T cells. *Journal of Immunology (Baltimore, Md.: 1950),* 176 (8), 4622-4631.
Baggerman, G., et al., 2005. Gel-based versus gel-free proteomics: a review. *Combinatorial Chemistry & High Throughput Screening,* 8 (8), 669-677.
Bailey, S.R., et al., 2014. Th17 cells in cancer: the ultimate identity crisis. *Frontiers in Immunology,* 5, 276.
Bailey-Bucktrout, S.L., et al., 2013. Self-antigen-driven activation induces instability of regulatory T cells during an inflammatory autoimmune response. *Immunity,* 39 (5), 949-962.
Bain, C., et al., 2017. TGFβR signalling controls CD103 CD11b dendritic cell development in the intestine. *Nature Communications,* 8 (1), 620.
Bajnok, A., et al., 2017. The distribution of activation markers and selectins on peripheral T lymphocytes in preeclampsia. *Mediators of Inflammation,* 2017.
Bakacs, T., Mehrishi, J.N. and Moss, R.W., 2012. Ipilimumab (Yervoy) and the TGN1412 catastrophe. *Immunobiology,* 217 (6), 583-589.
Banerjee, S. and Mazumdar, S., 2012. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International Journal of Analytical Chemistry,* 2012.
Bantscheff, M., et al., 2007. Quantitative mass spectrometry in proteomics: a critical review. *Analytical and Bioanalytical Chemistry,* 389 (4), 1017-1031.
Baron, U., et al., 2007. DNA demethylation in the human FOXP3 locus discriminates regulatory T cells from activated FOXP3 conventional T cells. *European Journal of Immunology,* 37 (9), 2378-2389.
Barton, G.M., 2008. A calculated response: control of inflammation by the innate immune system. *The Journal of Clinical Investigation,* 118 (2), 413-420.
Bateman, N.W., et al., 2014. Maximizing peptide identification events in proteomic workflows using data-dependent acquisition (DDA). *Molecular & Cellular Proteomics : MCP,* 13 (1), 329-338.
Baumgarth, N. and Roederer, M., 2000. A practical approach to multicolor flow cytometry for immunophenotyping. *Journal of Immunological Methods,* 243 (1-2), 77-97.
Beale, D.J., et al., 2018. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics,* 14 (11), 152.
Beatty, G.L. and Gladney, W.L., 2015. Immune escape mechanisms as a guide for cancer immunotherapy. *Clinical Cancer Research,* 21 (4), 687-692.
Behnes, C.L., et al., 2014. Tumor-associated macrophages are involved in tumor progression in papillary renal cell carcinoma. *Virchows Archiv,* 464 (2), 191-196.
Bellora, F., et al., 2010. The interaction of human natural killer cells with either unpolarized or polarized macrophages results in different functional outcomes. *Proceedings of the National Academy of Sciences of the United States of America,* 107 (50), 21659-21664.
Bending, D., et al., 2018. A timer for analyzing temporally dynamic changes in transcription during differentiation in vivo. *The Journal of Cell Biology,* 217 (8), 2931-2950.
Benencia, F., Muccioli, M. and Alnaeeli, M., 2014. Perspectives on reprograming cancer-associated dendritic cells for anti-tumor therapies. *Frontiers in Oncology,* 4, 72.
Benítez, F. and Najafian, N., 2008. Novel noninvasive assays to predict transplantation rejection and tolerance: enumeration of cytokine-producing alloreactive T cells. *Clinics in Laboratory Medicine,* 28 (3), 365-373.
Bernstein, C., et al., 2013. DNA damage, DNA repair and cancer. *In:* DNA damage, DNA repair and cancer. *New Research Directions in DNA Repair.* IntechOpen, 2013, .
Berthelot, J., et al., 2013. Regulatory B cells play a key role in immune system balance. *Joint Bone Spine,* 80 (1), 18-22.
Bettini, M.L. and Vignali, D.A., 2010. Development of thymically derived natural regulatory T cells. *Annals of the New York Academy of Sciences,* 1183, 1-12.
Beutler, B., 2004. Innate immunity: an overview. *Molecular Immunology,* 40 (12), 845-859.
Bhutia, S.K., Mallick, S.K. and Maiti, T.K., 2010. Tumour escape mechanisms and their therapeutic implications in combination tumour therapy. *Cell Biology International,* 34 (5), 553-563.
Bian, C., et al., 2015. Clinical outcome and expression of mutant P53, P16, and Smad4 in lung adenocarcinoma: a prospective study. *World Journal of Surgical Oncology,* 13 (1), 128.
Bian, J., et al., 2013. Mutation of TGF-β receptor II facilitates human bladder cancer progression through altered TGF-β1 signaling pathway. *International Journal of Oncology,* 43 (5), 1549-1559.

Bilbao, A., 2019. Proteomics Mass Spectrometry Data Analysis Tools.
Bilbao, A., Lisacek, F. and Hopfgartner, G., 2016. Dedicated software enhancing data-independent acquisition methods in mass spectrometry. *CHIMIA International Journal for Chemistry,* 70 (4), 293-293.
Bilbao, A., et al., 2015. Processing strategies and software solutions for data-independent acquisition in mass spectrometry. *Proteomics,* 15 (5-6), 964-980.
Blanco, P., et al., 2008. Dendritic cells and cytokines in human inflammatory and autoimmune diseases. *Cytokine & Growth Factor Reviews,* 19 (1), 41-52.
Blankenhaus, B., et al., 2014a. Foxp3 regulatory T cells delay expulsion of intestinal nematodes by suppression of IL-9-driven mast cell activation in BALB/c but not in C57BL/6 mice. *PLoS Pathogens,* 10 (2), e1003913.
Blankenhaus, B., et al., 2014b. Foxp3 regulatory T cells delay expulsion of intestinal nematodes by suppression of IL-9-driven mast cell activation in BALB/c but not in C57BL/6 mice. *PLoS Pathogens,* 10 (2), e1003913.
Bogdan, C., 2010. Regulation of lymphocytes by nitric oxide. *In:* Regulation of lymphocytes by nitric oxide. *Suppression and Regulation of Immune Responses.* Springer, 2010, pp. 375-393.
Boissel, N., et al., 2006. BCR/ABL oncogene directly controls MHC class I chain-related molecule A expression in chronic myelogenous leukemia. *Journal of Immunology (Baltimore, Md.: 1950),* 176 (8), 5108-5116.
Bollyky, P.L., et al., 2009. CD44 costimulation promotes FoxP3+ regulatory T cell persistence and function via production of IL-2, IL-10, and TGF-beta. *Journal of Immunology (Baltimore, Md.: 1950),* 183 (4), 2232-2241.
Bonilla, F.A. and Oettgen, H.C., 2010. Adaptive immunity. *Journal of Allergy and Clinical Immunology,* 125 (2), S33-S40.
Borcoman, E., et al., 2019. Inhibition of PI3K pathway increases immune infiltrate in muscle-invasive bladder cancer. *Oncoimmunology,* 8 (5), e1581556.
Borràs, E. and Sabidó, E., 2017. What is targeted proteomics? A concise revision of targeted acquisition and targeted data analysis in mass spectrometry. *Proteomics,* 17 (17-18), 1700180.
Boucard-Jourdin, M., et al., 2016. beta8 Integrin Expression and Activation of TGF-beta by Intestinal Dendritic Cells Are Determined by Both Tissue Microenvironment and Cell Lineage. *Journal of Immunology (Baltimore, Md.: 1950),* 197 (5), 1968-1978.
Bourmaud, A., Gallien, S. and Domon, B., 2016. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: principle and applications. *Proteomics,* 16 (15-16), 2146-2159.
Braun, R., et al., 2019. Single chromosome aneuploidy induces genome-wide perturbation of nuclear organization and gene expression. *Neoplasia,* 21 (4), 401-412.
Breitwieser, G.E., et al., 1997. Colonic epithelium-enriched protein A4 is a proteolipid that exhibits ion channel characteristics. *The American Journal of Physiology,* 272 (3 Pt 1), C957-65.
Brennan, P.J., Brigl, M. and Brenner, M.B., 2013. Invariant natural killer T cells: an innate activation scheme linked to diverse effector functions. *Nature Reviews Immunology,* 13 (2), 101.
Brilot, F., Strowig, T. and Munz, C., 2008. NK cells interactions with dendritic cells shape innate and adaptive immunity. *Frontiers in Bioscience : A Journal and Virtual Library,* 13, 6443-6454.
Broadhurst, D., et al., 2018. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics,* 14 (6), 72.
Bruder, D., et al., 2004. Frontline: Neuropilin-1: a surface marker of regulatory T cells. *European Journal of Immunology,* 34 (3), 623-630.
Bruno, A., et al., 2014. Orchestration of angiogenesis by immune cells. *Frontiers in Oncology,* 4, 131.
Bull, S.C. and Doig, A.J., 2015. Properties of protein drug target classes. *PloS One,* 10 (3), e0117955.
Burchill, M.A., et al., 2007. IL-2 receptor beta-dependent STAT5 activation is required for the development of Foxp3+ regulatory T cells. *Journal of Immunology (Baltimore, Md.: 1950),* 178 (1), 280-290.
Busch, S., et al., 2015. Loss of TGFbeta Receptor Type 2 Expression Impairs Estrogen Response and Confers Tamoxifen Resistance. *Cancer Research,* 75 (7), 1457-1469.
Bynoe, M.S. and Viret, C., 2008. Foxp3 CD4 T cell-mediated immunosuppression involves extracellular nucleotide catabolism. *Trends in Immunology,* 29 (3), 99-102.
Byrd, H.M. and McEwen, C.N., 2000. The limitations of MALDI-TOF mass spectrometry in the analysis of wide polydisperse polymers. *Analytical Chemistry,* 72 (19), 4568-4576.
Campoli, M. and Ferrone, S., 2008. HLA antigen changes in malignant cells: epigenetic mechanisms and biologic significance. *Oncogene,* 27 (45), 5869.

Cao, X., et al., 2007. Granzyme B and perforin are important for regulatory T cell-mediated suppression of tumor clearance. *Immunity,* 27 (4), 635-646.
Cao, Z., et al., 2008. Use of fluorescence-activated vesicle sorting for isolation of Naked2-associated, basolaterally targeted exocytic vesicles for proteomics analysis. *Molecular & Cellular Proteomics : MCP,* 7 (9), 1651-1667.
Carmeliet, P. and Jain, R.K., 2000. Angiogenesis in cancer and other diseases. *Nature,* 407 (6801), 249.
Carnevalli, L.S., et al., 2018. PI3Kα/δ inhibition promotes anti-tumor immunity through direct enhancement of effector CD8 T-cell activity. *Journal for Immunotherapy of Cancer,* 6 (1), 158.
Carvalho, P.C., et al., 2010. XDIA: improving on the label-free data-independent analysis. *Bioinformatics,* 26 (6), 847-848.
Casey, S.C., Bellovin, D.I. and Felsher, D.W., 2013. Noncanonical roles of the immune system in eliciting oncogene addiction. *Current Opinion in Immunology,* 25 (2), 246-258.
Casey, S.C., et al., 2016. MYC regulates the antitumor immune response through CD47 and PD-L1. *Science (New York, N.Y.),* 352 (6282), 227-231.
Castagnola, M., et al., 2012. The human salivary proteome: a critical overview of the results obtained by different proteomic platforms. *Expert Review of Proteomics,* 9 (1), 33-46.
Castañeda-Patlán, M.C., Fuentes-García, G. and Robles-Flores, M., 2018. Wnt Signaling as a Master Regulator of Immune Tolerance in a Tumor Microenvironment. *In:* Wnt Signaling as a Master Regulator of Immune Tolerance in a Tumor Microenvironment. *Cell Signalling-Thermodynamics and Molecular Control.* IntechOpen, 2018, .
Catherman, A.D., Skinner, O.S. and Kelleher, N.L., 2014. Top down proteomics: facts and perspectives. *Biochemical and Biophysical Research Communications,* 445 (4), 683-693.
Chahrour, O., Cobice, D. and Malone, J., 2015. Stable isotope labelling methods in mass spectrometry-based quantitative proteomics. *Journal of Pharmaceutical and Biomedical Analysis,* 113, 2-20.
Chait, B.T., 2006. Chemistry. Mass spectrometry: bottom-up or top-down? *Science (New York, N.Y.),* 314 (5796), 65-66.
Chang, S.H. and Dong, C., 2007. A novel heterodimeric cytokine consisting of IL-17 and IL-17F regulates inflammatory responses. *Cell Research,* 17 (5), 435.
Chang, Y.W., et al., 2015. Diverse Targets of beta-Catenin during the Epithelial-Mesenchymal Transition Define Cancer Stem Cells and Predict Disease Relapse. *Cancer Research,* 75 (16), 3398-3410.
Chanmee, T., et al., 2014. Tumor-associated macrophages as major players in the tumor microenvironment. *Cancers,* 6 (3), 1670-1690.
Chapman, J.D., Goodlett, D.R. and Masselon, C.D., 2014. Multiplexed and data-independent tandem mass spectrometry for global proteome profiling. *Mass Spectrometry Reviews,* 33 (6), 452-470.
Chaudhary, B. and Elkord, E., 2016. Regulatory T cells in the tumor microenvironment and cancer progression: role and therapeutic targeting. *Vaccines,* 4 (3), 28.
Chauvin, A. and Boisvert, F., 2018. Clinical proteomics in colorectal cancer, a promising tool for improving personalised medicine. *Proteomes,* 6 (4), 49.
Chen, L. and Flies, D.B., 2013. Molecular mechanisms of T cell co-stimulation and co-inhibition. *Nature Reviews Immunology,* 13 (4), 227.
Chen, T., et al., 2001a. Novel inactivating mutations of transforming growth factor-β type I receptor gene in head-and-neck cancer metastases. *International Journal of Cancer,* 93 (5), 653-661.
Chen, X., Oppenheim, J.J. and Howard, O.Z., 2005. BALB/c mice have more CD4 CD25 T regulatory cells and show greater susceptibility to suppression of their CD4 CD25− responder T cells than C57BL/6 mice. *Journal of Leukocyte Biology,* 78 (1), 114-121.
Chen, X., et al., 2015. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics,* 15 (18), 3175-3192.
Chen, Y., 2009. Endocytic regulation of TGF-β signaling. *Cell Research,* 19 (1), 58-70.
Chen, Y., Hueng, D. and Tsai, W., 2018. Proteolipid Protein 2 Overexpression Indicates Aggressive Tumor Behavior and Adverse Prognosis in Human Gliomas. *International Journal of Molecular Sciences,* 19 (11), 3353.
Chen, C.H., et al., 2003. Transforming growth factor beta blocks Tec kinase phosphorylation, Ca2+ influx, and NFATc translocation causing inhibition of T cell differentiation. *The Journal of Experimental Medicine,* 197 (12), 1689-1699.
Chen, T., et al., 2001b. Transforming growth factor-beta receptor type I gene is frequently mutated in ovarian carcinomas. *Cancer Research,* 61 (12), 4679-4682.
Cheroutre, H., Lambolez, F. and Mucida, D., 2011. The light and dark sides of intestinal intraepithelial lymphocytes. *Nature Reviews Immunology,* 11 (7), 445.

Chin, D., et al., 2004. What is transforming growth factor-beta (TGF-β)? *British Journal of Plastic Surgery,* 57 (3), 215-221.
Chow, M.T., Möller, A. and Smyth, M.J., 2012. Inflammation and immune surveillance in cancer. *In: Seminars in cancer biology,* Elsevier, pp. 23-32.
Chung, A.S., et al., 2013. An interleukin-17–mediated paracrine network promotes tumor resistance to anti-angiogenic therapy. *Nature Medicine,* 19 (9), 1114.
Collins, B.C., et al., 2017. Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. *Nature Communications,* 8 (1), 291.
Coombes, J.L., et al., 2007. A functionally specialized population of mucosal CD103+ DCs induces Foxp3+ regulatory T cells via a TGF-beta and retinoic acid-dependent mechanism. *The Journal of Experimental Medicine,* 204 (8), 1757-1764.
Correa, P. and Piazuelo, M.B., 2011. Helicobacter pylori Infection and Gastric Adenocarcinoma. *US Gastroenterology & Hepatology Review,* 7 (1), 59-64.
Corzo, C.A., et al., 2009. Mechanism regulating reactive oxygen species in tumor-induced myeloid-derived suppressor cells. *Journal of Immunology (Baltimore, Md.: 1950),* 182 (9), 5693-5701.
Cottrez, F. and Groux, H., 2001. Regulation of TGF-beta response during T cell activation is modulated by IL-10. *Journal of Immunology (Baltimore, Md.: 1950),* 167 (2), 773-778.
Coulie, P.G., et al., 2014. Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. *Nature Reviews Cancer,* 14 (2), 135.
Cowan, J.E., et al., 2018. Aire controls the recirculation of murine Foxp3 regulatory T-cells back to the thymus. *European Journal of Immunology,* 48 (5), 844-854.
Cowan, J.E., McCarthy, N.I. and Anderson, G., 2016. CCR7 controls thymus recirculation, but not production and emigration, of Foxp3 T cells. *Cell Reports,* 14 (5), 1041-1048.
Cowan, J.E., et al., 2013. The thymic medulla is required for Foxp3+ regulatory but not conventional CD4+ thymocyte development. *The Journal of Experimental Medicine,* 210 (4), 675-681.
Cox, K.L., et al., 2004. Immunoassay Methods. *In:* G.S. Sittampalam, et al., ed., *Assay Guidance Manual.* Bethesda (MD): 2004, .
Cristobal, A., et al., 2017. Toward an optimized workflow for middle-down proteomics. *Analytical Chemistry,* 89 (6), 3318-3325.
Croci, D.O., et al., 2007. Dynamic cross-talk between tumor and immune cells in orchestrating the immunosuppressive network at the tumor microenvironment. *Cancer Immunology, Immunotherapy,* 56 (11), 1687-1700.
Cuadrado, E., et al., 2018. Proteomic analyses of human regulatory T cells reveal adaptations in signaling pathways that protect cellular identity. *Immunity,* 48 (5), 1046-1059. e6.
Cui, J., et al., 2013. Accurate LC peak boundary detection for 16O/18O labeled LC-MS data. *PloS One,* 8 (10), e72951.
Curiel, T.J., et al., 2004. Specific recruitment of regulatory T cells in ovarian carcinoma fosters immune privilege and predicts reduced survival. *Nature Medicine,* 10 (9), 942.
Cutillas, P.R. and Vanhaesebroeck, B., 2007. Quantitative profile of five murine core proteomes using label-free functional proteomics. *Molecular & Cellular Proteomics : MCP,* 6 (9), 1560-1573.
Da Costa, A., et al., 2017. Extrinsic MAVS signaling is critical for Treg maintenance of Foxp3 expression following acute flavivirus infection. *Scientific Reports,* 7, 40720.
da Silva, F.B., et al., 2017. Single-nucleotide polymorphism array (SNP-A) improves the identification of chromosomal abnormalities by metaphase cytogenetics in myelodysplastic syndrome. *Journal of Clinical Pathology,* 70 (5), 435-442.
Dalili, N., et al., 2019. Urine and serum NMR-based metabolomics in pre-procedural prediction of contrast-induced nephropathy. *Internal and Emergency Medicine,* , 1-9.
Dalod, M., et al., 2014. Dendritic cell maturation: functional specialization through signaling specificity and transcriptional programming. *The EMBO Journal,* 33 (10), 1104-1116.
Dardalhon, V., et al., 2008. IL-4 inhibits TGF-β-induced Foxp3 T cells and, together with TGF-β, generates IL-9 IL-10 Foxp3− effector T cells. *Nature Immunology,* 9 (12), 1347.
Daroqui, M.C., et al., 2012. TGF-β autocrine pathway and MAPK signaling promote cell invasiveness and in vivo mammary adenocarcinoma tumor progression. *Oncology Reports,* 28 (2), 567-575.
Darrasse-Jeze, G., et al., 2009. Feedback control of regulatory T cell homeostasis by dendritic cells in vivo. *The Journal of Experimental Medicine,* 206 (9), 1853-1862.
Das, J., et al., 2009. Transforming growth factor beta is dispensable for the molecular orchestration of Th17 cell differentiation. *The Journal of Experimental Medicine,* 206 (11), 2407-2416.
Dawson, P.H., 2013. *Quadrupole mass spectrometry and its applications.* Elsevier.
De Kleer, I., et al., 2014. Ontogeny of myeloid cells. *Frontiers in Immunology,* 5, 423.

De Kruijf, E., et al., 2012. The prognostic role of TGF-β signaling pathway in breast cancer patients. *Annals of Oncology,* 24 (2), 384-390.
de Lafaille, Maria A Curotto and Lafaille, J.J., 2009. Natural and adaptive foxp3 regulatory T cells: more of the same or a division of labor? *Immunity,* 30 (5), 626-635.
Deaglio, S., et al., 2007. Adenosine generation catalyzed by CD39 and CD73 expressed on regulatory T cells mediates immune suppression. *The Journal of Experimental Medicine,* 204 (6), 1257-1265.
Deftos, M.L. and Bevan, M.J., 2000. Notch signaling in T cell development. *Current Opinion in Immunology,* 12 (2), 166-172.
del Campo, M., et al., 2015. Facilitating the validation of novel protein biomarkers for dementia: an optimal workflow for the development of sandwich immunoassays. *Frontiers in Neurology,* 6, 202.
Del Rio, M., et al., 2010. Development and functional specialization of CD103 dendritic cells. *Immunological Reviews,* 234 (1), 268-281.
Demelbauer, U.M., et al., 2004. Determination of glycopeptide structures by multistage mass spectrometry with low-energy collision-induced dissociation: comparison of electrospray ionization quadrupole ion trap and matrix-assisted laser desorption/ionization quadrupole ion trap reflectron time-of-flight approaches. *Rapid Communications in Mass Spectrometry,* 18 (14), 1575-1582.
Demeure, K., Gabelica, V. and De Pauw, E.A., 2010. New advances in the understanding of the in-source decay fragmentation of peptides in MALDI-TOF-MS. *Journal of the American Society for Mass Spectrometry,* 21 (11), 1906-1917.
Deng, G., 2018. Tumor-infiltrating regulatory T cells: origins and features. *American Journal of Clinical and Experimental Immunology,* 7 (5), 81.
Derynck, R. and Zhang, Y.E., 2003. Smad-dependent and Smad-independent pathways in TGF-β family signalling. *Nature,* 425 (6958), 577.
Desar, I.M., et al., 2011. Sorafenib reduces the percentage of tumour infiltrating regulatory T cells in renal cell carcinoma patients. *International Journal of Cancer,* 129 (2), 507-512.
Desruisseau, S., et al., 2006. Determination of TGFβ1 protein level in human primary breast cancers and its relationship with survival. *British Journal of Cancer,* 94 (2), 239.
DeVita, H., 2015. Rosenberg's cancer: principles and practice of oncology. TV DeVita Jr, TS Lawrence, SA Rosenberg (eds.).
Dilek, N., et al., 2012. Myeloid-derived suppressor cells: mechanisms of action and recent advances in their role in transplant tolerance. *Frontiers in Immunology,* 3, 208.
Ding, M., et al., 2016. Association between transforming growth factor-β1 expression and the clinical features of triple negative breast cancer. *Oncology Letters,* 11 (6), 4040-4044.
Dobbin, K.K., et al., 2016. Validation of biomarkers to predict response to immunotherapy in cancer: Volume II—clinical validation and regulatory considerations. *Journal for Immunotherapy of Cancer,* 4 (1), 77.
Doedens, A.L., et al., 2016. Molecular Programming of Tumor-Infiltrating CD8+ T Cells and IL15 Resistance. *Cancer Immunology Research,* 4 (9), 799-811.
Doisne, J.M., et al., 2009. iNKT cell development is orchestrated by different branches of TGF-beta signaling. *The Journal of Experimental Medicine,* 206 (6), 1365-1378.
Domon, B. and Aebersold, R., 2010. Options and considerations when selecting a quantitative proteomics strategy. *Nature Biotechnology,* 28 (7), 710.
Domon, B. and Gallien, S., 2015. Recent advances in targeted proteomics for clinical applications. *PROTEOMICS–Clinical Applications,* 9 (3-4), 423-431.
Dong, C., 2008. T H 17 cells in development: an updated view of their molecular identity and genetic programming. *Nature Reviews Immunology,* 8 (5), 337.
Drabsch, Y. and Ten Dijke, P., 2012. TGF-β signalling and its role in cancer progression and metastasis. *Cancer and Metastasis Reviews,* 31 (3-4), 553-568.
Dreger, M., 2003. Subcellular proteomics. *Mass Spectrometry Reviews,* 22 (1), 27-56.
Duguet, F., et al., 2017. Proteomic Analysis of Regulatory T Cells Reveals the Importance of Themis1 in the Control of Their Suppressive Function. *Molecular & Cellular Proteomics : MCP,* 16 (8), 1416-1432.
Dumont, N., Bakin, A.V. and Arteaga, C.L., 2003. Autocrine transforming growth factor-beta signaling mediates Smad-independent motility in human cancer cells. *The Journal of Biological Chemistry,* 278 (5), 3275-3285.
Dunn, G.P., Old, L.J. and Schreiber, R.D., 2004. The immunobiology of cancer immunosurveillance and immunoediting. *Immunity,* 21 (2), 137-148.

Ebhardt, H.A., 2014. Selected reaction monitoring mass spectrometry: a methodology overview. *In:* Selected reaction monitoring mass spectrometry: a methodology overview. *Plant Proteomics.* Springer, 2014, pp. 209-222.
El-Aneed, A., Cohen, A. and Banoub, J., 2009. Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews,* 44 (3), 210-230.
Elkord, E., et al., 2010. T regulatory cells in cancer: recent advances and therapeutic potential. *Expert Opinion on Biological Therapy,* 10 (11), 1573-1586.
Elo, L.L., et al., 2010. Genome-wide profiling of interleukin-4 and STAT6 transcription factor regulation of human Th2 cell programming. *Immunity,* 32 (6), 852-862.
Elpek, G.O., et al., 2001. The prognostic relevance of angiogenesis and mast cells in squamous cell carcinoma of the oesophagus. *Journal of Clinical Pathology,* 54 (12), 940-944.
Engelhardt, J.J., et al., 2012. Marginating dendritic cells of the tumor microenvironment cross-present tumor antigens and stably engage tumor-specific T cells. *Cancer Cell,* 21 (3), 402-417.
Evans, C., et al., 2012. An insight into iTRAQ: where do we stand now? *Analytical and Bioanalytical Chemistry,* 404 (4), 1011-1027.
Facciabene, A., et al., 2011. Tumour hypoxia promotes tolerance and angiogenesis via CCL28 and T reg cells. *Nature,* 475 (7355), 226.
Facciabene, A., Motz, G.T. and Coukos, G., 2012. T-regulatory cells: key players in tumor immune escape and angiogenesis. *Cancer Research,* 72 (9), 2162-2171.
Fallarino, F., et al., 2003. Modulation of tryptophan catabolism by regulatory T cells. *Nature Immunology,* 4 (12), 1206.
Falschlehner, C., et al., 2009. TRAIL and other TRAIL receptor agonists as novel cancer therapeutics. *In:* TRAIL and other TRAIL receptor agonists as novel cancer therapeutics. *Therapeutic Targets of the TNF Superfamily.* Springer, 2009, pp. 195-206.
Fan, X., et al., 2018. CD49b defines functionally mature Treg cells that survey skin and vascular tissues. *The Journal of Experimental Medicine,* 215 (11), 2796-2814.
Fang, X. and Zhang, W., 2008. Affinity separation and enrichment methods in proteomic analysis. *Journal of Proteomics,* 71 (3), 284-303.
Faria, S.S., et al., 2017. A timely shift from shotgun to targeted proteomics and how it can be groundbreaking for cancer research. *Frontiers in Oncology,* 7, 13.
Fehervari, Z. and Sakaguchi, S., 2004. CD4+ Tregs and immune control. *The Journal of Clinical Investigation,* 114 (9), 1209-1217.
Feist, P. and Hummon, A., 2015. Proteomic challenges: sample preparation techniques for microgram-quantity protein analysis from biological samples. *International Journal of Molecular Sciences,* 16 (2), 3537-3563.
Felsher, D.W., 2010. MYC inactivation elicits oncogene addiction through both tumor cell–intrinsic and host-dependent mechanisms. *Genes & Cancer,* 1 (6), 597-604.
Feng, S., et al., 2018. Myeloid-derived suppressor cells inhibit T cell activation through nitrating LCK in mouse cancers. *Proceedings of the National Academy of Sciences of the United States of America,* 115 (40), 10094-10099.
Fiehn, O., 2002. Metabolomics—the link between genotypes and phenotypes. *In:* Metabolomics—the link between genotypes and phenotypes. *Functional genomics.* Springer, 2002, pp. 155-171.
Fiehn, O., et al., 2007. The metabolomics standards initiative (MSI). *Metabolomics,* 3 (3), 175-178.
Floess, S., et al., 2007. Epigenetic control of the foxp3 locus in regulatory T cells. *PLoS Biology,* 5 (2), e38.
Fonslow, B.R. and Yates Iii, J.R., 2009. Capillary electrophoresis applied to proteomic analysis. *Journal of Separation Science,* 32 (8), 1175-1188.
Freudenberg, K., et al., 2018. critical role of TgF-β and il-2 receptor signaling in Foxp3 induction by an inhibitor of Dna Methylation. *Frontiers in Immunology,* 9, 125.
Fuertes, M.B., et al., 2011. Host type I IFN signals are required for antitumor CD8+ T cell responses through CD8{alpha}+ dendritic cells. *The Journal of Experimental Medicine,* 208 (10), 2005-2016.
Fukai, Y., et al., 2003. Reduced expression of transforming growth factor-β receptors is an unfavorable prognostic factor in human esophageal squamous cell carcinoma. *International Journal of Cancer,* 104 (2), 161-166.
Furini, G., et al., 2018. Proteomic Profiling Reveals the Transglutaminase-2 Externalization Pathway in Kidneys after Unilateral Ureteric Obstruction. *Journal of the American Society of Nephrology : JASN,* 29 (3), 880-905.
Gabay, M., Li, Y. and Felsher, D.W., 2014. MYC activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harbor Perspectives in Medicine,* 4 (6), a014241.

Gabrilovich, D.I. and Nagaraj, S., 2009. Myeloid-derived suppressor cells as regulators of the immune system. *Nature Reviews Immunology,* 9 (3), 162.
Gajewski, T.F., Schreiber, H. and Fu, Y., 2013. Innate and adaptive immune cells in the tumor microenvironment. *Nature Immunology,* 14 (10), 1014.
Galli, S.J., Grimbaldeston, M. and Tsai, M., 2008. Immunomodulatory mast cells: negative, as well as positive, regulators of immunity. *Nature Reviews Immunology,* 8 (6), 478.
Galli, S.J., Nakae, S. and Tsai, M., 2005. Mast cells in the development of adaptive immune responses. *Nature Immunology,* 6 (2), 135.
Gallien, S., Duriez, E. and Domon, B., 2011. Selected reaction monitoring applied to proteomics. *Journal of Mass Spectrometry,* 46 (3), 298-312.
Gallo, R.L. and Nizet, V., 2008. Innate barriers against skin infection and associated disorders. *Drug Discovery Today: Disease Mechanisms,* 5 (2), e145-e152.
Galluzzi, L., et al., 2012. The secret ally: immunostimulation by anticancer drugs. *Nature Reviews Drug Discovery,* 11 (3), 215.
Gambineri, E., Torgerson, T.R. and Ochs, H.D., 2003. Immune dysregulation, polyendocrinopathy, enteropathy, and X-linked inheritance (IPEX), a syndrome of systemic autoimmunity caused by mutations of FOXP3, a critical regulator of T-cell homeostasis. *Current Opinion in Rheumatology,* 15 (4), 430-435.
Garcia, B.A., 2010. What does the future hold for Top Down mass spectrometry? *Journal of the American Society for Mass Spectrometry,* 21 (2), 193-202.
Garin, M.I., et al., 2007. Galectin-1: a key effector of regulation mediated by CD4+CD25+ T cells. *Blood,* 109 (5), 2058-2065.
Gattinoni, L., Ji, Y. and Restifo, N.P., 2010. Wnt/beta-catenin signaling in T-cell immunity and cancer immunotherapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 16 (19), 4695-4701.
Ge, Y., et al., 2012. Metronomic cyclophosphamide treatment in metastasized breast cancer patients: immunological effects and clinical outcome. *Cancer Immunology, Immunotherapy,* 61 (3), 353-362.
Geiger, T., et al., 2011. Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nature Protocols,* 6 (2), 147.
Geiger, T.L. and Tauro, S., 2012. Nature and nurture in Foxp3 regulatory T cell development, stability, and function. *Human Immunology,* 73 (3), 232-239.
Gerber, A.L., et al., 2014. High expression of FOXP 3 in primary melanoma is associated with tumour progression. *British Journal of Dermatology,* 170 (1), 103-109.
Germain, R.N., 2002. T-cell development and the CD4–CD8 lineage decision. *Nature Reviews Immunology,* 2 (5), 309.
Ghiringhelli, F., et al., 2005. Tumor cells convert immature myeloid dendritic cells into TGF-beta-secreting cells inducing CD4+CD25+ regulatory T cell proliferation. *The Journal of Experimental Medicine,* 202 (7), 919-929.
Ghoreschi, K., et al., 2010. Generation of pathogenic T H 17 cells in the absence of TGF-β signalling. *Nature,* 467 (7318), 967.
Gillet, L.C., et al., 2012. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Molecular & Cellular Proteomics : MCP,* 11 (6), O111.016717.
Glimelius, I., et al., 2005. Angiogenesis and mast cells in Hodgkin lymphoma. *Leukemia,* 19 (12), 2360.
Gobert, M., et al., 2009. Regulatory T cells recruited through CCL22/CCR4 are selectively activated in lymphoid infiltrates surrounding primary breast tumors and lead to an adverse clinical outcome. *Cancer Research,* 69 (5), 2000-2009.
Goodacre, R., et al., 2004. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology,* 22 (5), 245-252.
Gordon, J. and Manley, N.R., 2011. Mechanisms of thymus organogenesis and morphogenesis. *Development (Cambridge, England),* 138 (18), 3865-3878.
Gorelik, L., Constant, S. and Flavell, R.A., 2002. Mechanism of transforming growth factor beta-induced inhibition of T helper type 1 differentiation. *The Journal of Experimental Medicine,* 195 (11), 1499-1505.
Gorelik, L., Fields, P.E. and Flavell, R.A., 2000. Cutting edge: TGF-beta inhibits Th type 2 development through inhibition of GATA-3 expression. *Journal of Immunology (Baltimore, Md.: 1950),* 165 (9), 4773-4777.
Gottschalk, R.A., Corse, E. and Allison, J.P., 2010. TCR ligand density and affinity determine peripheral induction of Foxp3 in vivo. *The Journal of Experimental Medicine,* 207 (8), 1701-1711.

Grade, M., Difilippantonio, M.J. and Camps, J., 2015. Patterns of chromosomal aberrations in solid tumors. *In:* Patterns of chromosomal aberrations in solid tumors. *Chromosomal instability in Cancer cells.* Springer, 2015, pp. 115-142.
Graham, J.M., 2001. Isolation of Golgi membranes from tissues and cells by differential and density gradient centrifugation. *Current Protocols in Cell Biology,* 10 (1), 3.9. 1-3.9. 24.
Grivennikov, S.I., Greten, F.R. and Karin, M., 2010. Immunity, inflammation, and cancer. *Cell,* 140 (6), 883-899.
Groh, V., et al., 2002. Tumour-derived soluble MIC ligands impair expression of NKG2D and T-cell activation. *Nature,* 419 (6908), 734.
Grossman, W.J., et al., 2004. Differential expression of granzymes A and B in human cytotoxic lymphocyte subsets and T regulatory cells. *Blood,* 104 (9), 2840-2848.
Gu, J., et al., 2014. TGF-beta-induced CD4+Foxp3+ T cells attenuate acute graft-versus-host disease by suppressing expansion and killing of effector CD8+ cells. *Journal of Immunology (Baltimore, Md.: 1950),* 193 (7), 3388-3397.
Guarino, M., Rubino, B. and Ballabio, G., 2007. The role of epithelial-mesenchymal transition in cancer pathology. *Pathology,* 39 (3), 305-318.
Guermonprez, P., et al., 2002. Antigen presentation and T cell stimulation by dendritic cells. *Annual Review of Immunology,* 20 (1), 621-667.
Guéry, L. and Hugues, S., 2015. Th17 cell plasticity and functions in cancer immunity. *BioMed Research International,* 2015.
Gurish, M.F. and Austen, K.F., 2001. The diverse roles of mast cells. *Journal of Experimental Medicine,* 194 (1), F1-F6.
Gutcher, I., et al., 2011. Autocrine transforming growth factor-β1 promotes in vivo Th17 cell differentiation. *Immunity,* 34 (3), 396-408.
Gygi, S.P., et al., 1999. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature Biotechnology,* 17 (10), 994.
Ha, D., et al., 2019. Differential control of human Treg and effector T cells in tumor immunity by Fc-engineered anti-CTLA-4 antibody. *Proceedings of the National Academy of Sciences of the United States of America,* 116 (2), 609-618.
Hadeiba, H., et al., 2012. Plasmacytoid dendritic cells transport peripheral antigens to the thymus to promote central tolerance. *Immunity,* 36 (3), 438-450.
Halim, L., et al., 2017. An atlas of human regulatory T helper-like cells reveals features of Th2-like Tregs that support a tumorigenic environment. *Cell Reports,* 20 (3), 757-770.
Hamerman, J.A., Ogasawara, K. and Lanier, L.L., 2005. NK cells in innate immunity. *Current Opinion in Immunology,* 17 (1), 29-35.
Hamm, A., et al., 2008. Frequent expression loss of Inter-alpha-trypsin inhibitor heavy chain (ITIH) genes in multiple human solid tumors: a systematic expression analysis. *BMC Cancer,* 8 (1), 25.
Hanagiri, T., et al., 2013. Clinical significance of the frequency of regulatory T cells in regional lymph node lymphocytes as a prognostic factor for non-small-cell lung cancer. *Lung Cancer,* 81 (3), 475-479.
Hanahan, D. and Weinberg, R.A., 2000. The hallmarks of cancer. *Cell,* 100 (1), 57-70.
Hanke, N., et al., 2013. Dendritic cell tumor killing activity and its potential applications in cancer immunotherapy. *Critical Reviews™ in Immunology,* 33 (1).
Haribhai, D., et al., 2011. A requisite role for induced regulatory T cells in tolerance based on expanding antigen receptor diversity. *Immunity,* 35 (1), 109-122.
Harlan, R. and Zhang, H., 2014. Targeted proteomics: a bridge between discovery and validation. *Expert Review of Proteomics,* 11 (6), 657-661.
Harrington, L.E., 2019. T-Cell Development. *In:* T-Cell Development. *Clinical Immunology.* Elsevier, 2019, pp. 119-125. e1.
Harrington, L.E., et al., 2005. Interleukin 17–producing CD4 effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nature Immunology,* 6 (11), 1123.
Harsha, H., Molina, H. and Pandey, A., 2008. Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nature Protocols,* 3 (3), 505.
Hathout, Y., 2015. *Proteomic Methods for Biomarker Discovery and Validation.are we there Yet?,* .
He, D., et al., 2010. IL-17 promotes tumor development through the induction of tumor promoting microenvironments at tumor sites and myeloid-derived suppressor cells. *The Journal of Immunology,* 184 (5), 2281-2288.
Hellmann, M.D., et al., 2018. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *New England Journal of Medicine,* 378 (22), 2093-2104.

Henderson, N.C., et al., 2013. Targeting of α v integrin identifies a core molecular pathway that regulates fibrosis in several organs. *Nature Medicine,* 19 (12), 1617.
Henry, N.L. and Hayes, D.F., 2012. Cancer biomarkers. *Molecular Oncology,* 6 (2), 140-146.
Herman, A.E., et al., 2004. CD4+CD25+ T regulatory cells dependent on ICOS promote regulation of effector cells in the prediabetic lesion. *The Journal of Experimental Medicine,* 199 (11), 1479-1489.
Hesketh, R., 2013. *Introduction to cancer biology.* Cambridge University Press.
Hikosaka, Y., et al., 2008. The cytokine RANKL produced by positively selected thymocytes fosters medullary thymic epithelial cells that express autoimmune regulator. *Immunity,* 29 (3), 438-450.
Hildner, K., et al., 2008. Batf3 deficiency reveals a critical role for CD8alpha+ dendritic cells in cytotoxic T cell immunity. *Science (New York, N.Y.),* 322 (5904), 1097-1100.
Himmel, M.E., et al., 2013. Helios+ and Helios- cells coexist within the natural FOXP3+ T regulatory cell subset in humans. *Journal of Immunology (Baltimore, Md.: 1950),* 190 (5), 2001-2008.
Hippen, K., et al., 2011. Generation and large-scale expansion of human inducible regulatory T cells that suppress graft-versus-host disease. *American Journal of Transplantation,* 11 (6), 1148-1157.
Ho, C.S., et al., 2003. Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist.Reviews,* 24 (1), 3-12.
Hogquist, K.A., Baldwin, T.A. and Jameson, S.C., 2005. Central tolerance: learning self-control in the thymus. *Nature Reviews Immunology,* 5 (10), 772.
Horgan, R.P. and Kenny, L.C., 2011. 'Omic'technologies: genomics, transcriptomics, proteomics and metabolomics. *The Obstetrician & Gynaecologist,* 13 (3), 189-195.
Hou, F., et al., 2011. MAVS forms functional prion-like aggregates to activate and propagate antiviral innate immune response. *Cell,* 146 (3), 448-461.
Hu, A., Noble, W.S. and Wolf-Yadlin, A., 2016. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research,* 5, 10.12688/f1000research.7042.1. eCollection 2016.
Huang, C., et al., 2004. Role of LAG-3 in regulatory T cells. *Immunity,* 21 (4), 503-513.
Huang, D.W., Sherman, B.T. and Lempicki, R.A., 2008. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research,* 37 (1), 1-13.
Huang, F., et al., 2017. Deep coverage of global protein expression and phosphorylation in breast tumor cell lines using TMT 10-plex isobaric labeling. *Journal of Proteome Research,* 16 (3), 1121-1132.
Huang, B., et al., 2008. SCF-mediated mast cell infiltration and activation exacerbate the inflammation and immunosuppression in tumor microenvironment. *Blood,* 112 (4), 1269-1279.
Huber, S. and Schramm, C., 2006. TGF-beta and CD4 CD25 regulatory T cells. *Front Biosci,* 11 (1), 1014-1023.
Hubert, P., et al., 2007. The cross-talk between dendritic and regulatory T cells: good or evil? *Journal of Leukocyte Biology,* 82 (4), 781-794.
Huehn, J., Polansky, J.K. and Hamann, A., 2009. Epigenetic control of FOXP3 expression: the key to a stable regulatory T-cell lineage? *Nature Reviews Immunology,* 9 (2), 83.
Hwang, H.K., et al., 2016. Prognostic impact of the tumor-infiltrating regulatory T-cell (Foxp3)/activated cytotoxic T lymphocyte (granzyme B) ratio on resected left-sided pancreatic cancer. *Oncology Letters,* 12 (6), 4477-4484.
Hyytiäinen, M., Penttinen, C. and Keski-Oja, J., 2004. Latent TGF-β binding proteins: extracellular matrix association and roles in TGF-β activation. *Critical Reviews in Clinical Laboratory Sciences,* 41 (3), 233-264.
Ichiyama, K., et al., 2011. Transcription factor Smad-independent T helper 17 cell induction by transforming-growth factor-β is mediated by suppression of eomesodermin. *Immunity,* 34 (5), 741-754.
Iliuk, A., Galan, J. and Tao, W.A., 2009. Playing tag with quantitative proteomics. *Analytical and Bioanalytical Chemistry,* 393 (2), 503-513.
Irla, M., et al., 2008. Autoantigen-specific interactions with CD4 thymocytes control mature medullary thymic epithelial cell cellularity. *Immunity,* 29 (3), 451-463.
Ishigame, H., et al., 2013. Excessive Th1 responses due to the absence of TGF-beta signaling cause autoimmune diabetes and dysregulated Treg cell homeostasis. *Proceedings of the National Academy of Sciences of the United States of America,* 110 (17), 6961-6966.
Ito, T., et al., 2013. The linkage of innate and adaptive immune response during granulomatous development. *Frontiers in Immunology,* 4, 10.
Ivanov, I.I., et al., 2006. The orphan nuclear receptor RORγt directs the differentiation program of proinflammatory IL-17 T helper cells. *Cell,* 126 (6), 1121-1133.

Iwata, M., et al., 2004. Retinoic acid imprints gut-homing specificity on T cells. *Immunity,* 21 (4), 527-538.
Jain, M., et al., 2002. Sustained loss of a neoplastic phenotype by brief inactivation of MYC. *Science (New York, N.Y.),* 297 (5578), 102-104.
Jaiswal, S., et al., 2009. CD47 is upregulated on circulating hematopoietic stem cells and leukemia cells to avoid phagocytosis. *Cell,* 138 (2), 271-285.
Janeway, C.A., et al., 1996. *Immunobiology: the immune system in health and disease.* Current Biology London.
Janusson, E., et al., 2015. Spatial effects on electrospray ionization response. *International Journal of Mass Spectrometry,* 388, 1-8.
Javle, M., et al., 2014. Biomarkers of TGF-β signaling pathway and prognosis of pancreatic cancer. *PloS One,* 9 (1), e85942.
Jenkins, G., 2008. The role of proteases in transforming growth factor-β activation. *The International Journal of Biochemistry & Cell Biology,* 40 (6-7), 1068-1078.
Jeon, S., et al., 2007. Mechanisms underlying TGF-β1-induced expression of VEGF and Flk-1 in mouse macrophages and their implications for angiogenesis. *Journal of Leukocyte Biology,* 81 (2), 557-566.
Johnson, L.D. and Jameson, S.C., 2012. TGF-β sensitivity restrains CD8 T cell homeostatic proliferation by enforcing sensitivity to IL-7 and IL-15. *PLoS One,* 7 (8), e42268.
Josefowicz, S.Z., Lu, L. and Rudensky, A.Y., 2012. Regulatory T cells: mechanisms of differentiation and function. *Annual Review of Immunology,* 30, 531-564.
Kainosho, M., et al., 2006. Optimal isotope labelling for NMR protein structure determinations. *Nature,* 440 (7080), 52.
Kakugawa, K., et al., 2017. Essential roles of SATB1 in specifying T lymphocyte subsets. *Cell Reports,* 19 (6), 1176-1188.
Kamburov, A., et al., 2011. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. *Bioinformatics,* 27 (20), 2917-2918.
Kanamori, M., et al., 2016. Induced regulatory T cells: their development, stability, and applications. *Trends in Immunology,* 37 (11), 803-811.
Karp, N.A., et al., 2010. Addressing accuracy and precision issues in iTRAQ quantitation. *Molecular & Cellular Proteomics : MCP,* 9 (9), 1885-1897.
Kashyap, R.S., et al., 2009. Inter-α-trypsin inhibitor heavy chain 4 is a novel marker of acute ischemic stroke. *Clinica Chimica Acta,* 402 (1-2), 160-163.
Katsila, T., et al., 2014. Circulating pEGFR is a candidate response biomarker of cetuximab therapy in colorectal cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 20 (24), 6346-6356.
Kauder, S.E., et al., 2018. ALX148 blocks CD47 and enhances innate and adaptive antitumor immunity with a favorable safety profile. *PloS One,* 13 (8), e0201832.
Kawano, Y., et al., 2018. Blocking IFNRA1 inhibits multiple myeloma-driven Treg expansion and immunosuppression. *The Journal of Clinical Investigation,* 128 (6).
Keller, A., et al., 2015. Automated Validation of Results and Removal of Fragment Ion Interferences in Targeted Analysis of Data-independent Acquisition Mass Spectrometry (MS) using SWATHProphet. *Molecular & Cellular Proteomics : MCP,* 14 (5), 1411-1418.
Kelley, M. and DeSilva, B., 2007. Key elements of bioanalytical method validation for macromolecules. *The AAPS Journal,* 9 (2), E156-E163.
Kew, M., 2013. Hepatitis viruses (other than hepatitis B and C viruses) as causes of hepatocellular carcinoma: an update. *Journal of Viral Hepatitis,* 20 (3), 149-157.
Kim, E.Y., et al., 2017. MYC expression correlates with PD-L1 expression in non-small cell lung cancer. *Lung Cancer,* 110, 63-67.
Kim, R., Emi, M. and Tanabe, K., 2007. Cancer immunoediting from immune surveillance to immune escape. *Immunology,* 121 (1), 1-14.
Kim, S., et al., 2014. Zonal difference and prognostic significance of foxp3 regulatory T cell infiltration in breast cancer. *Journal of Breast Cancer,* 17 (1), 8-17.
Kim, J.W., et al., 2013. Clinical Implications of VEGF, TGF-beta1, and IL-1beta in Patients with Advanced Non-small Cell Lung Cancer. *Cancer Research and Treatment : Official Journal of Korean Cancer Association,* 45 (4), 325-333.
Kindt, T.J., et al., 2007. *Kuby immunology.* Macmillan.
Kinter, M. and Sherman, N.E., 2005. *Protein sequencing and identification using tandem mass spectrometry.* John Wiley & Sons.

Kitamura, T., Qian, B. and Pollard, J.W., 2015. Immune cell promotion of metastasis. *Nature Reviews Immunology,* 15 (2), 73.
Klein, L., et al., 2014a. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nature Reviews Immunology,* 14 (6), 377.
Klein, L., et al., 2014b. Positive and negative selection of the T cell repertoire: what thymocytes see (and don't see). *Nature Reviews Immunology,* 14 (6), 377.
Knochelmann, H.M., et al., 2018. When worlds collide: Th17 and Treg cells in cancer and autoimmunity. *Cellular & Molecular Immunology,* 15 (5), 458.
Knowles, M. and Selby, P., 2005. *Introduction to the cellular and molecular biology of cancer.* Oxford university press.
Kobayashi, S.D. and DeLeo, F.R., 2009. Role of neutrophils in innate immunity: a systems biology-level approach. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine,* 1 (3), 309-333.
Koch, U. and Radtke, F., 2011. Mechanisms of T cell development and transformation. *Annual Review of Cell and Developmental Biology,* 27, 539-562.
Koebel, C.M., et al., 2007. Adaptive immunity maintains occult cancer in an equilibrium state. *Nature,* 450 (7171), 903.
Koehler, H., et al., 2007. CD28 costimulation overcomes transforming growth factor-beta-mediated repression of proliferation of redirected human CD4+ and CD8+ T cells in an antitumor cell attack. *Cancer Research,* 67 (5), 2265-2273.
Kohla, M.A.S., et al., 2016. Association of serum levels of epidermal growth factor with disease severity in patients with unresectable hepatocellular carcinoma. *Hepatoma Res,* 2, 18-25.
Konkel, J.E., et al., 2011. Control of the development of CD8αα intestinal intraepithelial lymphocytes by TGF-β. *Nature Immunology,* 12 (4), 312.
Kortlever, R.M., et al., 2017. Myc cooperates with Ras by programming inflammation and immune suppression. *Cell,* 171 (6), 1301-1315. e14.
Kortylewski, M., et al., 2005. Inhibiting Stat3 signaling in the hematopoietic system elicits multicomponent antitumor immunity. *Nature Medicine,* 11 (12), 1314.
Kotsakis, A., et al., 2016. Prognostic value of circulating regulatory T cell subsets in untreated non-small cell lung cancer patients. *Scientific Reports,* 6, 39247.
Krijgsman, D., Hokland, M. and Kuppen, P.J., 2018. The role of natural killer T cells in cancer—a phenotypical and functional approach. *Frontiers in Immunology,* 9, 367.
Kryczek, I., et al., 2009. Phenotype, distribution, generation, and functional and clinical relevance of Th17 cells in the human tumor environments. *Blood,* 114 (6), 1141-1149.
Kubiczkova, L., et al., 2012. TGF-β–an excellent servant but a bad master. *Journal of Translational Medicine,* 10 (1), 183.
Kujawski, M., et al., 2008. Stat3 mediates myeloid cell-dependent tumor angiogenesis in mice. *The Journal of Clinical Investigation,* 118 (10), 3367-3377.
Kumar, D., et al., 2016. Integrating transcriptome and proteome profiling: strategies and applications. *Proteomics,* 16 (19), 2533-2544.
Kurose, K., et al., 2015. Phase Ia Study of FoxP3+ CD4 Treg Depletion by Infusion of a Humanized Anti-CCR4 Antibody, KW-0761, in Cancer Patients. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 21 (19), 4327-4336.
Kurtz, J., 2004. Memory in the innate and adaptive immune systems. *Microbes and Infection,* 6 (15), 1410-1417.
Kushwah, R. and Hu, J., 2011. Role of dendritic cells in the induction of regulatory T cells. *Cell & Bioscience,* 1 (1), 20.
Kuwahara, M., et al., 2012. The transcription factor Sox4 is a downstream target of signaling by the cytokine TGF-β and suppresses T H 2 differentiation. *Nature Immunology,* 13 (8), 778.
Lambert, J., et al., 2013. Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nature Methods,* 10 (12), 1239.
Lancashire, L.J., Lemetre, C. and Ball, G.R., 2009. An introduction to artificial neural networks in bioinformatics—application to complex microarray and mass spectrometry datasets in cancer studies. *Briefings in Bioinformatics,* 10 (3), 315-329.
Lancashire, L.J., et al., 2010. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Research and Treatment,* 120 (1), 83-93.
Lancashire, L.J., Rees, R.C. and Ball, G.R., 2008a. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. *Artificial Intelligence in Medicine,* 43 (2), 99-111.

Lancashire, L.J., Rees, R.C. and Ball, G.R., 2008b. Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach. *Artificial Intelligence in Medicine,* 43 (2), 99-111.
Langenskiöld, M., et al., 2008. Increased TGF-Beta1 protein expression in patients with advanced colorectal cancer. *Journal of Surgical Oncology,* 97 (5), 409-415.
Larkin, J., et al., 2015. Combined nivolumab and ipilimumab or monotherapy in untreated melanoma. *New England Journal of Medicine,* 373 (1), 23-34.
Lathrop, S.K., et al., 2011. Peripheral education of the immune system by colonic commensal microbiota. *Nature,* 478 (7368), 250.
Latosinska, A., et al., 2015. Comparative analysis of label-free and 8-Plex iTRAQ approach for quantitative tissue proteomic analysis. *PloS One,* 10 (9), e0137048.
Lawrence, D.A., 2001. Latent-TGF-β: an overview. *Molecular and Cellular Biochemistry,* 219 (1-2), 163-170.
Le Bras, S. and Geha, R.S., 2006. IPEX and the role of Foxp3 in the development and function of human Tregs. *The Journal of Clinical Investigation,* 116 (6), 1473-1475.
LeBien, T.W. and Tedder, T.F., 2008. B lymphocytes: how they develop and function. *Blood,* 112 (5), 1570-1580.
Lee, H., Bautista, J.L. and Hsieh, C., 2011. Thymic and peripheral differentiation of regulatory T cells. *In:* Thymic and peripheral differentiation of regulatory T cells. *Advances in immunology.* Elsevier, 2011, pp. 25-71.
Lee, J.H., Lydon, J.P. and Kim, C.H., 2012. Progesterone suppresses the m TOR pathway and promotes generation of induced regulatory T cells with increased stability. *European Journal of Immunology,* 42 (10), 2683-2696.
Lee, S.M., et al., 2004. PLP2/A4 interacts with CCR1 and stimulates migration of CCR1-expressing HOS cells. *Biochemical and Biophysical Research Communications,* 324 (2), 768-772.
Lee, S., et al., 2018. Matrix metalloproteinase-9 in monocytic myeloid-derived suppressor cells correlate with early infections and clinical outcomes in allogeneic hematopoietic stem cell transplantation. *Biology of Blood and Marrow Transplantation,* 24 (1), 32-42.
Lee, Y.H., Tan, H.T. and Chung, M.C., 2010. Subcellular fractionation methods and strategies for proteomics. *Proteomics,* 10 (22), 3935-3956.
Lei, Y., et al., 2011. Aire-dependent production of XCL1 mediates medullary accumulation of thymic dendritic cells and contributes to regulatory T cell development. *The Journal of Experimental Medicine,* 208 (2), 383-394.
Lemière, F., 2001. Mass analysers for LC-MS. *Guide to LC-MS, Sl, LC-GC Europe,* , 22-28.
Leone, P., et al., 2013. MHC class I antigen processing and presenting machinery: organization, function, and defects in tumor cells. *Journal of the National Cancer Institute,* 105 (16), 1172-1187.
Leth-Larsen, R., Lund, R.R. and Ditzel, H.J., 2010. Plasma membrane proteomics and its application in clinical cancer biomarker discovery. *Molecular & Cellular Proteomics : MCP,* 9 (7), 1369-1382.
Levy, L. and Hill, C.S., 2006. Alterations in components of the TGF-β superfamily signaling pathways in human cancer. *Cytokine & Growth Factor Reviews,* 17 (1-2), 41-58.
Li, C.X., et al., 2016. CXCL10/CXCR3 signaling mobilized-regulatory T cells promote liver tumor recurrence after transplantation. *Journal of Hepatology,* 65 (5), 944-952.
Li, J., et al., 2011. Interleukin 17A promotes hepatocellular carcinoma metastasis via NF-kB induced matrix metalloproteinases 2 and 9 expression. *PloS One,* 6 (7), e21816.
Li, J., et al., 2010. Renal cell carcinoma may evade the immune system by converting CD4 Foxp3-T cells into CD4 CD25 Foxp3 regulatory T cells: Role of tumor COX-2-derived PGE2. *Molecular Medicine Reports,* 3 (6), 959-963.
Li, J., et al., 2019. Prognostic value of TGF-β in lung cancer: systematic review and meta-analysis. *BMC Cancer,* 19 (1), 691.
Li, M.O., Sanjabi, S. and Flavell, R.A., 2006. Transforming growth factor-β controls development, homeostasis, and tolerance of T cells by regulatory T cell-dependent and-independent mechanisms. *Immunity,* 25 (3), 455-471.
Li, M.O., et al., 2006. Transforming growth factor-β regulation of immune responses. *Annu.Rev.Immunol.,* 24, 99-146.
Li, M., et al., 2014. CD4 Foxp3 regulatory T cell differentiation mediated by endometrial stromal cell-derived TECK promotes the growth and invasion of endometriotic lesions. *Cell Death & Disease,* 5 (10), e1436.
Li, Z., et al., 2015. COX-2 promotes metastasis in nasopharyngeal carcinoma by mediating interactions between cancer cells and myeloid-derived suppressor cells. *Oncoimmunology,* 4 (11), e1044712.

Li, C., Ebert, P.J. and Li, Q.J., 2013. T cell receptor (TCR) and transforming growth factor beta (TGF-beta) signaling converge on DNA (cytosine-5)-methyltransferase to control forkhead box protein 3 (foxp3) locus methylation and inducible regulatory T cell differentiation. *The Journal of Biological Chemistry,* 288 (26), 19127-19139.
Li, H., et al., 2009. Cancer-expanded myeloid-derived suppressor cells induce anergy of NK cells through membrane-bound TGF-beta 1. *Journal of Immunology (Baltimore, Md.: 1950),* 182 (1), 240-249.
Li, Q., et al., 2014. Interleukin-17 Indirectly Promotes M2 Macrophage Differentiation through Stimulation of COX-2/PGE2 Pathway in the Cancer Cells. *Cancer Research and Treatment : Official Journal of Korean Cancer Association,* 46 (3), 297-306.
Liang, Q., Wang, C. and Li, B., 2015. Metabolomic analysis using liquid chromatography/mass spectrometry for gastric cancer. *Applied Biochemistry and Biotechnology,* 176 (8), 2170-2184.
Liang, B., et al., 2008. Regulatory T cells inhibit dendritic cells by lymphocyte activation gene-3 engagement of MHC class II. *Journal of Immunology (Baltimore, Md.: 1950),* 180 (9), 5916-5926.
Liang, S.C., et al., 2006. Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *The Journal of Experimental Medicine,* 203 (10), 2271-2279.
Lieberman, J., 2003. Cell death and immunity: the ABCs of granule-mediated cytotoxicity: new weapons in the arsenal. *Nature Reviews Immunology,* 3 (5), 361.
Liebler, D.C., 2001. *Introduction to proteomics: tools for the new biology.* Springer Science & Business Media.
Lin, C.Y., et al., 2012. Transcriptional amplification in tumor cells with elevated c-Myc. *Cell,* 151 (1), 56-67.
Lin, J.T., et al., 2005. TGF-beta 1 uses distinct mechanisms to inhibit IFN-gamma expression in CD4+ T cells at priming and at recall: differential involvement of Stat4 and T-bet. *Journal of Immunology (Baltimore, Md.: 1950),* 174 (10), 5950-5958.
Lin, T.H., et al., 2015. High Serum Transforming Growth Factor-beta1 Levels Predict Outcome in Hepatocellular Carcinoma Patients Treated with Sorafenib. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 21 (16), 3678-3684.
Lin, X., et al., 2013. Advances in distinguishing natural from induced Foxp3(+) regulatory T cells. *International Journal of Clinical and Experimental Pathology,* 6 (2), 116-123.
Lindner, S., et al., 2013. Interleukin 21-induced granzyme B-expressing B cells infiltrate tumors and regulate T cells. *Cancer Research,* 73 (8), 2468-2479.
Lio, C.J. and Hsieh, C., 2008. A two-step process for thymic regulatory T cell development. *Immunity,* 28 (1), 100-111.
Listgarten, J. and Emili, A., 2005. Statistical and computational methods for comparative proteomic profiling using liquid chromatography-tandem mass spectrometry. *Molecular & Cellular Proteomics,* 4 (4), 419-434.
Liu, C., Workman, C.J. and Vignali, D.A., 2016a. Targeting regulatory T cells in tumors. *The FEBS Journal,* 283 (14), 2731-2748.
Liu, C., Workman, C.J. and Vignali, D.A., 2016b. Targeting regulatory T cells in tumors. *The FEBS Journal,* 283 (14), 2731-2748.
Liu, C., Workman, C.J. and Vignali, D.A., 2016c. Targeting regulatory T cells in tumors. *The FEBS Journal,* 283 (14), 2731-2748.
Liu, C., et al., 2017a. Circulating regulatory T cell subsets predict overall survival of patients with unresectable pancreatic cancer. *International Journal of Oncology,* 51 (2), 686-694.
Liu, K. and Nussenzweig, M.C., 2010. Development and homeostasis of dendritic cells. *European Journal of Immunology,* 40 (8), 2099-2102.
Liu, N., et al., 2015. SMAD4 expression in breast ductal carcinoma correlates with prognosis. *Oncology Letters,* 10 (3), 1709-1715.
Liu, X., et al., 2009. T cell receptor CDR3 sequence but not recognition characteristics distinguish autoreactive effector and Foxp3 regulatory T cells. *Immunity,* 31 (6), 909-920.
Liu, Y., et al., 2008. A critical function for TGF-β signaling in the development of natural CD4 CD25 Foxp3 regulatory T cells. *Nature Immunology,* 9 (6), 632.
Liu, C., et al., 2017b. Predictive value of peripheral regulatory T cells in non-small cell lung cancer patients undergoing radiotherapy. *Oncotarget,* 8 (26), 43427-43438.
Liu, K., et al., 2009. In vivo analysis of dendritic cell development and homeostasis. *Science (New York, N.Y.),* 324 (5925), 392-397.

Liu, V.C., et al., 2007. Tumor evasion of the immune system by converting CD4+CD25- T cells into CD4+CD25+ T regulatory cells: role of tumor-derived TGF-beta. *Journal of Immunology (Baltimore, Md.: 1950),* 178 (5), 2883-2892.
Lowe, R., et al., 2017. Transcriptomics technologies. *PLoS Computational Biology,* 13 (5), e1005457.
Lu, C., et al., 2016. The expression profiles and regulation of PD-L1 in tumor-induced myeloid-derived suppressor cells. *Oncoimmunology,* 5 (12), e1247135.
Lu, L., et al., 2010. Characterization of protective human CD4 CD25 FOXP3 regulatory T cells generated with IL-2, TGF-β and retinoic acid. *PloS One,* 5 (12), e15150.
Lucas, B., et al., 2017. Progressive changes in CXCR4 expression that define thymocyte positive selection are dispensable for both innate and conventional αβT-cell development. *Scientific Reports,* 7 (1), 5068.
Ludwig, C., et al., 2018. Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Molecular Systems Biology,* 14 (8).
Luke, J.J., et al., 2016. *Correlation of WNT/B-Catenin Pathway Activation with Immune Exclusion Across most Human Cancers.,* .
Luo, J., et al., 2018. TgF-beta 1 levels are associated with lymphocyte percentages in patients with lung cancer treated with radiation therapy. *OncoTargets and Therapy,* 11, 8349.
Luscombe, N.M., Greenbaum, D. and Gerstein, M., 2001. What is bioinformatics? A proposed definition and overview of the field. *Methods of Information in Medicine,* 40 (04), 346-358.
Luu, M., Steinhoff, U. and Visekruna, A., 2017. Functional heterogeneity of gut-resident regulatory T cells. *Clinical & Translational Immunology,* 6 (9), e156.
Lv, Y., et al., 2019. Increased intratumoral mast cells foster immune suppression and gastric cancer progression through TNF-α-PD-L1 pathway. *Journal for Immunotherapy of Cancer,* 7 (1), 54.
MacLean, B., et al., 2010. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics,* 26 (7), 966-968.
Maecker, H.T., McCoy, J.P. and Nussenblatt, R., 2012. Standardizing immunophenotyping for the human immunology project. *Nature Reviews Immunology,* 12 (3), 191.
Mailliard, R.B., et al., 2003. Dendritic cells mediate NK cell help for Th1 and CTL responses: two-signal requirement for the induction of NK cell helper function. *Journal of Immunology (Baltimore, Md.: 1950),* 171 (5), 2366-2373.
Malkoski, S.P., et al., 2012. Loss of transforming growth factor beta type II receptor increases aggressive tumor behavior and reduces survival in lung adenocarcinoma and squamous cell carcinoma. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 18 (8), 2173-2183.
Mann, M., 2006. Innovations: Functional and quantitative proteomics using SILAC. *Nature Reviews Molecular Cell Biology,* 7 (12), 952.
Mantovani, A., et al., 2008. Cancer-related inflammation. *Nature,* 454 (7203), 436.
Marie, J.C., Liggitt, D. and Rudensky, A.Y., 2006. Cellular mechanisms of fatal early-onset autoimmunity in mice with the T cell-specific targeting of transforming growth factor-β receptor. *Immunity,* 25 (3), 441-454.
Marigo, I., et al., 2010. Tumor-induced tolerance and immune suppression depend on the C/EBPβ transcription factor. *Immunity,* 32 (6), 790-802.
Markiewicz, M.A., et al., 2009. IL-12 enhances CTL synapse formation and induces self-reactivity. *Journal of Immunology (Baltimore, Md.: 1950),* 182 (3), 1351-1361.
Martens, A., et al., 2016. Baseline Peripheral Blood Biomarkers Associated with Clinical Outcome of Advanced Melanoma Patients Treated with Ipilimumab. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 22 (12), 2908-2918.
Martinez, G.J., et al., 2010. Smad2 positively regulates the generation of Th17 cells. *The Journal of Biological Chemistry,* 285 (38), 29039-29043.
Martins-de-Souza, D., et al., 2010. The role of proteomics in depression research. *European Archives of Psychiatry and Clinical Neuroscience,* 260 (6), 499-506.
Marvel, D. and Gabrilovich, D.I., 2015. Myeloid-derived suppressor cells in the tumor microenvironment: expect the unexpected. *The Journal of Clinical Investigation,* 125 (9), 3356-3364.
Marx, V., 2013. Targeted proteomics. *Nature Methods,* 10 (1), 19.
Massagué, J., 2008. TGFβ in cancer. *Cell,* 134 (2), 215-230.
Massague, J., Seoane, J. and Wotton, D., 2005. Smad transcription factors. *Genes & Development,* 19 (23), 2783-2810.
Masucci, G.V., et al., 2016. Validation of biomarkers to predict response to immunotherapy in cancer: volume I—pre-analytical and analytical validation. *Journal for Immunotherapy of Cancer,* 4 (1), 76.

Matthiesen, R. and Bunkenborg, J., 2013. Introduction to mass spectrometry-based proteomics. *In:* Introduction to mass spectrometry-based proteomics. *Mass spectrometry data analysis in proteomics.* Springer, 2013, pp. 1-45.
Mbongue, J.C., et al., 2017. The role of dendritic cell maturation in the induction of insulin-dependent diabetes mellitus. *Frontiers in Immunology,* 8, 327.
McCarthy, F.M., et al., 2005. Differential detergent fractionation for non-electrophoretic eukaryote cell proteomics. *Journal of Proteome Research,* 4 (2), 316-324.
McCarthy, N.I., et al., 2015. Osteoprotegerin-Mediated Homeostasis of Rank+ Thymic Epithelial Cells Does Not Limit Foxp3+ Regulatory T Cell Development. *Journal of Immunology (Baltimore, Md.: 1950),* 195 (6), 2675-2682.
McCracken, M.N., Cha, A.C. and Weissman, I.L., 2015. Molecular Pathways: Activating T Cells after Cancer Cell Phagocytosis from Blockade of CD47 "Don't Eat Me" Signals. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 21 (16), 3597-3601.
McDonald, W.H. and Yates, J.R., 2002. Shotgun proteomics and biomarker discovery. *Disease Markers,* 18 (2), 99-105.
McDonald, T., et al., 2006. Expanding the subproteome of the inner mitochondria using protein separation technologies: one- and two-dimensional liquid chromatography and two-dimensional gel electrophoresis. *Molecular & Cellular Proteomics : MCP,* 5 (12), 2392-2411.
McDonnell, A.M., Robinson, B.W. and Currie, A.J., 2010. Tumor antigen cross-presentation and the dendritic cell: where it all begins? *Clinical & Developmental Immunology,* 2010, 539519.
McEarchern, J.A., et al., 2001. Invasion and metastasis of a mammary tumor involves TGF-β signaling. *International Journal of Cancer,* 91 (1), 76-82.
McGeachy, M.J., et al., 2007. TGF-β and IL-6 drive the production of IL-17 and IL-10 by T cells and restrain T H-17 cell–mediated pathology. *Nature Immunology,* 8 (12), 1390.
McHugh, R.S., et al., 2002. CD4 CD25 immunoregulatory T cells: gene expression analysis reveals a functional role for the glucocorticoid-induced TNF receptor. *Immunity,* 16 (2), 311-323.
McKarns, S.C. and Schwartz, R.H., 2005. Distinct effects of TGF-beta 1 on CD4+ and CD8+ T cell survival, division, and IL-2 production: a role for T cell intrinsic Smad3. *Journal of Immunology (Baltimore, Md.: 1950),* 174 (4), 2071-2083.
McKarns, S.C., Schwartz, R.H. and Kaminski, N.E., 2004. Smad3 is essential for TGF-beta 1 to suppress IL-2 production and TCR-induced proliferation, but not IL-2-induced proliferation. *Journal of Immunology (Baltimore, Md.: 1950),* 172 (7), 4275-4284.
McWhirter, S.M. and Maniatis, T., 2005. Connecting mitochondria and innate immunity. *Cell,* 122 (5), 645-647.
Medzhitov, R., 2008. Origin and physiological roles of inflammation. *Nature,* 454 (7203), 428.
Megger, D.A., et al., 2013. Label-free quantification in clinical proteomics. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics,* 1834 (8), 1581-1590.
Megger, D.A., et al., 2014. Comparison of label-free and label-based strategies for proteome analysis of hepatoma cell lines. *Biochimica Et Biophysica Acta (BBA)-Proteins and Proteomics,* 1844 (5), 967-976.
Merlo, A., et al., 2009. FOXP3 expression and overall survival in breast cancer. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology,* 27 (11), 1746-1752.
Merrick, B.A., et al., 2011. Platforms for biomarker analysis using high-throughput approaches in genomics, transcriptomics, proteomics, metabolomics, and bioinformatics. *IARC Scientific Publications,* (163) (163), 121-142.
Messerschmidt, J.L., Prendergast, G.C. and Messerschmidt, G.L., 2016. How Cancers Escape Immune Destruction and Mechanisms of Action for the New Significantly Active Immune Therapies: Helping Nonimmunologists Decipher Recent Advances. *The Oncologist,* 21 (2), 233-243.
Metrock, L.K., et al., 2017. Utility of peripheral blood immunophenotyping by flow cytometry in the diagnosis of pediatric acute leukemia. *Pediatric Blood & Cancer,* 64 (10), e26526.
Millán, O., et al., 2010. Biomarkers of immunoregulatory status in stable liver transplant recipients undergoing weaning of immunosuppressive therapy. *Clinical Immunology,* 137 (3), 337-346.
Millán, O. and Brunet, M., 2015. Flow Cytometry as Platform for Biomarker Discovery and Clinical Validation. *General Methods in Biomarker Research and their Applications,* , 141-164.
Milpied, P., et al., 2009. Neuropilin-1 is not a marker of human Foxp3 Treg. *European Journal of Immunology,* 39 (6), 1466-1471.
Mishra, N.C., 2010. *Introduction to proteomics: principles and applications.* Wiley Online Library.
Misra, B.B., et al., 2018. Integrated omics: tools, advances, and future approaches. *Journal of Molecular Endocrinology,* 1 (aop).

Mittal, D., et al., 2014. New insights into cancer immunoediting and its three component phases—elimination, equilibrium and escape. *Current Opinion in Immunology,* 27, 16-25.
Mitulović, G. and Mechtler, K., 2006. HPLC techniques for proteomics analysis—a short overview of latest developments. *Briefings in Functional Genomics,* 5 (4), 249-260.
Mizrahi, M. and Ilan, Y., 2009. The gut mucosa as a site for induction of regulatory T-cells. *Current Pharmaceutical Design,* 15 (11), 1191-1202.
Mizukami, Y., et al., 2008. CCL17 and CCL22 chemokines within tumor microenvironment are related to accumulation of Foxp3 regulatory T cells in gastric cancer. *International Journal of Cancer,* 122 (10), 2286-2293.
Mogensen, T.H., 2009. Pathogen recognition and inflammatory signaling in innate immune defenses. *Clinical Microbiology Reviews,* 22 (2), 240-73, Table of Contents.
Mohammad, I., et al., 2018. Quantitative proteomic characterization and comparison of T helper 17 and induced regulatory T cells. *PLoS Biology,* 16 (5), e2004194.
Molica, S., et al., 2003. Tryptase-positive mast cells predict clinical outcome of patients with early B-cell chronic lymphocytic leukemia.
Monroe, M.E., et al., 2007. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinformatics,* 23 (15), 2021-2023.
Moreno, P., et al., 2018. Metabolomic profiling of human lung tumor tissues–nucleotide metabolism as a candidate for therapeutic interventions and biomarkers. *Molecular Oncology,* 12 (10), 1778-1796.
Moser, M., 2003. Dendritic cells in immunity and tolerance—do they display opposite functions? *Immunity,* 19 (1), 5-8.
Mosser, D.M. and Edwards, J.P., 2008. Exploring the full spectrum of macrophage activation. *Nature Reviews Immunology,* 8 (12), 958.
Mount, D.W. and Pandey, R., 2005. Using bioinformatics and genome analysis for new therapeutic interventions. *Molecular Cancer Therapeutics,* 4 (10), 1636-1643.
Moustakas, A. and de Herreros, A.G., 2017. Epithelial-mesenchymal transition in cancer. *Molecular Oncology,* 11 (7), 715-717.
Mueller, D.L., 2010. Mechanisms maintaining peripheral tolerance. *Nature Immunology,* 11 (1), 21.
Mueller, L.N., et al., 2007. SuperHirn–a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics,* 7 (19), 3470-3480.
Müller, M., et al., 2001. Limits for the detection of (poly-) phosphoinositides by matrix-assisted laser desorption and ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Chemistry and Physics of Lipids,* 110 (2), 151-164.
Munz, C., Steinman, R.M. and Fujii, S., 2005. Dendritic cell maturation by innate lymphocytes: coordinated stimulation of innate and adaptive immunity. *The Journal of Experimental Medicine,* 202 (2), 203-207.
Murphy, K. and Weaver, C., 2016. *Janeway's immunobiology.* Garland Science.
Murphy-Ullrich, J.E. and Poczatek, M., 2000. Activation of latent TGF-β by thrombospondin-1: mechanisms and physiology. *Cytokine & Growth Factor Reviews,* 11 (1-2), 59-69.
Murshid, A., Gong, J. and Calderwood, S.K., 2012. The role of heat shock proteins in antigen cross presentation. *Frontiers in Immunology,* 3, 63.
Nakamura, T. and Ushigome, H., 2018. Myeloid-derived suppressor cells as a regulator of immunity in organ transplantation. *International Journal of Molecular Sciences,* 19 (8), 2357.
Narasimhan, M., et al., 2019. Clinical biomarker discovery by SWATH-MS based label-free quantitative proteomics: impact of criteria for identification of differentiators and data normalization method. *Journal of Translational Medicine,* 17 (1), 184.
Navarro, P., et al., 2016. A multicenter study benchmarks software tools for label-free proteome quantification. *Nature Biotechnology,* 34 (11), 1130.
Neilson, K.A., et al., 2011. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics,* 11 (4), 535-553.
Nesbitt, C.A., Zhang, H. and Yeung, K.K., 2008. Recent applications of capillary electrophoresis–mass spectrometry (CE–MS): CE performing functions beyond separation. *Analytica Chimica Acta,* 627 (1), 3-24.
Ng, T., et al., 2013. Regulation of adaptive immunity; the role of interleukin-10. *Frontiers in Immunology,* 4, 129.
Nguyen, N.T., et al., 2014. Aryl hydrocarbon receptor and kynurenine: recent advances in autoimmune disease research. *Frontiers in Immunology,* 5, 551.
Nonomura, N., et al., 2007. Decreased number of mast cells infiltrating into needle biopsy specimens leads to a better prognosis of prostate cancer. *British Journal of Cancer,* 97 (7), 952.

Noy, R. and Pollard, J.W., 2014. Tumor-associated macrophages: from mechanisms to therapy. *Immunity,* 41 (1), 49-61.
Oberle, N., et al., 2007. Rapid suppression of cytokine transcription in human CD4+CD25 T cells by CD4+Foxp3+ regulatory T cells: independence of IL-2 consumption, TGF-beta, and various inhibitors of TCR signaling. *Journal of Immunology (Baltimore, Md.: 1950),* 179 (6), 3578-3587.
Obermajer, N., et al., 2011. Positive feedback between PGE2 and COX2 redirects the differentiation of human dendritic cells toward stable myeloid-derived suppressor cells. *Blood,* 118 (20), 5498-5505.
Oda, Y., et al., 1999. Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America,* 96 (12), 6591-6596.
Oderup, C., et al., 2006. Cytotoxic T lymphocyte antigen-4-dependent down-modulation of costimulatory molecules on dendritic cells in CD4 CD25 regulatory T-cell-mediated suppression. *Immunology,* 118 (2), 240-249.
Oh, S.A. and Li, M.O., 2013. TGF-beta: guardian of T cell function. *Journal of Immunology (Baltimore, Md.: 1950),* 191 (8), 3973-3979.
Oliva, M., Wu, T. and Yang, V., 1993. Isolation and characterization of a differentiation-dependent gene in the human colonic cell line HT29-18. *Archives of Biochemistry and Biophysics,* 302 (1), 183-192.
Olkhanud, P.B., et al., 2011. Tumor-evoked regulatory B cells promote breast cancer metastasis by converting resting CD4(+) T cells to T-regulatory cells. *Cancer Research,* 71 (10), 3505-3515.
Olsen, J.V., Ong, S.E. and Mann, M., 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Molecular & Cellular Proteomics : MCP,* 3 (6), 608-614.
Ong, S., et al., 2006. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *In:* Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Cell Biology.* Elsevier, 2006, pp. 427-436.
Ong, S.E., et al., 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Molecular & Cellular Proteomics : MCP,* 1 (5), 376-386.
Orsburn, B.C., Stockwin, L.H. and Newton, D.L., 2011. Challenges in plasma membrane phosphoproteomics. *Expert Review of Proteomics,* 8 (4), 483-494.
Ortea, I., et al., 2016. Discovery of potential protein biomarkers of lung adenocarcinoma in bronchoalveolar lavage fluid by SWATH MS data-independent acquisition and targeted data extraction. *Journal of Proteomics,* 138, 106-114.
Ostrand-Rosenberg, S., et al., 2012. Cross-talk between myeloid-derived suppressor cells (MDSC), macrophages, and dendritic cells enhances tumor-induced immune suppression. *In: Seminars in cancer biology,* Elsevier, pp. 275-281.
Ostrowski, J., et al., 2018. Redefining the Practical Utility of Blood Transcriptome Biomarkers in Inflammatory Bowel Diseases. *Journal of Crohn's and Colitis,* 13 (5), 626-633.
Ouyang, W., et al., 2010. Transforming growth factor-β signaling curbs thymic negative selection promoting regulatory T cell development. *Immunity,* 32 (5), 642-653.
Ouyang, W., et al., 2013. TGF-β cytokine signaling promotes CD8 T cell development and low-affinity CD4 T cell homeostasis by regulation of interleukin-7 receptor α expression. *Immunity,* 39 (2), 335-346.
Pai, S.G., et al., 2017. Wnt/beta-catenin pathway: modulating anticancer immune response. *Journal of Hematology & Oncology,* 10 (1), 101.
Palucka, K., et al., 2011. Dendritic cells and immunity against cancer. *Journal of Internal Medicine,* 269 (1), 64-73.
Pan, P.Y., et al., 2010. Immune stimulatory receptor CD40 is required for T-cell suppression and T regulatory cell activation mediated by myeloid-derived suppressor cells in cancer. *Cancer Research,* 70 (1), 99-108.
Pandeswari, P.B. and Sabareesh, V., 2019. Middle-down approach: a choice to sequence and characterize proteins/proteomes by mass spectrometry. *RSC Advances,* 9 (1), 313-344.
Pandiyan, P., et al., 2007. CD4 CD25 Foxp3 regulatory T cells induce cytokine deprivation–mediated apoptosis of effector CD4 T cells. *Nature Immunology,* 8 (12), 1353.
Park, H., et al., 2018. *The Role of Soluble TGF-Beta and its Dynamics for Predicting the Prognosis in Unresectable Pancreatic Cancer Patients Treated with Chemotherapy.,* .
Park, J., et al., 2010. Signaling by intrathymic cytokines, not T cell antigen receptors, specifies CD8 lineage choice and promotes the differentiation of cytotoxic-lineage T cells. *Nature Immunology,* 11 (3), 257.
Park, J.B. and Koo, J.S., 2014. Helicobacter pylori infection in gastric mucosa-associated lymphoid tissue lymphoma. *World Journal of Gastroenterology,* 20 (11), 2751-2759.

Parker, C.E. and Borchers, C.H., 2014. Mass spectrometry based biomarker discovery, verification, and validation–quality assurance and control of protein biomarker assays. *Molecular Oncology,* 8 (4), 840-858.
Parker, K.H., Beury, D.W. and Ostrand-Rosenberg, S., 2015. Myeloid-derived suppressor cells: critical cells driving immune suppression in the tumor microenvironment. *In:* Myeloid-derived suppressor cells: critical cells driving immune suppression in the tumor microenvironment. *Advances in cancer research.* Elsevier, 2015, pp. 95-139.
Pecorino, L., 2012. *Molecular biology of cancer: mechanisms, targets, and therapeutics.* Oxford university press.
Pedrotty, D.M., Morley, M.P. and Cappola, T.P., 2012. Transcriptomic biomarkers of cardiovascular disease. *Progress in Cardiovascular Diseases,* 55 (1), 64-69.
Peng, H. and Tian, Z., 2017. Natural killer cell memory: progress and implications. *Frontiers in Immunology,* 8, 1143.
Peng, W., et al., 2016. Loss of PTEN Promotes Resistance to T Cell-Mediated Immunotherapy. *Cancer Discovery,* 6 (2), 202-216.
Perfilyeva, Y., et al., 2015. Hyaluronan-binding T regulatory cells in peripheral blood of breast cancer patients. *J Clin Cell Immunol,* 6 (286), 2.
Perry, J.S., et al., 2014. Distinct contributions of Aire and antigen-presenting-cell subsets to the generation of self-tolerance in the thymus. *Immunity,* 41 (3), 414-426.
Petricoin, E.F., et al., 2002. Clinical proteomics: translating benchside promise into bedside reality. *Nature Reviews Drug Discovery,* 1 (9), 683.
Petzold, C., et al., 2014. Fluorochrome-based definition of naturally occurring Foxp3 regulatory T cells of intra-and extrathymic origin. *European Journal of Immunology,* 44 (12), 3632-3645.
Picotti, P. and Aebersold, R., 2012a. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods,* 9 (6), 555.
Picotti, P. and Aebersold, R., 2012b. Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions. *Nature Methods,* 9 (6), 555.
Pino, L.K., et al., 2017. The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrometry Reviews,* .
Pinu, F.R., et al., 2019. Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites,* 9 (4), 76.
Pitt, J.J., 2009. Principles and applications of liquid chromatography-mass spectrometry in clinical biochemistry. *The Clinical Biochemist.Reviews,* 30 (1), 19-34.
Pluskal, T., et al., 2010. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics,* 11 (1), 395.
Pobezinsky, L.A., et al., 2012. Clonal deletion and the fate of autoreactive thymocytes that survive negative selection. *Nature Immunology,* 13 (6), 569.
Podwojski, K., et al., 2010. Peek a peak: a glance at statistics for quantitative label-free proteomics. *Expert Review of Proteomics,* 7 (2), 249-261.
Porta, C., Paglino, C. and Mosca, A., 2014. Targeting PI3K/Akt/mTOR signaling in cancer. *Frontiers in Oncology,* 4, 64.
Prendergast, G., 2008. Immune escape as a fundamental trait of cancer: focus on IDO. *Oncogene,* 27 (28), 3889.
Procaccini, C., et al., 2016. The proteomic landscape of human ex vivo regulatory and conventional T cells reveals specific metabolic requirements. *Immunity,* 44 (2), 406-421.
Proserpio, V., et al., 2016. Single-cell analysis of CD4 T-cell differentiation reveals three major cell states and progressive acceleration of proliferation. *Genome Biology,* 17 (1), 103.
Qian, X., et al., 2011. Generation of human regulatory T cells de novo with suppressive function prevent xenogeneic graft versus host disease. *International Immunopharmacology,* 11 (5), 630-637.
Qu, P., Wang, L. and Lin, P.C., 2016. Expansion and functions of myeloid-derived suppressor cells in the tumor microenvironment. *Cancer Letters,* 380 (1), 253-256.
Quezada, H., et al., 2017. Omics-based biomarkers: current status and potential use in the clinic. *Boletín Médico Del Hospital Infantil De México (English Edition),* 74 (3), 219-226.
Quezada, S.A., et al., 2011. Shifting the equilibrium in cancer immunoediting: from tumor tolerance to eradication. *Immunological Reviews,* 241 (1), 104-118.
Rakhra, K., et al., 2010. CD4 T cells contribute to the remodeling of the microenvironment required for sustained tumor regression upon oncogene inactivation. *Cancer Cell,* 18 (5), 485-498.

Ramos-Fernandez, A., Lopez-Ferrer, D. and Vazquez, J., 2007. Improved method for differential expression proteomics using trypsin-catalyzed 18O labeling with a correction for labeling efficiency. *Molecular & Cellular Proteomics : MCP,* 6 (7), 1274-1286.
Ramos-Vara, J., 2005. Technical aspects of immunohistochemistry. *Veterinary Pathology,* 42 (4), 405-426.
Ratan, Z.A., et al., 2017. Application of Fluorescence In Situ Hybridization (FISH) Technique for the Detection of Genetic Aberration in Medical Science. *Cureus,* 9 (6), e1325.
Rauniyar, N., 2015. Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry. *International Journal of Molecular Sciences,* 16 (12), 28566-28581.
Read, S., Malmstrom, V. and Powrie, F., 2000. Cytotoxic T lymphocyte-associated antigen 4 plays an essential role in the function of CD25(+)CD4(+) regulatory cells that control intestinal inflammation. *The Journal of Experimental Medicine,* 192 (2), 295-302.
Rech, A.J., et al., 2012. CD25 blockade depletes and selectively reprograms regulatory T cells in concert with immunotherapy in cancer patients. *Science Translational Medicine,* 4 (134), 134ra62.
Reis, B.S., et al., 2013. Mutual expression of the transcription factors Runx3 and ThPOK regulates intestinal CD4 T cell immunity. *Nature Immunology,* 14 (3), 271.
Reis, S.T.d., et al., 2011. Tgf-β1 expression as a biomarker of poor prognosis in prostate cancer. *Clinics,* 66 (7), 1143-1147.
Ren, X., et al., 2007. Involvement of cellular death in TRAIL/DR5-dependent suppression induced by CD4 CD25 regulatory T cells. *Cell Death and Differentiation,* 14 (12), 2076.
Ren, L., et al., 2016. Hypoxia-induced CCL28 promotes recruitment of regulatory T cells and tumor growth in liver cancer. *Oncotarget,* 7 (46), 75763-75773.
Reynolds, K.J., Yao, X. and Fenselau, C., 2002. Proteolytic 18O labeling for comparative proteomics: evaluation of endoprotease Glu-C as the catalytic agent. *Journal of Proteome Research,* 1 (1), 27-33.
Ribatti, D. and Crivellato, E., 2012. Mast cells, angiogenesis, and tumour growth. *Biochimica Et Biophysica Acta (BBA)-Molecular Basis of Disease,* 1822 (1), 2-8.
Ribatti, D., et al., 2003. Tumor vascularity and tryptase-positive mast cells correlate with a poor prognosis in melanoma. *European Journal of Clinical Investigation,* 33 (5), 420-425.
Roane, B.M., Arend, R.C. and Birrer, M.J., 2019. Targeting the Transforming Growth Factor-Beta Pathway in Ovarian Cancer. *Cancers,* 11 (5), 668.
Robinson, R.T. and Gorham, J.D., 2007. TGF-beta 1 regulates antigen-specific CD4+ T cell responses in the periphery. *Journal of Immunology (Baltimore, Md.: 1950),* 179 (1), 71-79.
Roitt, I., Brostoff, J. and Male, D., 2001. Immunology 6th eds. *USA, Mosby,* , 323-383.
Rossi, S.W., et al., 2007. RANK signals from CD4(+)3(-) inducer cells regulate development of Aire-expressing epithelial cells in the thymic medulla. *The Journal of Experimental Medicine,* 204 (6), 1267-1272.
Röst, H.L., et al., 2014. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nature Biotechnology,* 32 (3), 219.
Round, J.L. and Mazmanian, S.K., 2010. Inducible Foxp3+ regulatory T-cell development by a commensal bacterium of the intestinal microbiota. *Proceedings of the National Academy of Sciences of the United States of America,* 107 (27), 12204-12209.
Ruan, Q., et al., 2009. Development of Foxp3 regulatory T cells is driven by the c-Rel enhanceosome. *Immunity,* 31 (6), 932-940.
Rudd, C.E., Taylor, A. and Schneider, H., 2009. CD28 and CTLA-4 coreceptor expression and signal transduction. *Immunological Reviews,* 229 (1), 12-26.
Russo, A., et al., 2015. Genomic instability: Crossing pathways at the origin of structural and numerical chromosome changes. *Environmental and Molecular Mutagenesis,* 56 (7), 563-580.
Sabroe, I., et al., 2003. Selective roles for Toll-like receptor (TLR)2 and TLR4 in the regulation of neutrophil activation and life span. *Journal of Immunology (Baltimore, Md.: 1950),* 170 (10), 5268-5275.
Sager, M., et al., 2015. Transcriptomics in cancer diagnostics: Developments in technology, clinical research and commercialization. *Expert Review of Molecular Diagnostics,* 15 (12), 1589-1603.
Sakaguchi, S., et al., 2008. Regulatory T cells and immune tolerance. *Cell,* 133 (5), 775-787.
Samanta, D. and Datta, P.K., 2012. Alterations in the Smad pathway in human cancers. *Frontiers in Bioscience (Landmark Edition),* 17, 1281-1293.
Samstein, R.M., et al., 2012. Extrathymic generation of regulatory T cells in placental mammals mitigates maternal-fetal conflict. *Cell,* 150 (1), 29-38.

Sánchez-Paulete, A.R., et al., 2017. Antigen cross-presentation and T-cell cross-priming in cancer immunology and immunotherapy. *Annals of Oncology,* 28 (suppl_12), xii44-xii55.
Sanjabi, S., Mosaheb, M.M. and Flavell, R.A., 2009. Opposing effects of TGF-β and IL-15 cytokines control the number of short-lived effector CD8 T cells. *Immunity,* 31 (1), 131-144.
Sanjabi, S., Oh, S.A. and Li, M.O., 2017. Regulation of the Immune Response by TGF-beta: From Conception to Autoimmunity and Infection. *Cold Spring Harbor Perspectives in Biology,* 9 (6), 10.1101/cshperspect.a022236.
Sarris, M., et al., 2008. Neuropilin-1 expression on regulatory T cells enhances their interactions with dendritic cells during antigen recognition. *Immunity,* 28 (3), 402-413.
Sato, E., et al., 2005. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. *Proceedings of the National Academy of Sciences of the United States of America,* 102 (51), 18538-18543.
Satori, C.P., Kostal, V. and Arriaga, E.A., 2012. Review on recent advances in the analysis of isolated organelles. *Analytica Chimica Acta,* 753, 8-18.
Savaryn, J.P., et al., 2013. The emergence of top-down proteomics in clinical research. *Genome Medicine,* 5 (6), 53.
Savitski, M.M., et al., 2013. Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *Journal of Proteome Research,* 12 (8), 3586-3598.
Sayour, E.J., et al., 2015. Increased proportion of FoxP3 regulatory T cells in tumor infiltrating lymphocytes is associated with tumor recurrence and reduced survival in patients with glioblastoma. *Cancer Immunology, Immunotherapy,* 64 (4), 419-427.
Schlenner, S.M., et al., 2012. Smad3 binding to the foxp3 enhancer is dispensable for the development of regulatory T cells with the exception of the gut. *The Journal of Experimental Medicine,* 209 (9), 1529-1535.
Schlitzer, A., et al., 2015. Identification of cDC1-and cDC2-committed DC progenitors reveals early lineage priming at the common DC progenitor stage in the bone marrow. *Nature Immunology,* 16 (7), 718.
Schmidlin, T., et al., 2016. Assessment of SRM, MRM3, and DIA for the targeted analysis of phosphorylation dynamics in non-small cell lung cancer. *Proteomics,* 16 (15-16), 2193-2205.
Schmidt, A., et al., 2016. Comparative analysis of protocols to induce human CD4 Foxp3 regulatory T cells by combinations of IL-2, TGF-beta, retinoic acid, rapamycin and butyrate. *PloS One,* 11 (2), e0148474.
Schmidt, A., et al., 2018. Time-resolved transcriptome and proteome landscape of human regulatory T cell (Treg) differentiation reveals novel regulators of FOXP3. *BMC Biology,* 16 (1), 47.
Schmitt, E.G. and Williams, C.B., 2013. Generation and function of induced regulatory T cells. *Frontiers in Immunology,* 4, 152.
Schmitt, N., et al., 2014. The cytokine TGF-β co-opts signaling via STAT3-STAT4 to promote the differentiation of human T FH cells. *Nature Immunology,* 15 (9), 856.
Schreiber, L., et al., 2014. The Treg-specific demethylated region stabilizes Foxp3 expression independently of NF-κB signaling. *PLoS One,* 9 (2), e88318.
Schreiber, R.D., Old, L.J. and Smyth, M.J., 2011. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. *Science,* 331 (6024), 1565-1570.
Schwartz, R.H., 2003. T cell anergy. *Annual Review of Immunology,* 21 (1), 305-334.
Schwarz, B.A. and Bhandoola, A., 2006. Trafficking from the bone marrow to the thymus: a prerequisite for thymopoiesis. *Immunological Reviews,* 209 (1), 47-57.
Seo, J. and Lee, K.J., 2004. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *Journal of Biochemistry and Molecular Biology,* 37 (1), 35-44.
Serrano, A.E., et al., 2011. Interleukin 10 decreases MICA expression on melanoma cell surface. *Immunology and Cell Biology,* 89 (3), 447-457.
Seth, R.B., et al., 2005. Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-κB and IRF3. *Cell,* 122 (5), 669-682.
Shang, B., et al., 2015a. Prognostic value of tumor-infiltrating FoxP3 regulatory T cells in cancers: a systematic review and meta-analysis. *Scientific Reports,* 5, 15179.
Shang, B., et al., 2015b. Prognostic value of tumor-infiltrating FoxP3 regulatory T cells in cancers: a systematic review and meta-analysis. *Scientific Reports,* 5, 15179.
Shankaran, V., et al., 2001. IFNgamma and lymphocytes prevent primary tumour development and shape tumour immunogenicity Nature.
Sharma, M., Kaveri, S.V. and Bayry, J., 2013. Th17 cells, pathogenic or not? TGF-β3 imposes the embargo. *Cellular & Molecular Immunology,* 10 (2), 101.

Sheen-Chen, S., et al., 2001. Serum levels of transforming growth factor β1 in patients with breast cancer. *Archives of Surgery,* 136 (8), 937-940.
Shi, J., et al., 2014. Perioperative changes in peripheral regulatory B cells of patients with esophageal cancer. *Molecular Medicine Reports,* 10 (3), 1525-1530.
Shi, Y., et al., 2004. The role of liquid chromatography in proteomics. *Journal of Chromatography A,* 1053 (1-2), 27-36.
Shi, Y. and Massagué, J., 2003. Mechanisms of TGF-β signaling from cell membrane to the nucleus. *Cell,* 113 (6), 685-700.
Shi, L.Z., et al., 2011. HIF1alpha-dependent glycolytic pathway orchestrates a metabolic checkpoint for the differentiation of TH17 and Treg cells. *The Journal of Experimental Medicine,* 208 (7), 1367-1376.
Shirabe, K., et al., 2012. Role of tumor-associated macrophages in the progression of hepatocellular carcinoma. *Surgery Today,* 42 (1), 1-7.
Shitara, K. and Nishikawa, H., 2018. Regulatory T cells: a potential target in cancer immunotherapy. *Annals of the New York Academy of Sciences,* 1417 (1), 104-115.
Shugang, X., et al., 2016. Prognostic value of SMAD4 in pancreatic cancer: A meta-analysis. *Translational Oncology,* 9 (1), 1-7.
Sica, A., et al., 2008. Macrophage polarization in tumour progression. *In: Seminars in cancer biology,* Elsevier, pp. 349-355.
Sica, A., et al., 2007. Targeting tumour-associated macrophages. *Expert Opinion on Therapeutic Targets,* 11 (9), 1219-1229.
Sica, A., et al., 2006. Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: potential targets of anti-cancer therapy. *European Journal of Cancer,* 42 (6), 717-727.
Simpson, D.C. and Smith, R.D., 2005. Combining capillary electrophoresis with mass spectrometry for applications in proteomics. *Electrophoresis,* 26 (7-8), 1291-1305.
Singer, A., Adoro, S. and Park, J., 2008. Lineage fate and intense debate: myths, models and mechanisms of CD4-versus CD8-lineage choice. *Nature Reviews Immunology,* 8 (10), 788.
Singh, N.J., Bando, J.K. and Schwartz, R.H., 2012. Subsets of nonclonal neighboring CD4 T cells specifically regulate the frequency of individual antigen-reactive T cells. *Immunity,* 37 (4), 735-746.
Sinha, P., et al., 2007. Prostaglandin E2 promotes tumor progression by inducing myeloid-derived suppressor cells. *Cancer Research,* 67 (9), 4507-4513.
Smigiel, K.S., et al., 2014. CCR7 provides localized access to IL-2 and defines homeostatically distinct regulatory T cell subsets. *The Journal of Experimental Medicine,* 211 (1), 121-136.
Smith-Garvin, J.E., Koretzky, G.A. and Jordan, M.S., 2009. T cell activation. *Annual Review of Immunology,* 27, 591-619.
Smolka, M.B., et al., 2001. Optimization of the isotope-coded affinity tag-labeling procedure for quantitative proteome analysis. *Analytical Biochemistry,* 297 (1), 25-31.
Smyth, M.J., Dunn, G.P. and Schreiber, R.D., 2006. Cancer immunosurveillance and immunoediting: the roles of immunity in suppressing tumor development and shaping tumor immunogenicity. *Advances in Immunology,* 90, 1-50.
Solinas, G., et al., 2009. Tumor-associated macrophages (TAM) as major players of the cancer-related inflammation. *Journal of Leukocyte Biology,* 86 (5), 1065-1073.
Song, M.S., Salmena, L. and Pandolfi, P.P., 2012. The functions and regulation of the PTEN tumour suppressor. *Nature Reviews Molecular Cell Biology,* 13 (5), 283.
Song, Q., et al., 2017. Techniques for detecting chromosomal aberrations in myelodysplastic syndromes. *Oncotarget,* 8 (37), 62716-62729.
Sonoda, Y., et al., 2010. Proteolipid protein 2 is associated with melanoma metastasis. *Oncology Reports,* 23 (2), 371-376.
Soucek, L., et al., 2007. Mast cells are required for angiogenesis and macroscopic expansion of Myc-induced pancreatic islet tumors. *Nature Medicine,* 13 (10), 1211.
Spits, H., 2002. Development of αβ T cells in the human thymus. *Nature Reviews Immunology,* 2 (10), 760.
Spranger, S., Bao, R. and Gajewski, T.F., 2015. Melanoma-intrinsic β-catenin signalling prevents anti-tumour immunity. *Nature,* 523 (7559), 231.
Sprent, J. and Kishimoto, H., 2002. The thymus and negative selection. *Immunological Reviews,* 185 (1), 126-135.
Starling, S., 2017. MHC molecules: Immune editing shapes the cancer landscape. *Nature Reviews Immunology,* 17 (12), 729.

Starr, T.K., Jameson, S.C. and Hogquist, K.A., 2003. Positive and negative selection of T cells. *Annual Review of Immunology,* 21 (1), 139-176.
Steinman, R.M., 2006. Linking innate to adaptive immunity through dendritic cells. *In: Novartis Foundation symposium,* Wiley Online Library, pp. 101.
Storms, H.F., et al., 2004. Capillary isoelectric focusing-mass spectrometry for shotgun approach in proteomics. *Electrophoresis,* 25 (20), 3461-3467.
Strainic, M.G., et al., 2013. Absence of signaling into CD4 cells via C3aR and C5aR enables autoinductive TGF-β1 signaling and induction of Foxp3 regulatory T cells. *Nature Immunology,* 14 (2), 162.
Strauss, L., et al., 2007. A unique subset of CD4+CD25highFoxp3+ T cells secreting interleukin-10 and transforming growth factor-beta1 mediates suppression in the tumor microenvironment. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 13 (15 Pt 1), 4345-4354.
Stritesky, G.L., Jameson, S.C. and Hogquist, K.A., 2012. Selection of self-reactive T cells in the thymus. *Annual Review of Immunology,* 30, 95-114.
Stritesky, G.L., Yeh, N. and Kaplan, M.H., 2008. IL-23 promotes maintenance but not commitment to the Th17 lineage. *Journal of Immunology (Baltimore, Md.: 1950),* 181 (9), 5948-5955.
Strouch, M.J., et al., 2010. Crosstalk between mast cells and pancreatic cancer cells contributes to pancreatic tumor progression. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 16 (8), 2257-2265.
Su, S., et al., 2017a. Blocking the recruitment of naive CD4 T cells reverses immunosuppression in breast cancer. *Cell Research,* 27 (4), 461.
Su, S., et al., 2017b. Blocking the recruitment of naive CD4 T cells reverses immunosuppression in breast cancer. *Cell Research,* 27 (4), 461.
Suciu-Foca, N., Berloco, P. and Cortesini, R., 2009. Tolerogenic dendritic cells in cancer, transplantation, and autoimmune diseases. *Human Immunology,* 70 (5), 277-280.
Sun, L., et al., 2016. Capillary zone electrophoresis for bottom-up analysis of complex proteomes. *Proteomics,* 16 (2), 188-196.
Sun, S., et al., 2014. PD-1 immune cell infiltration inversely correlates with survival of operable breast cancer patients. *Cancer Immunology, Immunotherapy,* 63 (4), 395-406.
Sun, T., et al., 2014. Inhibition of tumor angiogenesis by interferon-γ by suppression of tumor-associated macrophage differentiation. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics,* 21 (5), 227-235.
Sun, W., et al., 2017. A positive-feedback loop between tumour infiltrating activated Treg cells and type 2-skewed macrophages is essential for progression of laryngeal squamous cell carcinoma. *British Journal of Cancer,* 117 (11), 1631.
Sun, L., et al., 2017. Clinicopathologic and prognostic significance of regulatory T cells in patients with hepatocellular carcinoma: a meta-analysis. *Oncotarget,* 8 (24), 39658-39672.
Sun, S.P., et al., 2007. Serum transforming growth factor-beta1 level reflects disease status in patients with esophageal carcinoma after radiotherapy. *World Journal of Gastroenterology,* 13 (39), 5267-5272.
Sung, J.L., Lin, J.T. and Gorham, J.D., 2003. CD28 co-stimulation regulates the effect of transforming growth factor-β1 on the proliferation of naïve CD4 T cells. *International Immunopharmacology,* 3 (2), 233-245.
Surinova, S., et al., 2010. On the development of plasma protein biomarkers. *Journal of Proteome Research,* 10 (1), 5-16.
Suthar, M.S., et al., 2010. IPS-1 is essential for the control of West Nile virus infection and immunity. *PLoS Pathogens,* 6 (2), e1000757.
Suthar, M.S., et al., 2012. The RIG-I-like receptor LGP2 controls CD8 T cell survival and fitness. *Immunity,* 37 (2), 235-248.
Swan, A.L., et al., 2013. Application of machine learning to proteomics data: classification and biomarker identification in postgenomics biology. *Omics: A Journal of Integrative Biology,* 17 (12), 595-610.
Sweis, R.F., et al., 2016. Molecular Drivers of the Non-T-cell-Inflamed Tumor Microenvironment in Urothelial Bladder Cancer. *Cancer Immunology Research,* 4 (7), 563-568.
Tadokoro, C.E., et al., 2006. Regulatory T cells inhibit stable contacts between CD4+ T cells and dendritic cells in vivo. *The Journal of Experimental Medicine,* 203 (3), 505-511.
Tai, X., et al., 2013. Foxp3 transcription factor is proapoptotic and lethal to developing regulatory T cells unless counterbalanced by cytokine survival signals. *Immunity,* 38 (6), 1116-1128.

Takaba, H. and Takayanagi, H., 2017. The mechanisms of T cell selection in the thymus. *Trends in Immunology,* 38 (11), 805-816.
Takahama, Y., 2006. Journey through the thymus: stromal guides for T-cell development and selection. *Nature Reviews Immunology,* 6 (2), 127.
Takahama, Y., et al., 2017. Generation of diversity in thymic epithelial cells. *Nature Reviews Immunology,* 17 (5), 295.
Takeuchi, Y. and Nishikawa, H., 2016. Roles of regulatory T cells in cancer immunity. *International Immunology,* 28 (8), 401-409.
Takimoto, T., et al., 2010. Smad2 and Smad3 are redundantly essential for the TGF-beta-mediated regulation of regulatory T plasticity and Th1 development. *Journal of Immunology (Baltimore, Md.: 1950),* 185 (2), 842-855.
Tanaka, A. and Sakaguchi, S., 2017. Regulatory T cells in cancer immunotherapy. *Cell Research,* 27 (1), 109.
Tang, Y., et al., 2014. An increased abundance of tumor-infiltrating regulatory T cells is correlated with the progression and prognosis of pancreatic ductal adenocarcinoma. *PloS One,* 9 (3), e91551.
Taniuchi, I., 2018. CD4 helper and CD8 cytotoxic T cell differentiation. *Annual Review of Immunology,* 36, 579-601.
Tao, H., et al., 2012. Prognostic potential of FOXP3 expression in non-small cell lung cancer cells combined with tumor-infiltrating regulatory T cells. *Lung Cancer,* 75 (1), 95-101.
Tcyganov, E., et al., 2018. Plasticity of myeloid-derived suppressor cells in cancer. *Current Opinion in Immunology,* 51, 76-82.
Teleman, J., et al., 2014. DIANA—algorithmic improvements for analysis of data-independent acquisition MS data. *Bioinformatics,* 31 (4), 555-562.
Teng, M.W., et al., 2008. Immune-mediated dormancy: an equilibrium with cancer. *Journal of Leukocyte Biology,* 84 (4), 988-993.
Teng, M.W., et al., 2012. Opposing roles for IL-23 and IL-12 in maintaining occult cancer in an equilibrium state. *Cancer Research,* 72 (16), 3987-3996.
Theisen, A. and Shaffer, L.G., 2010. Disorders caused by chromosome abnormalities. *The Application of Clinical Genetics,* 3, 159-174.
Theodorescu, D. and Mischak, H., 2007. Mass spectrometry based proteomics in urine biomarker discovery. *World Journal of Urology,* 25 (5), 435-443.
Thiault, N., et al., 2015. Peripheral regulatory T lymphocytes recirculating to the thymus suppress the development of their precursors. *Nature Immunology,* 16 (6), 628.
Thomas, S.N. and Zhang, H., 2016. *Targeted Proteomic Assays for the Verification of Global Proteomics Insights,* .
Thornton, A.M., et al., 2010. Expression of Helios, an Ikaros transcription factor family member, differentiates thymic-derived from peripherally induced Foxp3+ T regulatory cells. *Journal of Immunology (Baltimore, Md.: 1950),* 184 (7), 3433-3441.
Thuring, C., et al., 2015. HLA class I is most tightly linked to levels of tapasin compared with other antigen-processing proteins in glioblastoma. *British Journal of Cancer,* 113 (6), 952.
Toby, T.K., Fornelli, L. and Kelleher, N.L., 2016. Progress in top-down proteomics and the analysis of proteoforms. *Annual Review of Analytical Chemistry,* 9, 499-519.
Togashi, Y., Shitara, K. and Nishikawa, H., 2019. Regulatory T cells in cancer immunosuppression—implications for anticancer therapy. *Nature Reviews Clinical Oncology,* , 1.
Tokunaga, H., et al., 1999. Decreased expression of transforming growth factor beta receptor type I is associated with poor prognosis in bladder transitional cell carcinoma patients. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 5 (9), 2520-2525.
Tone, Y., et al., 2008. Smad3 and NFAT cooperate to induce Foxp3 expression through its enhancer. *Nature Immunology,* 9 (2), 194.
Topfer, K., et al., 2011. Tumor evasion from T cell surveillance. *Journal of Biomedicine & Biotechnology,* 2011, 918471.
Tosi, M.F., 2005. Innate immune responses to infection. *Journal of Allergy and Clinical Immunology,* 116 (2), 241-249.
Toso, A., et al., 2014. Enhancing chemotherapy efficacy in Pten-deficient prostate tumors by activating the senescence-associated antitumor immunity. *Cell Reports,* 9 (1), 75-89.
Tran Janco, J.M., et al., 2015. Tumor-infiltrating dendritic cells in cancer pathogenesis. *Journal of Immunology (Baltimore, Md.: 1950),* 194 (7), 2985-2991.
Tran, D.Q., 2011. TGF-β: the sword, the wand, and the shield of FOXP3 regulatory T cells. *Journal of Molecular Cell Biology,* 4 (1), 29-37.

Trujillo, J.A., et al., 2018. T Cell-Inflamed versus Non-T Cell-Inflamed Tumors: A Conceptual Framework for Cancer Immunotherapy Drug Development and Combination Therapy Selection. *Cancer Immunology Research,* 6 (9), 990-1000.
Tsiatsiani, L. and Heck, A.J., 2015. Proteomics beyond trypsin. *The FEBS Journal,* 282 (14), 2612-2626.
Tsutsui, H., et al., 2011. Biomarker discovery in biological specimens (plasma, hair, liver and kidney) of diabetic mice based upon metabolite profiling using ultra-performance liquid chromatography with electrospray ionization time-of-flight mass spectrometry. *Clinica Chimica Acta,* 412 (11-12), 861-872.
Tu, W.H., et al., 2003. The loss of TGF-beta signaling promotes prostate cancer metastasis. *Neoplasia (New York, N.Y.),* 5 (3), 267-277.
Tuomela, S., et al., 2016. Comparative analysis of human and mouse transcriptomes of Th17 cell priming. *Oncotarget,* 7 (12), 13416-13428.
Tuomela, S., et al., 2012. Identification of early gene expression changes during human Th17 cell differentiation. *Blood,* 119 (23), e151-60.
Tyers, M. and Mann, M., 2003. From genomics to proteomics. *Nature,* 422 (6928), 193.
Ueda, Y., et al., 2010. Serum biomarkers for early detection of gynecologic cancers. *Cancers,* 2 (2), 1312-1327.
Unsold, C., et al., 2001. Latent TGF-beta binding protein LTBP-1 contains three potential extracellular matrix interacting domains. *Journal of Cell Science,* 114 (Pt 1), 187-197.
Urbanczyk-Wochniak, E., et al., 2003. Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO Reports,* 4 (10), 989-993.
Vaeth, M., et al., 2012. Dependence on nuclear factor of activated T-cells (NFAT) levels discriminates conventional T cells from Foxp3+ regulatory T cells. *Proceedings of the National Academy of Sciences of the United States of America,* 109 (40), 16258-16263.
Vahl, J.C., et al., 2014. Continuous T cell receptor signals maintain a functional regulatory T cell pool. *Immunity,* 41 (5), 722-736.
Valzasina, B., et al., 2006. Tumor-induced expansion of regulatory T cells by conversion of CD4+CD25- lymphocytes is thymus and proliferation independent. *Cancer Research,* 66 (8), 4488-4495.
van den Ham, H., et al., 2019. Proteomic Profiling of Mouse Helper T Cell Differentiation. *Proteomics,* 19 (7), 1800045.
van den Ham, H., et al., 2013. Early divergence of Th1 and Th2 transcriptomes involves a small core response and sets of transiently expressed genes. *European Journal of Immunology,* 43 (4), 1074-1084.
Van Rensburg, I.C. and Loxton, A.G., 2015. Transcriptomics: the key to biomarker discovery during tuberculosis? *Biomarkers in Medicine,* 9 (5), 483-495.
Vargas-Rondón, N., Villegas, V. and Rondón-Lagos, M., 2017. The role of chromosomal instability in cancer and therapeutic responses. *Cancers,* 10 (1), 4.
Vesely, M.D., et al., 2011. Natural innate and adaptive immunity to cancer. *Annual Review of Immunology,* 29, 235-271.
Vesely, M.D. and Schreiber, R.D., 2013. Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. *Annals of the New York Academy of Sciences,* 1284, 1-5.
Vignali, D.A., Collison, L.W. and Workman, C.J., 2008. How regulatory T cells work. *Nature Reviews Immunology,* 8 (7), 523.
Vitale, M., et al., 2014. Effect of tumor cells and tumor microenvironment on NK-cell function. *European Journal of Immunology,* 44 (6), 1582-1592.
Vogelsang, P., et al., 2009a. Levels of dendritic cell populations and regulatory T cells vary significantly between two commonly used mouse strains. *Scandinavian Journal of Immunology,* 70 (6), 541-546.
Vogelsang, P., et al., 2009b. Levels of dendritic cell populations and regulatory T cells vary significantly between two commonly used mouse strains. *Scandinavian Journal of Immunology,* 70 (6), 541-546.
von Rahden, B.H., et al., 2006. Overexpression of TGF-β1 in esophageal (Barrett's) adenocarcinoma is associated with advanced stage of disease and poor prognosis. *Molecular Carcinogenesis: Published in Cooperation with the University of Texas MD Anderson Cancer Center,* 45 (10), 786-794.
Voskuil, J., 2014. Commercial antibodies and their validation. *F1000Research,* 3, 232.
Wagner, S., et al., 2019. A parsimonious 3-gene signature predicts clinical outcomes in an acute myeloid leukemia multicohort study. *Blood Advances,* 3 (8), 1330-1346.
Waight, J.D., et al., 2013. Myeloid-derived suppressor cell development is regulated by a STAT/IRF-8 axis. *The Journal of Clinical Investigation,* 123 (10), 4464-4478.

Waldhauer, I. and Steinle, A., 2008. NK cells and cancer immunosurveillance. *Oncogene,* 27 (45), 5932.
Wang, J., et al., 2017a. Tumor cells induced-M2 macrophage favors accumulation of Treg in nasopharyngeal carcinoma. *Int.J.Clin.Exp.Pathol,* 10, 8389-8401.
Wang, J., et al., 2017b. In-depth method assessments of differentially expressed protein detection for shotgun proteomics data with missing values. *Scientific Reports,* 7 (1), 3367.
Wang, M., et al., 2008. Label-free mass spectrometry-based protein quantification technologies in proteomic analysis. *Briefings in Functional Genomics and Proteomics,* 7 (5), 329-339.
Wang, R., et al., 2012. GARP regulates the bioavailability and activation of TGFβ. *Molecular Biology of the Cell,* 23 (6), 1129-1139.
Wang, K., Liu, J. and Li, J., 2018. IL-35-producing B cells in gastric cancer patients. *Medicine,* 97 (19), e0710.
Wang, X., et al., 2016. PD-L1 expression in human cancers and its association with clinical outcomes. *OncoTargets and Therapy,* 9, 5023-5039.
Weber, G., et al., 2004. Efficient separation and analysis of peroxisomal membrane proteins using free-flow isoelectric focusing. *Electrophoresis,* 25 (12), 1735-1747.
Weber, S.M., et al., 2008. Role of Transforming Growth Factor-β in the Pathogenesis of Human Head and Neck Squamous Cell Carcinoma. *In:* Role of Transforming Growth Factor-β in the Pathogenesis of Human Head and Neck Squamous Cell Carcinoma. *Transforming Growth Factor-β in Cancer Therapy, Volume II.* Springer, 2008, pp. 21-32.
Wei, X., et al., 2017. Reciprocal expression of IL-35 and IL-10 defines two distinct effector Treg subsets that are required for maintenance of immune tolerance. *Cell Reports,* 21 (7), 1853-1869.
Wendt, M.K., Tian, M. and Schiemann, W.P., 2012. Deconstructing the mechanisms and consequences of TGF-β-induced EMT during cancer progression. *Cell and Tissue Research,* 347 (1), 85-101.
Wherry, E.J., 2011. T cell exhaustion. *Nature Immunology,* 12 (6), 492.
Whiteside, T.L., 2012. What are regulatory T cells (Treg) regulating in cancer and why? *In: Seminars in cancer biology,* Elsevier, pp. 327-334.
Williams, C.B., Yeh, E.S. and Soloff, A.C., 2016. Tumor-associated macrophages: unwitting accomplices in breast cancer malignancy. *NPJ Breast Cancer,* 2, 15025.
Wirtz, S., et al., 2011. Interleukin-35 mediates mucosal immune responses that protect against t-cell–dependent colitis. *Gastroenterology,* 141 (5), 1875-1886.
Wishart, D.S., 2016. Emerging applications of metabolomics in drug discovery and precision medicine. *Nature Reviews Drug Discovery,* 15 (7), 473.
Witt, C.M. and Robey, E.A., 2005. Thymopoiesis in 4 dimensions. *In: Seminars in immunology,* Elsevier, pp. 95-102.
Wong, J., Mathis, D. and Benoist, C., 2007. TCR-based lineage tracing: no evidence for conversion of conventional into regulatory T cells in response to a natural self-antigen in pancreatic islets. *The Journal of Experimental Medicine,* 204 (9), 2039-2045.
Workman, C.J., et al., 2009. The development and function of regulatory T cells. *Cellular and Molecular Life Sciences,* 66 (16), 2603.
Wu, C., et al., 2015. TGF beta1 expression correlates with survival and tumor aggressiveness of prostate cancer. *Annals of Surgical Oncology,* 22 (3), 1587-1593.
Wu, C., et al., 2015. TGF-β1 mediates the radiation response of prostate cancer. *Journal of Molecular Medicine,* 93 (1), 73-82.
Wu, L., 2006. T lineage progenitors: the earliest steps en route to T lymphocytes. *Current Opinion in Immunology,* 18 (2), 121-126.
Wu, X., et al., 2013. Immune microenvironment profiles of tumor immune equilibrium and immune escape states of mouse sarcoma. *Cancer Letters,* 340 (1), 124-133.
Wu, C.C. and MacCoss, M.J., 2002. Shotgun proteomics: tools for the analysis of complex biological systems. *Current Opinion in Molecular Therapeutics,* 4 (3), 242-250.
Wyss, L., et al., 2016. Affinity for self antigen selects T reg cells with distinct functional properties. *Nature Immunology,* 17 (9), 1093.
Xie, F., Smith, R.D. and Shen, Y., 2012. Advanced proteomic liquid chromatography. *Journal of Chromatography A,* 1261, 78-90.
Xie, M., He, C. and Wei, S., 2015. Relationship between Expression of TGF-beta1, Smad2, Smad4 and Prognosis of Patients with Resected Non-small Cell Lung Cancer. *Zhongguo Fei Ai Za Zhi = Chinese Journal of Lung Cancer,* 18 (9), 543-548.
Xiong, B., et al., 2002. Transforming growth factor-beta1 in invasion and metastasis in colorectal cancer. *World Journal of Gastroenterology,* 8 (4), 674-678.

Xu, G., Chakraborty, C. and Lala, P.K., 2003. Reconstitution of Smad3 restores TGF-β response of tissue inhibitor of metalloprotease-1 upregulation in human choriocarcinoma cells. *Biochemical and Biophysical Research Communications,* 300 (2), 383-390.
Xu, L., Kitani, A. and Strober, W., 2010. Molecular mechanisms regulating TGF-β-induced Foxp3 expression. *Mucosal Immunology,* 3 (3), 230.
Yadav, M., Bluestone, J.A. and Stephan, S., 2013. Peripherally induced tregs–role in immune homeostasis and autoimmunity. *Frontiers in Immunology,* 4, 232.
Yamazaki, S., et al., 2007. Dendritic cells are specialized accessory cells along with TGF- for the differentiation of Foxp3+ CD4+ regulatory T cells from peripheral Foxp3 precursors. *Blood,* 110 (13), 4293-4302.
Yan, P., et al., 2016. Reduced Expression of SMAD4 Is Associated with Poor Survival in Colon Cancer. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research,* 22 (12), 3037-3047.
Yang, E., et al., 2014. Both retention and recirculation contribute to long-lived regulatory T-cell accumulation in the thymus. *European Journal of Immunology,* 44 (9), 2712-2720.
Yang, L., et al., 2008. Abrogation of TGFβ signaling in mammary carcinomas recruits Gr-1 CD11b myeloid cells that promote metastasis. *Cancer Cell,* 13 (1), 23-35.
Yang, Z.Z., et al., 2006. Intratumoral CD4+CD25+ regulatory T-cell-mediated suppression of infiltrating CD4+ T cells in B-cell non-Hodgkin lymphoma. *Blood,* 107 (9), 3639-3646.
Yap, N.Y., et al., 2015. Genetic and chromosomal aberrations and their clinical significance in renal neoplasms. *BioMed Research International,* 2015.
Yates Iii, J.R., et al., 2005. Proteomics of organelles and large cellular structures. *Nature Reviews Molecular Cell Biology,* 6 (9), 702.
Yen, H.R., et al., 2009. Tc17 CD8 T cells: functional plasticity and subset diversity. *Journal of Immunology (Baltimore, Md.: 1950),* 183 (11), 7161-7168.
Yue, Q., et al., 2014. The prognostic value of Foxp3 tumor-infiltrating lymphocytes in patients with glioblastoma. *Journal of Neuro-Oncology,* 116 (2), 251-259.
Zafeiris, D., Rutella, S. and Ball, G.R., 2018. An Artificial Neural Network Integrated Pipeline for Biomarker Discovery Using Alzheimer's Disease as a Case Study. *Computational and Structural Biotechnology Journal,* 16, 77-87.
Zarek, P.E., et al., 2008. A2A receptor signaling promotes peripheral tolerance by inducing T-cell anergy and the generation of adaptive regulatory T cells. *Blood,* 111 (1), 251-259.
Zhang, A., et al., 2015. Metabolomics for biomarker discovery: moving to the clinic. *BioMed Research International,* 2015.
Zhang, G., et al., 2010. Protein quantitation using mass spectrometry. *In:* Protein quantitation using mass spectrometry. *Computational biology.* Springer, 2010, pp. 211-222.
Zhang, R., et al., 2001. Fractionation of isotopically labeled peptides in quantitative proteomics. *Analytical Chemistry,* 73 (21), 5142-5149.
Zhang, Y., et al., 2013. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews,* 113 (4), 2343-2394.
Zhang, Y., et al., 2015. The use of variable Q1 isolation windows improves selectivity in LC–SWATH–MS acquisition. *Journal of Proteome Research,* 14 (10), 4359-4371.
Zhang, Y., Gallastegui, N. and Rosenblatt, J.D., 2015. Regulatory B cells in anti-tumor immunity. *International Immunology,* 27 (10), 521-530.
Zhang, Y., et al., 2014. Th1/Th2 cell's function in immune system. *In:* Th1/Th2 cell's function in immune system. *T Helper Cell Differentiation and Their Function.* Springer, 2014, pp. 45-65.
Zhang, Z., et al., 2016. Interconversion of peptide mass spectral libraries derivatized with iTRAQ or TMT labels. *Journal of Proteome Research,* 15 (9), 3180-3187.
Zhang, Y., Alexander, P.B. and Wang, X.F., 2017. TGF-beta Family Signaling in the Control of Cell Proliferation and Survival. *Cold Spring Harbor Perspectives in Biology,* 9 (4), 10.1101/cshperspect.a022145.
Zhang, Y.E., 2017. Non-Smad Signaling Pathways of the TGF-beta Family. *Cold Spring Harbor Perspectives in Biology,* 9 (2), 10.1101/cshperspect.a022129.
Zhang, Z., et al., 2017. Activation and Functional Specialization of Regulatory T Cells Lead to the Generation of Foxp3 Instability. *Journal of Immunology (Baltimore, Md.: 1950),* 198 (7), 2612-2625.
Zhao, H., Liao, X. and Kang, Y., 2017. Tregs: where we are and what comes next? *Frontiers in Immunology,* 8, 1578.

Zhao, J., et al., 2016. Clinical and prognostic significance of serum transforming growth factor-beta1 levels in patients with pancreatic ductal adenocarcinoma. *Brazilian Journal of Medical and Biological Research,* 49 (8).
Zhao, Y., et al., 2017. Regulatory B cells induced by pancreatic cancer cell-derived interleukin-18 promote immune tolerance via the PD-1/PD-L1 pathway. *Oncotarget,* 9 (19), 14803-14814.
Zheng, Y., et al., 2010a. Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. *Nature,* 463 (7282), 808.
Zheng, Y., et al., 2010b. Role of conserved non-coding DNA elements in the Foxp3 gene in regulatory T-cell fate. *Nature,* 463 (7282), 808.
Zheng, X., et al., 2017. Prognostic role of tumor-infiltrating lymphocytes in gastric cancer: a meta-analysis. *Oncotarget,* 8 (34), 57386-57398.
Zhong, Z., Sanchez-Lopez, E. and Karin, M., 2016. Autophagy, inflammation, and immunity: a troika governing cancer and its treatment. *Cell,* 166 (2), 288-298.
Zhou, J., et al., 2014. Enhanced frequency and potential mechanism of B regulatory cells in patients with lung cancer. *Journal of Translational Medicine,* 12 (1), 304.
Zhou, Y., et al., 2017. Prognostic value of tumor-infiltrating Foxp3+ regulatory T cells in patients with breast cancer: a meta-analysis. *Journal of Cancer,* 8 (19), 4098-4105.
Zhu, J., et al., 2009. Down-regulation of Gfi-1 expression by TGF-beta is important for differentiation of Th17 and CD103+ inducible regulatory T cells. *The Journal of Experimental Medicine,* 206 (2), 329-341.
Zitvogel, L., Tesniere, A. and Kroemer, G., 2006. Cancer despite immunosurveillance: immunoselection and immunosubversion. *Nature Reviews Immunology,* 6 (10), 715.

Web references

WHO, 2018. Cancer fact sheet [online]. Available at https://www.who.int/news-room/fact-sheets/detail/cancer [Accessed on 20th October 2019].

Cancer Research UK, 2016. Cancer statistics for the UK, Cancer Incidence [online]. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk#heading-Zero [Accessed on 20th October 2019].

Cancer Research UK, 2016. Cancer statistics for the UK, Cancer Mortality [online]. Available at https://www.cancerresearchuk.org/health-professional/cancer-statistics-for-the-uk#heading-One [Accessed on 20th October 2019].

Uniprot database. UniProtKB - Q04941 (PLP2_HUMAN) [online]. Available at https://www.uniprot.org/uniprot/Q04941 [Accessed on 25th November 2019].

Uniprot database. UniProtKB - Q9R1Q7 (PLP2_MOUSE) [online]. Available at https://www.uniprot.org/uniprot/Q9R1Q7 [Accessed on 25th November 2019].

Uniprot database. UniProtKB - P36552 (HEM6_MOUSE) [online]. Available at https://www.uniprot.org/uniprot/P36552#function [Accessed on 25th November 2019].

Metacore. Clarivate analysis [online]. Available at https://portal.genego.com/ [Accessed 15 Oct 2019.

BloodSpot. [online]. Available at http://servers.binf.ku.dk/bloodspot/ [Accessed on 15 Oct 2019].

OncoLnc. [online]. Available at http://www.oncolnc.org/ [Accessed on 15 Oct 2019].

Morpheus. [online]. Available at https://software.broadinstitute.org/morpheus/ [Accessed on 15 Sep 2019].

Appendix

Table 1A summary of the amount and concentration of reagents that were used to prepare samples for proteomic MS analysis.

Cell Type (Splenocyte)	Cell No. (x10⁶)	EB (µL)	Protein Conc. (µg/µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
Experiment 1	10	100	2.36	10.6	89.4	1	2.7	103.7	0.97	1.25
	5	50	1.72	14.5	85.5	1	2.7	103.7	1.33	1.25
	2.5	50	1.25	20	80	1	2.7	103.7	1.83	1.25
	1	50	0.39	63.8	There two samples were excluded from preparation because there were no enough proteins yielded from cell lysis					
	0.5	50	0.15	166.6						
Experiment 2	10	100	2.63	9.5	90.5	1	2.7	103.7	0.87	1.25
	5	50	1.62	15.5	84.5	1	2.7	103.7	1.42	1.25
	2.5	50	0.83	30.2	69.8	1	2.7	103.7	2.77	1.25
	1	50	0.34	73.5						
	0.5	50	0.18	136.4						
Experiment 3	10	100	1.41	17.8	82.2	1	2.7	103.7	1.63	1.25
	5	50	1.48	16.8	83.2	1	2.7	103.7	1.54	1.25
	2.5	50	0.67	37.4	62.6	1	2.7	103.7	3.43	1.25
	1	50	0.29	85.6						
	0.5	50	0.27	91.5						
Experiment 4	10	100	1.68	14.9	85.1	1	2.7	103.7	1.36	1.25
	5	50	1.40	17.9	82.1	1	2.7	103.7	1.64	1.25
	2.5	50	0.62	40.5	59.5	1	2.7	103.7	3.71	1.25
	1	50	0.32	77.9						
	0.5	50	0	0						

Table 2A. summary of summarises the concentration of proteins yielded in each sample, and the amount and concentration of reagents that were used to prepare samples for MS analysis.

Cell Type (Splenocytes)	Cell No. (x10⁶)	EB (µL)	Protein Conc. (µg/µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
Experiment 1	10	100	1.15	21.7	78.3	1	2.7	103.7	1.98	1.25
	5	50	1	25	75	1	2.7	103.7	2.37	1.25
	2.5	50	0.99	25.3	74.7	1	2.7	103.7	2.40	1.25
	1	50	0.76	32.9	67.1	1	2.7	103.7	3.13	1.25
	0.5	50	0.66	29.9	70.1	1	2.7	103.7	2.84	1.25
Experiment 2	10	100	1.8	13.9	86.1	1	2.7	103.7	1.32	1.25
	5	50	1.1	22.7	77.3	1	2.7	103.7	2.15	1.25
	2.5	50	0.82	30.5	69.5	1	2.7	103.7	2.89	1.25
	1	50	0.8	31.3	68.7	1	2.7	103.7	2.97	1.25
	0.5	50	0.64	35.7	64.3	1	2.7	103.7	3.39	1.25

Table 3A. Summary of the concentration of proteins resulted in each sample and the amount and concentration of reagents used for samples preparation.

Cell Type (Splenocytes)	Cell No. ($x10^6$)	EB (µL)	Protein Conc. (µg/µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
Experiment 1	10	60	2.5	10	90	1	2.7	103.7	0.95	1.25
	5	40	1.4	17.9	82.1	1	2.7	103.7	1.70	1.25
	2.5	30	1.35	18.5	81.5	1	2.7	103.7	1.75	1.25
	1	24	1.08	23.1	76.9	1	2.7	103.7	2.19	1.25
	0.5	18	0.85	29.4	70.6	1	2.7	103.7	2.79	1.25
Experiment 2	10	60	2.3	10.9	89.1	1	2.7	103.7	1.03	1.25
	5	40	1.6	15.6	84.4	1	2.7	103.7	1.48	1.25
	2.5	30	1.2	20.8	79.2	1	2.7	103.7	1.97	1.25
	1	24	0.95	26.3	73.7	1	2.7	103.7	2.49	1.25
	0.5	18	0.76	32.9	67.1	1	2.7	103.7	3.12	1.25

Table 4A. Summary of the concentration of proteins resulted in each sample and the amount and concentration of reagents used for samples preparation.

Cell Type (Splenocytes)	Cell No. ($x10^6$)	EB (µL)	Protein Conc. (µg/µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)	P. Max (µL)
Experiment 1	10	60	1.8	13.9	86.1	1	2.7	103.7	1.32	1.25	1.25
	5	40	1.4	17.9	82.1	1	2.7	103.7	1.70	1.25	1.25
	2.5	30	1.09	22.9	77.1	1	2.7	103.7	2.18	1.25	1.25
	1	24	1.0	25	75	1	2.7	103.7	2.37	1.25	1.25
	0.5	18	0.77	32.5	67.5	1	2.7	103.7	3.08	1.25	1.25
Experiment 2	10	60	2.3	10.9	89.1	1	2.7	103.7	1.04	1.25	1.25
	5	40	1.6	15.6	84.4	1	2.7	103.7	1.48	1.25	1.25
	2.5	30	1.23	20.3	79.7	1	2.7	103.7	1.93	1.25	1.25
	1	24	1.02	24.5	75.5	1	2.7	103.7	2.33	1.25	1.25
	0.5	18	0.85	29.4	70.6	1	2.7	103.7	2.79	1.25	1.25

Table 5A. Summary of the concentration of proteins yielded in each sample in all experiments, and the amount and concentration of reagents used for preparation samples.

Experiments	Cell No. (x10⁶)	EB (µL)	Protein Conc. (µg/µL)	25 µg of proteins in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin / *Trypsin Lys-C (µL)	P. Max (µL)
Exp 1	2.5	30	0.9	27.8	172.2	1	2.7	203.7	1.3	1.25	
	2.5	30	1.2	20.8	179.2	1	2.7	203.7	0.97	1.25	1.25
	2.5	30	1.14	21.9	178.1	1	2.7	203.7	1.02	*1.25	
	2.5	30	0.99	25.3	174.7	1	2.7	203.7	1.18	*1.25	1.25
Exp 2	2.5	30	0.98	25.5	174.5	1	2.7	203.7	1.19	1.25	
	2.5	30	0.96	26	174	1	2.7	203.7	1.21	1.25	1.25
	2.5	30	1.21	20.7	179.3	1	2.7	203.7	0.96	*1.25	
	2.5	30	1.10	22.7	177.3	1	2.7	203.7	1.06	*1.25	1.25
Exp 3	2.5	30	1.15	21.7	178.3	1	2.7	203.7	1.01	1.25	
	2.5	30	0.99	25.3	174.7	1	2.7	203.7	1.18	1.25	1.25
	2.5	30	1.25	20	180	1	2.7	203.7	0.93	*1.25	
	2.5	30	1.18	21.2	178.8	1	2.7	203.7	0.99	*1.25	1.25
Exp 4	2.5	30	0.96	26	174	1	2.7	203.7	1.21	1.25	
	2.5	30	0.92	27.2	172.8	1	2.7	203.7	1.27	1.25	1.25
	2.5	30	1.3	19.2	180.8	1	2.7	203.7	0.90	*1.25	
	2.5	30	1.05	23.8	176.2	1	2.7	203.7	1.11	*1.25	1.25

Table 6A. Summary of the concentration of proteins of each sample, and the amounts and concentrations of each reagent used for processing samples.

Cell No. (x10⁶)	[1] EB 9.5M [2] EB 8M [3] EB 7M (µL)	[1] Urea 9.5M [2] Urea 8M [3] Urea 7M (µL)	Protein Conc. (µg/ µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
2.5	[1] 30		1.05	23.8	176.2	1	2.7	203.7	1.11	1.25
2.5		[1] 30	1.10	22.7	177.3	1	2.7	203.7	1.05	1.25
2.5	[2] 30		1.2	20.8	179.2	1	2.7	203.7	0.82	1.25
2.5		[2] 30	1.08	23.1	176.9	1	2.7	203.7	0.91	1.25
2.5	[3] 30		0.9	27.8	172.2	1	2.7	203.7	0.95	1.25
2.5		[3] 30	0.96	26.0	174	1	2.7	203.7	0.89	1.25
2.5	[1] 30		1.2	20.8	179.2	1	2.7	203.7	0.97	1.25
2.5		[1] 30	1.1	22.7	177.3	1	2.7	203.7	1.05	1.25
2.5	[2] 30		1.02	24.5	175.5	1	2.7	203.7	0.96	1.25
2.5		[2] 30	1.06	23.6	176.4	1	2.7	203.7	0.92	1.25
2.5	[3] 30		0.87	28.7	171.3	1	2.7	203.7	0.98	1.25
2.5		[3] 30	0.84	29.8	170.2	1	2.7	203.7	1.02	1.25

Table 7A. Protein concentrations of each sample and the concentrations and amounts of reagents used for preparation samples.

Cell No. (x10^6)	[1] EB 9.5M [2] EB 8M [3] EB 7M (µL)	[1] Urea 9.5M [2] Urea 8M [3] Urea 7M (µL)	Protein Conc. (µg/µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
2.5	[1] 30		1.24	20.2	179.8	1	2.7	203.7	0.94	1.25
2.5		[1] 30	1.1	22.7	177.3	1	2.7	203.7	1.06	1.25
2.5	[2] 30		1.5	16.7	183.3	1	2.7	203.7	0.65	1.25
2.5		[2] 30	1.04	24	176	1	2.7	203.7	0.94	1.25
2.5	[3] 30		0.85	29.4	170.6	1	2.7	203.7	1.01	1.25
2.5		[3] 30	0.91	27.5	172.5	1	2.7	203.7	0.94	1.25
2.5	[1] 30		1.07	23.4	176.6	1	2.7	203.7	1.09	1.25
2.5		[1] 30	1.17	21.4	178.6	1	2.7	203.7	1.00	1.25
2.5	[2] 30		1.6	15.6	184.4	1	2.7	203.7	0.61	1.25
2.5		[2] 30	1.2	20.8	179.2	1	2.7	203.7	0.82	1.25
2.5	[3] 30		0.82	30.5	169.5	1	2.7	203.7	1.05	1.25
2.5		[3] 30	0.89	28.1	171.9	1	2.7	203.7	0.97	1.25

Table 8A. Summary of the Protein concentrations of each sample and the concentrations and amounts of reagents used for preparation samples.

Cell No. (x10^6)	EB 8M (µL)	Protein Conc. (µg/ µL)	25 µg of protein in volume (µL)	TEAB (µL)	DTT (µL)	IAA (µL)	Total volume at digestion (µL)	Urea Conc. At digestion (M)	Trypsin (µL)
2.5	30	2.02	12.4	187.6	1	2.7	203.7	0.49	1.25
2.5	30	1.86	13.4	186.6	1	2.7	203.7	0.53	1.25
2.5	30	1.76	14.2	185.8	1	2.7	203.7	0.56	1.25
2.5	30	1.98	12.6	187.4	1	2.7	203.7	0.49	1.25

Table 9A. List of upregulated and downregulated proteins in iTregs vs Stim CD4, sorted based on the fold-change and Confidence score ≥50%.

Protein	Fold-change	Confidence Score
EWS	4.42	64%
A0A0R4J187	4.01	51%
Q3UMA3	4.00	57%
MTMRC	3.75	52%
A6PW84	3.68	57%
TTC28	3.30	59%
MT2	3.24	51%
D5MCW4	3.22	63%
ARGL1	2.96	59%
MESD	2.84	58%
Q9EQ08	2.69	52%
RBBP5	2.68	62%
NAB2	2.60	52%
DOPD	2.58	64%
CASP7	2.46	69%
ARF1	2.34	50%
HBB2	2.18	82%
Q9CPX4	2.18	59%
WASH1	2.07	54%
CS012	2.02	74%
STAT2	1.98	54%
HEM6	1.92	80%
CAPG	1.92	87%
HBB1	1.88	53%
RING2	1.86	65%
Q9DCC5	1.83	68%
Q99LB4	1.81	86%
RS23	1.78	51%
CRYL1	1.77	52%
CYTB	1.67	52%
PTMS	1.56	63%
PPP5	1.43	65%
ALDR	1.36	90%
MCRI1	1.35	61%
CYH3	1.35	72%
SUOX	1.13	58%
TT39B	1.12	63%
ASSY	1.10	77%
PLRG1	1.08	54%
A0A0A6YY34	1.07	84%
DJC16	1.02	70%
ORN	0.90	57%
TBA1A	0.90	71%
PNKP	0.87	56%
E9Q9M1	0.84	60%
SRSF1	0.81	72%
P85A	0.77	51%
E9PYG6	0.74	51%
RUXG	0.74	65%
RL30	0.72	54%
RNH2C	0.69	60%
UBCP1	0.68	71%
TBCA	0.67	52%
CNDP2	0.65	79%
E9QA45	0.65	79%
Q6ZWZ4	0.64	53%
GSH0	0.64	53%
LPXN	0.60	55%
LKHA4	0.60	53%
LGUL	0.59	59%
KAPCB	0.58	62%
GDIB	0.52	66%
VIME	0.50	69%
UBA1	0.48	75%
COPZ1	0.47	61%
PDCD4	0.44	59%
TPM3	-0.29	59%
MARK4	-0.30	52%
MCM5	-0.31	62%
RSSA	-0.35	58%
A2A547	-0.38	80%
RS28	-0.39	66%
RL19	-0.42	71%
EYA4	-0.42	70%
EF2	-0.42	54%

ARP5L	-0.43	59%
CLIC1	-0.43	51%
PSDE	-0.44	58%
A0A0A6YX26	-0.44	63%
PSB10	-0.44	57%
D3Z6X7	-0.44	68%
SH3L3	-0.45	54%
SET	-0.46	63%
LRRF1	-0.46	58%
PSA4	-0.47	52%
Sep-05	-0.47	57%
Q6ZWZ7	-0.47	63%
IF4G1	-0.47	55%
PDC10	-0.47	62%
SRSF4	-0.47	57%
EF1B	-0.48	54%
UH1BL	-0.49	63%
Sep-06	-0.50	65%
Q9Z1A1	-0.51	57%
SETD2	-0.51	80%
CDK2	-0.52	74%
D3Z5F7	-0.52	55%
ALKB5	-0.52	54%
PA2G4	-0.52	62%
TCPD	-0.53	62%
IMB1	-0.53	63%
RL11	-0.53	55%
E9PYJ6	-0.54	56%
LASP1	-0.54	79%
CAC1H	-0.54	61%
PKHH2	-0.55	62%
RL13	-0.55	78%
RL31	-0.55	69%
PLSL	-0.55	65%
MYG1	-0.55	57%
A0A1B0GSG5	-0.55	58%
RINI	-0.55	57%
AMRA1	-0.56	66%
GBLP	-0.56	50%
ACLY	-0.56	65%
PCBP2	-0.57	53%
SMYD5	-0.57	50%
M3K12	-0.57	75%
EI2BE	-0.57	54%
SBNO2	-0.57	72%
RLA0	-0.57	80%
ASC	-0.57	57%
DPY30	-0.58	54%
TXD17	-0.58	68%
RL17	-0.59	79%
COR1C	-0.59	55%
COR1A	-0.59	67%
ARBK1	-0.60	60%
E9PWW6	-0.60	71%
BUB3	-0.60	64%
MYO9B	-0.60	52%
A0A1L1SQA8	-0.61	78%
PSB2	-0.61	60%
PSA3	-0.61	58%
CAPZB	-0.61	61%
ZCCHV	-0.61	55%
PFKAL	-0.62	51%
COF1	-0.62	56%
INPP	-0.63	73%
RS20	-0.63	82%
PSA1	-0.63	70%
NH2L1	-0.63	77%
FUS	-0.64	56%
RAC2	-0.64	83%
Q8C605	-0.64	58%
O08797	-0.64	64%
NUCL	-0.65	73%
SRA1	-0.65	52%
BCL10	-0.66	68%
SYFB	-0.66	55%
ALDOC	-0.67	66%
GABR1	-0.67	54%
IF2G	-0.67	55%
SYHC	-0.67	74%
1433G	-0.67	66%

PAIRB	-0.68	77%
THADA	-0.68	55%
RL35	-0.68	66%
S10AB	-0.70	60%
RLA1	-0.70	72%
COX17	-0.70	56%
FKBP5	-0.70	55%
DGKA	-0.70	79%
PGP	-0.70	53%
RADI	-0.70	60%
ARC1B	-0.70	80%
TLK1	-0.71	61%
ENOA	-0.71	73%
ZER1	-0.71	68%
ACTC	-0.72	59%
RL34	-0.72	80%
RL27A	-0.72	70%
ASPM	-0.72	72%
MIF	-0.73	70%
TEX10	-0.73	58%
RL14	-0.73	61%
RL26	-0.73	61%
HNRPD	-0.73	53%
HDDC2	-0.73	55%
SYYC	-0.73	69%
RIPK3	-0.73	69%
SYG	-0.73	70%
ISG15	-0.74	68%
Q3U741	-0.74	53%
GRB2	-0.75	68%
PYM1	-0.75	70%
IF6	-0.76	59%
HINT1	-0.77	73%
RUVB1	-0.77	76%
TPM4	-0.77	74%
PGAM1	-0.78	69%
THIO	-0.79	68%
RLA2	-0.80	77%
AIMP1	-0.80	56%
FKB1A	-0.80	84%
BIN2	-0.80	77%
6PGD	-0.80	74%
RL29	-0.80	97%
ELMO1	-0.81	62%
F7DEU6	-0.81	80%
UB2V1	-0.81	51%
SR140	-0.82	55%
GNPI1	-0.82	61%
AEDO	-0.82	61%
ADK	-0.83	52%
RS25	-0.83	70%
ECHD1	-0.83	68%
SAP	-0.83	55%
F8WJB9	-0.84	50%
NOP56	-0.85	51%
RUVB2	-0.85	68%
GIMA4	-0.85	89%
TBCD1	-0.85	53%
SYAC	-0.86	63%
SH3K1	-0.87	52%
PRDBP	-0.87	56%
STK4	-0.88	54%
G5E8E4	-0.88	75%
PP1G	-0.88	55%
UFC1	-0.88	56%
HDAC7	-0.88	57%
H2AV	-0.89	52%
ROAA	-0.89	72%
A6ZI44	-0.90	63%
GALK1	-0.90	78%
GMFB	-0.90	52%
SERC	-0.90	66%
SYSC	-0.91	68%
CYC	-0.91	63%
KIF15	-0.91	76%
KCC4	-0.92	64%
E9Q5C9	-0.92	64%
TAGL2	-0.92	68%
A0A087WRY3	-0.93	54%
Q6PE70	-0.93	67%

RAD50	-0.94	59%
THIC	-0.94	55%
RL10A	-0.94	71%
F8WJI3	-0.95	74%
TITIN	-0.95	52%
PGM1	-0.95	77%
GBP1	-0.96	64%
KCRB	-0.96	76%
A2AMV1	-0.98	57%
MDHM	-0.98	76%
KPYM	-0.99	71%
AB17C	-1.02	91%
SAMH1	-1.02	79%
TSN	-1.02	50%
SERB	-1.02	58%
PFKAP	-1.03	82%
STS	-1.03	66%
TPIS	-1.03	75%
A0A0R4J0J3	-1.05	67%
SAP3	-1.05	78%
DDRGK	-1.05	52%
DHPR	-1.06	78%
FPPS	-1.07	74%
CLCB	-1.08	58%
CH10	-1.08	53%
GBB2	-1.09	58%
A0A087WR57	-1.09	73%
HPCL1	-1.10	80%
ALDOA	-1.10	73%
LRCH1	-1.11	57%
PGK1	-1.12	75%
KAD2	-1.13	64%
TCOF	-1.13	63%
AATM	-1.14	76%
SON	-1.16	71%
LSP1	-1.18	84%
Q8BFQ1	-1.19	82%
CCD66	-1.21	68%
PML	-1.21	83%
THMS1	-1.22	82%
SEC20	-1.23	66%
HA1D	-1.23	55%
GAPR1	-1.24	76%
F107B	-1.26	68%
F8WJE0	-1.27	88%
NDRG1	-1.28	93%
RU1C	-1.30	76%
RS2	-1.32	54%
A2A6J4	-1.32	82%
RL7A	-1.32	77%
RED	-1.32	56%
PTPRC	-1.33	75%
H9H9T1	-1.34	57%
A0A0B4J1E7	-1.34	70%
MLKL	-1.34	57%
TCP4	-1.36	60%
IKZF1	-1.37	54%
F6ZHD8	-1.38	65%
E9QP46	-1.38	55%
KI13A	-1.39	75%
CA198	-1.41	53%
PRAF3	-1.41	72%
KBTBB	-1.42	76%
PLXA2	-1.42	71%
GNAS1	-1.45	54%
PDLI2	-1.46	63%
E9PUA7	-1.46	57%
BRD2	-1.50	58%
TT21B	-1.54	70%
ESYT1	-1.54	60%
CO6A2	-1.54	63%
HXK2	-1.67	56%
E9Q8L9	-1.76	67%
KTHY	-1.77	58%
G3X8R0	-1.79	55%
Q5PPR2	-1.80	55%
HIKES	-1.80	57%
GRIN3	-1.82	64%
PPIL2	-1.84	78%
Q8CBB6	-1.84	57%

DYST	-1.86	58%
RBP1	-1.96	53%
SBP1	-2.00	75%
APOB	-2.05	69%
RHG22	-2.07	93%
SMU1	-2.17	65%
KIN17	-2.17	57%
SAHH3	-2.18	63%
UGDH	-2.20	56%
GOGA5	-2.24	53%
AQR	-2.24	50%
COX7C	-2.31	61%
UBE3C	-2.32	51%
ANXA2	-2.32	83%
RP1	-2.40	51%
CWC27	-2.47	54%
CD97	-2.52	52%
H14	-2.53	68%
H15	-2.54	56%
NOL10	-2.57	50%
H2A2A	-2.59	58%
H4	-2.61	80%
CO4B	-2.64	52%
ATG3	-2.64	52%
RHG25	-2.69	54%
GTR3	-2.72	82%
AP4A	-2.77	55%
FOLH1	-3.01	88%
H2A1K	-3.03	60%
ASB6	-3.08	63%
B1AW21	-3.09	55%
VP26A	-3.14	59%
BOLA2	-3.19	83%
CNO10	-3.25	54%
FIP1	-3.28	53%
H2AX	-3.30	60%
SF3B5	-3.39	63%
H13	-3.45	70%
B1ASU9	-3.48	63%
CHD9	-3.61	59%
TGM2	-3.77	84%
BAG3	-3.83	62%
NHEJ1	-4.16	60%
HINT2	-6.49	55%
SDCG8	-6.64	76%
AMRA1	5.16	66%
A0A0R4J0F6	4.69	53%
AT1B1	3.29	74%
PTH2	3.22	60%
WASH1	3.16	59%
NRK1	3.16	65%
ENDD1	2.92	65%
TBB1	2.82	52%
DAAM2	2.78	56%
M3K11	2.45	77%
DPP9	2.37	51%
CARF	2.31	54%
E9Q6R7	2.19	66%
CAF17	2.09	50%
SIVA	2.08	51%
XYLT1	2.03	60%
MAVS	1.92	84%
ITIH4	1.84	72%
Q0VBL3	1.78	66%
PLP2	1.73	56%
A2A6J4	1.67	71%
E9QP00	1.61	52%
A0A0N4SVQ1	1.58	89%
Q91VB8	1.54	52%
NDUA4	1.46	91%
OAT	1.35	74%
MYH9	1.19	74%
MYL6	1.13	58%
ML12B	1.07	83%
COR1A	1.03	80%
CS012	0.99	67%
ACON	0.94	76%
D3Z5F7	0.86	59%
A2AP32	0.79	60%
CX6B1	0.79	72%

VIME	0.78	75%
CD47	0.73	56%
P5CS	0.71	64%
SP110	0.70	50%
ADT2	0.70	64%
NDUS7	0.67	68%
PTN1	0.64	72%
H32	0.63	52%
CASP7	0.61	51%
LAT	0.57	65%
SUMO2	0.55	88%
NDUB3	0.50	67%
A2BIE1	0.46	54%
E41L2	0.44	57%
NDUA5	0.44	55%
RAB8B	0.43	56%
COX6C	0.42	55%
NDUS3	0.42	57%
NDUV2	0.32	55%
CBR4	-0.27	52%
RL31	-0.34	51%
TM109	-0.40	57%
RAC2	-0.40	57%
FKBP2	-0.40	51%
RL27A	-0.41	61%
Q6PE70	-0.43	59%
RRBP1	-0.45	51%
AT1A1	-0.47	57%
H2AV	-0.49	51%
PDIA3	-0.50	71%
GNAS1	-0.51	54%
ROAA	-0.51	54%
RL10A	-0.51	69%
A0A1B0GX27	-0.53	52%
RS28	-0.53	56%
K2C79	-0.54	74%
H14	-0.54	51%
A0A1L1SQA8	-0.54	51%
GRP78	-0.56	69%
CBX5	-0.56	52%
CD3G	-0.57	56%
MOT1	-0.57	67%
RAP1B	-0.58	55%
RRAS2	-0.58	52%
RL40	-0.59	50%
LIMD2	-0.59	57%
HA1D	-0.60	59%
DNJC9	-0.60	60%
SET1B	-0.61	55%
CH60	-0.61	60%
GLU2B	-0.62	74%
Sep-05	-0.63	50%
ITB2	-0.65	53%
CD2	-0.66	53%
AT1B3	-0.68	62%
EF1B	-0.68	55%
URP2	-0.70	55%
SPTN1	-0.71	81%
FLOT1	-0.73	55%
ENPL	-0.74	78%
CATE	-0.75	86%
HMGN2	-0.75	58%
ITAL	-0.75	71%
BASI	-0.76	79%
SPTB2	-0.76	74%
RLA2	-0.77	70%
MANF	-0.83	52%
Sep-07	-0.83	56%
TXD12	-0.83	64%
ARHG2	-0.86	66%
RFC4	-0.87	62%
RASA3	-0.89	57%
KAD4	-0.89	58%
H2AY	-0.90	67%
CD6	-0.93	74%
FBLN1	-0.96	53%
IKZF1	-0.99	50%
PTCA	-1.00	56%
DHC24	-1.00	56%
HXK2	-1.00	64%

D3YX85	-1.01	65%
Q5SUH6	-1.01	54%
CD48	-1.01	72%
PTPRC	-1.01	81%
CD3D	-1.02	83%
K3W4Q8	-1.02	72%
CD5	-1.03	93%
TCB1	-1.05	86%
SYYC	-1.05	57%
LY9	-1.06	81%
ERO1A	-1.08	74%
PDIA1	-1.09	69%
GAPR1	-1.12	87%
NUCB1	-1.13	53%
KI13A	-1.17	68%
GTR1	-1.23	77%
GIMA5	-1.24	61%
MIF	-1.24	57%
A0A0A0MQF6	-1.29	76%
G3P	-1.33	67%
ALDOA	-1.34	55%
G3UZK1	-1.38	70%
E9Q0A7	-1.39	51%
Q9DC42	-1.44	53%
WASP	-1.45	64%
MOT4	-1.49	84%
PAXB1	-1.51	53%
HMGA1	-1.52	60%
KPYM	-1.53	68%
WAPL	-1.61	51%
AB17C	-1.63	90%
KRT85	-1.64	72%
CD97	-1.66	81%
SH3K1	-1.67	63%
DDX54	-1.67	57%
PROS	-1.83	50%
THMS1	-1.98	72%
GRAB	-2.00	75%
SATB2	-2.09	86%
SATB1	-2.16	82%
INCE	-2.19	61%
RBX1	-2.19	52%
MIO	-2.48	61%
UH1BL	-2.51	54%
Q3U125	-2.61	70%
ITB3	-2.75	66%
Z4YJR1	-2.88	64%
NUSAP	-2.92	64%
CR1L	-3.00	62%
E9Q0G1	-3.05	60%
ANXA2	-3.07	75%
GTR3	-3.12	88%
SDCG8	-3.20	64%
MET13	-3.29	55%
Q8CAH8	-3.36	59%
TPX2	-3.38	74%
Q792Z1	-3.62	62%
G3X920	-3.66	73%
A0A0R4J0J3	-4.10	58%
PSMD8	-4.36	66%
DYHC2	-5.34	62%

Table 10A. List of upregulated and downregulated proteins in iTregs vs nTregs, sorted based on the fold-change and Confidence score ≥50%.

Protein	Fold-change	Confidence score
HAUS1	5.976176337	71%
E9Q1M6	3.947804713	75%
Q3UMA3	3.915017274	57%
NOP9	3.783097989	59%
VPS4A	3.500980564	53%
TT21B	3.39013993	64%
A2A4U6	3.34175052	57%
PSMD8	3.241019382	83%
FPPS	3.232690314	78%
E9PWW6	3.217973811	60%
BAIP2	3.165469907	69%
TPIS	3.153365419	87%
HEM6	3.110451987	59%
LRC47	3.055577367	65%
MTMR2	3.052956905	61%
D3Z5N2	3.022273833	77%
PRI2	2.99301149	65%
TRAF6	2.960251077	71%
AB17C	2.936095762	91%
Q9CQ43	2.893899119	64%
PIGT	2.867384406	58%
3BP1	2.862409805	57%
TYSY	2.841712078	76%
STMN1	2.838682374	83%
MA2B1	2.789088052	50%
SETD3	2.77106176	63%
NUBP2	2.768789464	70%
ALDOC	2.764492699	85%
AACS	2.756559219	68%
DHX57	2.741935108	52%
SPA3F	2.728216736	77%
BGAL	2.688703985	52%
NXN	2.687786201	51%
PSME3	2.662295271	53%
ESPL1	2.656615938	56%
PGK1	2.643863611	83%
A6ZI44	2.639074038	83%
RNH2C	2.636608515	53%
MA2B2	2.592720359	59%
VIME	2.576606747	82%
FAS	2.575171276	77%
GALK1	2.560187458	77%
STAT2	2.542360335	77%
Q5SRW8	2.524306234	51%
E9Q496	2.507824385	56%
Q8C605	2.473801304	73%
ALDOA	2.464310915	85%
PTN9	2.461547455	52%
A2AGL5	2.443456543	53%
4ET	2.43542392	54%
PGAM1	2.429740756	79%
Q9DCC5	2.408305778	54%
DYR	2.400061967	74%
PCH2	2.383706385	71%
ENOPH	2.359927238	57%
CCD25	2.356476431	53%
PPP5	2.346188422	55%
XYLT1	2.325469223	71%
PFKAM	2.318423243	62%
E9PW66	2.314192542	52%
CCD43	2.311145198	51%
ITPA	2.310923112	55%
KIF15	2.30297472	53%
DHX29	2.295129906	57%
RL7L	2.264498293	61%
SPF30	2.242098664	56%
TBA1B	2.241235524	62%
MCM4	2.236390131	74%
PFKAL	2.229996817	63%
NAMPT	2.227069358	74%
BROX	2.225183009	57%
ENOA	2.22360986	90%
LDHA	2.218481401	77%
KPYM	2.213124427	88%
NDRG1	2.209128209	71%

TBB5	2.204592625	89%
TF2AA	2.182018603	50%
CNDP2	2.163824104	83%
MARK4	2.150629986	70%
MCM7	2.132796062	78%
D3Z4B2	2.132703544	55%
PCNA	2.130458947	64%
RL5	2.12375492	57%
Q564E2	2.070048597	81%
NP1L1	2.04881197	51%
D3Z7B5	2.045407879	68%
RL8	2.032850948	53%
RIR1	2.031533295	74%
XPO1	2.025047568	51%
OGFR	2.018750591	79%
E9PYG6	2.006828693	58%
FABP5	2.004846918	83%
VAMP7	2.003756692	58%
RL10	1.980682052	54%
CRADD	1.971719346	76%
IMPA2	1.944394486	73%
PFKAP	1.911252199	73%
SEC20	1.894501153	51%
THADA	1.891698906	77%
PA2G4	1.88381308	81%
SERC	1.88318566	56%
MCM2	1.860913467	77%
SYK	1.851941703	57%
NFAC2	1.834407739	71%
CC186	1.831752371	58%
MCM5	1.825241863	82%
HPRT	1.794280606	74%
OSTF1	1.783409931	61%
IPO5	1.768907712	69%
EF1B	1.757933732	77%
TNC18	1.755830362	69%
MCM3	1.752651383	80%
PGM1	1.736356876	72%
Q9QXJ2	1.731975426	84%
RL27	1.728826923	55%
ORN	1.72550424	57%
P85A	1.710538315	57%
ZN827	1.707143544	73%
A2BHP6	1.706029751	54%
PYR1	1.695693529	72%
IMDH2	1.690839328	80%
B1B1A8	1.688550334	63%
MCM6	1.688462923	75%
RL26	1.687992864	72%
RIPK3	1.681862973	50%
SYWC	1.679250802	70%
Q5XJF6	1.66836229	70%
COX17	1.658044526	58%
MIF	1.645344938	89%
UBQL1	1.636652037	69%
UBE2K	1.631591734	59%
G3P	1.5724389	83%
IPYR	1.561162765	76%
TLK1	1.558614959	50%
TCPZ	1.557336829	70%
EF1G	1.531146268	74%
A0A0R4J1E2	1.519931528	63%
A0A0A0MQF6	1.519015517	81%
NUDC	1.517516049	57%
E9PYH2	1.500569451	69%
LEG1	1.495421145	53%
KRT85	1.487862328	80%
EF1A1	1.487034822	73%
RL39	1.474656244	74%
A0A0R4J138	1.466401373	53%
RL29	1.464346999	88%
ZER1	1.463157379	71%
GRHPR	1.45800026	84%
TBB4B	1.456429503	86%
EEA1	1.454507141	50%
IF4A1	1.451667011	59%
CUL4A	1.451307141	50%
G6PI	1.447645256	79%
SAP3	1.444361431	75%
SPEE	1.398736585	62%

TBA1A	1.388304409	68%
IF5A1	1.374934077	80%
SYVC	1.367969209	69%
SERA	1.367452431	81%
A0A0A6YX26	1.358182054	65%
RS3A	1.356923041	71%
THIO	1.351840481	82%
RL23A	1.341293138	56%
ANKL2	1.34027211	60%
EF2	1.327727052	81%
NDKA	1.32356167	70%
EF1D	1.314252729	86%
RLA2	1.310930885	82%
RS15A	1.30910125	74%
RS25	1.302317658	70%
GALE	1.300421309	58%
RFC4	1.285967475	57%
RL10A	1.280392803	71%
RS27	1.274944358	77%
RL19	1.262343666	82%
PTN6	1.259582568	74%
BTF3	1.240861495	52%
PAIRB	1.239814502	74%
A2A5V4	1.221328507	53%
F8VPV0	1.218844752	62%
RS14	1.214249505	71%
IMA4	1.211509959	61%
Q6ZWZ7	1.206416727	75%
A0A1B0GTA4	1.184590631	74%
HNRH2	1.181683292	51%
RL35	1.179136411	68%
PRPS1	1.170845851	81%
RL7	1.167450664	64%
E9PZF0	1.164873514	71%
G3UYV7	1.163791019	85%
PDLI4	1.150034377	55%
SYMC	1.147962468	66%
M3K12	1.143498381	89%
SYHC	1.14221261	60%
NDKB	1.142086175	79%
E9PVU0	1.140218058	58%
Q8BFQ1	1.137245214	67%
RS11	1.135658301	51%
Q6ZWZ6	1.135390247	61%
PUR6	1.133971147	52%
RS13	1.133128273	68%
AIMP1	1.126484354	58%
SYYC	1.126305311	66%
NPM	1.120060761	77%
URP2	1.115548349	61%
RL17	1.112931528	69%
RL11	1.111591543	75%
PPIA	1.1094904	81%
PTMA	1.105381192	81%
RS29	1.100403605	52%
Q8R2P8	1.100069998	51%
CALM	1.098891859	57%
Q3UKW2	1.098891859	57%
UMPS	1.098349016	51%
TCPE	1.094121076	76%
HS90B	1.093053431	77%
NTAN1	1.091822827	56%
RS18	1.080926284	79%
GLRX3	1.061449173	51%
RS20	1.059206356	85%
1433F	1.058567154	74%
RL13	1.058281259	71%
RL27A	1.050752801	62%
TCTP	1.050407451	66%
RS7	1.050101572	72%
Q8C3V4	1.043198793	68%
RLA0	1.039473793	81%
UBA1	1.034484634	79%
CDK2	1.028900088	61%
RL12	1.026560044	81%
IIGP1	1.024528229	58%
ERF1	0.999784978	55%
A0A1L1SQA8	0.99586945	74%
RS16	0.991578779	79%
RS2	0.991412376	80%

RL34	0.986133915	50%
EIF3G	0.985728436	50%
DOK2	0.983235507	56%
RS4X	0.983096934	73%
TCPB	0.971956541	68%
RL23	0.962146625	74%
PRDX1	0.961155009	82%
1433E	0.958981736	70%
Q91V55	0.95615805	74%
IF4E	0.951740658	50%
PSD11	0.946182502	66%
RL31	0.938696744	87%
RL4	0.936012172	69%
IMB1	0.932987815	60%
SYSC	0.926338362	58%
RS19	0.920804793	75%
Q6ZWZ4	0.919466223	59%
CAPG	0.918621095	74%
RL6	0.916754952	54%
PNKP	0.914344555	100%
AGAP2	0.914032087	58%
RS26	0.908594553	54%
CATZ	0.899439817	54%
TCPD	0.88990838	82%
STAT1	0.881640735	62%
HS90A	0.86865842	69%
RSSA	0.863926767	82%
Q99LB4	0.862002244	78%
A2A547	0.862001121	80%
ARPC3	0.861398782	57%
EIF3I	0.857038637	55%
TCPG	0.841910771	70%
RS28	0.841035908	76%
DPYL2	0.839158146	75%
PSMG3	0.83889422	68%
GDIB	0.832597611	74%
TCPQ	0.827081546	72%
Q9CPS5	0.803562305	64%
CSK22	0.79620706	52%
PKHH2	0.79501471	52%
CATC	0.777690394	70%
RAN	0.775742365	64%
TKT	0.763445033	72%
TCPH	0.759557535	71%
AATC	0.754347194	62%
RL18	0.75393001	57%
TCPA	0.740563185	73%
FKB1A	0.740391013	79%
RL24	0.736755791	62%
A0A0A0MQA5	0.735928861	58%
TPP2	0.71591284	52%
Q8C2Q7	0.710130508	53%
PSB9	0.704755524	61%
PSA1	0.704544265	60%
5NT3B	0.685513127	55%
SNX27	0.683191104	54%
CCS	0.678453029	52%
DDX3X	0.669260891	50%
SERB	0.665037647	55%
PCBP1	0.646977304	78%
RL14	0.642916608	51%
PSDE	0.632522582	65%
RAB10	0.605001071	51%
NUCL	0.600883627	67%
A0A1C7CYV0	0.577432374	67%
PSB4	0.577338146	62%
AN32E	0.563990502	71%
SAR1B	0.558312564	50%
HSP7C	0.533118294	69%
COF1	0.506969854	60%
IF4G1	0.458176452	54%
RAB3A	0.402610527	51%
ML12B	0.350682789	50%
SUMO2	-0.437906772	74%
BAG5	-0.488128881	57%
ARP5L	-0.515388739	55%
APEH	-0.555301136	54%
KBTBB	-0.582750783	73%
GIMA4	-0.586958532	82%
LASP1	-0.605192482	81%

ISG15	-0.615980192	52%
MYO9A	-0.633907244	55%
LSM7	-0.643097025	68%
RAC2	-0.649705384	67%
COR1A	-0.656773075	74%
SPRE	-0.66364491	61%
ANX11	-0.668745171	51%
A2AQN4	-0.676396627	58%
EYA4	-0.676961195	70%
MIO	-0.683987047	52%
Q3UDM0	-0.69896521	54%
NSF1C	-0.699207077	68%
GDIR2	-0.708360074	57%
ANXA5	-0.716944179	73%
NH2L1	-0.780835972	76%
IL16	-0.790376457	77%
Q924B0	-0.820365322	73%
RAB35	-0.825073078	56%
DCPS	-0.832661562	57%
MGDP1	-0.836458482	69%
NAA35	-0.839041148	67%
NC2B	-0.847865948	56%
PP1G	-0.85088241	61%
SH3L1	-0.870319261	66%
ALDR	-0.913791336	76%
F193A	-0.941189771	59%
B7FAU9	-0.943674952	76%
FLNA	-0.98740393	83%
E9QMV2	-0.990226805	80%
SUMO1	-1.011031426	51%
TRNT1	-1.023985754	61%
COTL1	-1.070273899	66%
THUM1	-1.090353667	68%
ARHL2	-1.090762814	90%
H4	-1.101741029	69%
K1671	-1.127870326	71%
COA7	-1.205892258	61%
ADK	-1.209179347	74%
NCK5L	-1.253254215	70%
NT5D1	-1.260329357	52%
FETUA	-1.269541086	67%
F8WIV2	-1.310106855	64%
CRYL1	-1.318626918	75%
Q9D3L3	-1.335593554	50%
PI42A	-1.350272376	51%
ACTN1	-1.355551085	78%
SH3L3	-1.382962303	84%
CBR1	-1.384967919	73%
Q9Z1R9	-1.407690967	64%
RASF2	-1.432516526	77%
LDHB	-1.450042124	73%
HBB2	-1.452646404	71%
LSP1	-1.459802441	82%
ALBU	-1.472934882	54%
RP1	-1.478193371	60%
SAMH1	-1.509903903	84%
PGM2	-1.530752445	63%
NUF2	-1.562914648	68%
E9QA45	-1.615720156	87%
MDHC	-1.618283509	84%
ALG5	-1.643446007	55%
CCD66	-1.716910781	55%
A0A0N4SVU1	-1.741225317	51%
AL3A1	-1.762505859	69%
A2A6J4	-1.76861937	82%
ZNFX1	-1.799250899	60%
HBB1	-1.947805141	69%
SMU1	-1.952024419	54%
GDE1	-2.122937652	58%
GSHR	-2.167151945	87%
NXP20	-2.188672871	51%
F6QFD1	-2.203205713	82%
2A5A	-2.308269452	69%
CO4B	-2.363107995	59%
PPM1A	-2.412952867	52%
DEGS1	-2.428721472	61%
Q9CPX4	-2.474955864	70%
A7L3B	-2.502508629	55%
SPT6H	-2.566254597	63%
PARP8	-2.784452396	53%

ADA10	-2.786631474	62%
NBN	-2.874122083	50%
MEA1	-2.932147383	51%
PLS3	-2.944291467	57%
EMC2	-3.028877048	51%
CN37	-3.06436379	53%
SCAM3	-3.259336624	50%
RM38	-3.280625557	52%
Q0PD20	-3.29537635	56%
COPE	-3.416868567	63%
VWA8	-3.430355112	51%
REV1	-3.490632586	57%
MKS3	-3.493715443	57%
B0R091	-3.534464722	56%
CD109	-3.53944641	83%
ARMX3	-3.880537772	57%
BOLA2	-4.010263922	61%
PAXB1	-4.201515599	61%
NHEJ1	-4.496686829	59%
A2A4U6	4.982689692	55%
M3K11	4.740146814	57%
NHEJ1	4.515149001	55%
A0A0R4J0F6	3.789269464	56%
NDRG2	3.750100815	64%
CSN7B	3.719436073	61%
PLP2	3.689022256	86%
D3YXP6	3.628123708	50%
XYLT1	3.519154323	72%
UBE2C	3.41624116	60%
PIGT	3.295881594	63%
KI67	3.178351861	77%
K3W4Q8	3.176821273	78%
MOT4	3.133896329	55%
SRA1	3.128443865	53%
MAVS	3.116232429	55%
RM27	3.089410774	51%
PYC	3.089339057	80%
A0A0R4J0T5	3.068419193	59%
NSMA3	3.052409218	55%
E9PZ97	3.032865818	65%
MFS10	3.029930613	60%
RRS1	2.992391584	53%
AT131	2.986363154	58%
TFR1	2.950759968	77%
BASI	2.938760399	77%
RL39	2.85225458	53%
E9PUQ5	2.840505581	51%
AB17C	2.812938361	78%
E9QP62	2.802208636	54%
DNMT1	2.727523308	66%
KAD4	2.724352103	69%
DDAH1	2.710950923	50%
NCBP3	2.708044334	51%
A2AUM0	2.691828829	50%
MIC26	2.682100351	72%
GTR1	2.668670345	61%
CBX5	2.662932355	62%
E9PWG6	2.64845045	72%
MYEF2	2.640070827	51%
MSMO1	2.619052916	57%
SMC2	2.589416922	66%
SYNE3	2.523543649	63%
DHB7	2.510119858	56%
MCM5	2.508739148	66%
Q3UNN4	2.461155731	63%
NRM	2.449179585	59%
TBA1A	2.445783927	51%
BRI3B	2.442106192	56%
SEC63	2.428361781	71%
ERO1A	2.425634386	65%
BCL7C	2.411474052	56%
LRC59	2.385233562	76%
TT21B	2.375570528	55%
NCBP2	2.341750529	53%
IL2RA	2.310505317	76%
LRC8C	2.281477925	61%
TERF2	2.273376489	64%
S2539	2.260478731	57%
K2013	2.237671604	51%
RM17	2.225342186	57%

E9QAS5	2.224122917	58%
A0JNY3	2.201726637	51%
A2A5V4	2.201716211	50%
NSDHL	2.193123487	77%
E9PX68	2.101974524	57%
E9QP00	2.081551819	51%
VIME	2.073260383	82%
PPIF	2.045078012	64%
DDX27	2.025132612	58%
AP2S1	2.017330939	64%
Q9CZN7	2.008627512	81%
MBB1A	1.952073224	75%
F6ZBR8	1.944945527	52%
ATIF1	1.844961515	51%
GNAS1	1.828819526	57%
A0A0R4J170	1.772732171	73%
IMP3	1.755311086	58%
A0A0R4J0V1	1.745982981	62%
TNC18	1.744082202	60%
ITIH4	1.720990902	69%
GTR3	1.719255828	61%
A0A0R4J0D3	1.715367511	70%
Q8C2Q7	1.705611861	56%
LAP2A	1.694158517	50%
ABHGA	1.684692205	50%
PGES2	1.674929385	65%
SYAM	1.674713361	52%
IIGP1	1.589790297	62%
LAT	1.571947656	52%
MOT1	1.543162215	57%
FKBP8	1.530549418	78%
CD47	1.521756522	57%
TM9S4	1.511747048	71%
4F2	1.50855694	54%
GBG2	1.426500688	82%
GAPR1	1.403211843	77%
PYRD	1.393353099	59%
ERG28	1.386334816	58%
RSSA	1.370983006	52%
FKBP2	1.341284766	57%
PHB2	1.316208488	74%
RPN1	1.307906372	78%
SFXN1	1.240247886	79%
ADT2	1.230661374	66%
P5CS	1.222021762	72%
FLOT1	1.215065514	51%
SSRG	1.204912186	70%
Q31093	1.20450616	52%
PDIA6	1.200289106	80%
ERP29	1.189546739	66%
OAT	1.169299517	69%
ENPL	1.163797268	72%
CTL2	1.135686634	60%
ICAM1	1.111363701	61%
PHB	1.105551485	73%
LS14A	1.101804273	54%
SDF2L	1.08763238	70%
Q3UJB0	1.081230351	54%
SUV3	1.072920317	56%
MYH7B	1.04837306	72%
NOP2	1.047483349	53%
RL27A	1.030620908	55%
RBBP4	1.0130215	61%
MTA2	1.009859995	54%
E41L2	0.99525125	51%
RL10A	0.967143477	69%
J7NUP1	0.958626237	56%
A0A1L1SUX8	0.933950976	87%
THY1	0.899042313	61%
LONM	0.896113946	51%
PR40A	0.891082324	58%
Q9D8L3	0.888433694	67%
FUBP2	0.885628316	64%
MFGM	0.879667225	64%
PNKP	0.87918808	85%
A0A0A0MQF6	0.860111589	60%
RNH2B	0.850019537	58%
IRGM1	0.833009637	79%
ESYT1	0.811507111	73%
RL19	0.799144763	75%

RUXF	0.794192962	52%
Q3TCU5	0.786387113	53%
SNRPA	0.780766765	58%
URP2	0.723522629	51%
F208B	0.723052921	58%
VDAC1	0.668898776	59%
G3UYV7	0.656685347	56%
ROAA	0.647674179	52%
VDAC2	0.620943323	54%
GRP75	0.557654145	57%
CATE	-0.608240043	54%
RAC2	-0.649188436	78%
D39U1	-0.699734246	53%
COR1A	-0.710889411	68%
CD5	-0.726835796	67%
EFTU	-0.729583272	66%
PTCA	-0.787351171	65%
CHM1A	-0.821679604	52%
THIM	-0.83078699	52%
UBC9	-0.832849377	64%
ACD10	-0.855542919	51%
SODM	-0.954900893	52%
GIMA1	-0.981637575	60%
ACO13	-1.002283757	55%
LIMD2	-1.106030929	69%
G3X9U9	-1.120982422	57%
IVD	-1.230495099	69%
ACOT9	-1.234025867	64%
ES1	-1.311787078	51%
IL16	-1.335764945	66%
ACTC	-1.344187262	61%
COX7C	-1.35312901	65%
CECR5	-1.405217941	54%
ECI1	-1.470161359	67%
ARHL2	-1.47219518	58%
PRDX5	-1.54121843	65%
FETUA	-1.56103197	64%
CYB	-1.629090902	61%
K1C17	-1.667139482	58%
K1C16	-1.781354311	55%
ACOT2	-1.866607248	87%
EYA4	-1.941342071	63%
NAA35	-1.987904367	57%
TBC8B	-1.99922149	63%
K1C14	-2.011751053	56%
E0CZ27	-2.03178395	72%
K22E	-2.045144758	58%
MA2B2	-2.050546159	55%
E9Q0F0	-2.075959586	55%
K2C79	-2.167485783	81%
ROCK2	-2.174853235	56%
NEST	-2.177505857	56%
K1C10	-2.264287146	80%
PDPR	-2.291383666	56%
F16B1	-2.292925895	52%
K2C1B	-2.440165523	76%
HMGB1	-2.53477992	58%
CCD66	-2.576701269	53%
A2A513	-2.664352983	91%
E9Q0G1	-2.778194168	55%
Q9Z1R9	-2.897566049	85%
A0A0R4J0J3	-3.268143068	50%
HYEP	-3.342961114	76%
Q9CZP3	-3.401367187	65%
E9QA45	-3.507008954	89%
H2B2B	-3.846890226	57%
MIO	-4.010402747	76%
Q792Z1	-4.472107504	77%

Table 11A. List of upregulated and downregulated proteins in iTregs vs naïve CD4, sorted based on the fold-change and Confidence score ≥50%.

Protein	Fold-change	Confidence score
E9PWW6	5.559155236	64%
PIGT	4.832151703	60%
A2CG76	4.683056553	55%
CUL4A	4.218718279	52%
A2A4U6	4.004331271	54%
IL2RA	3.989840648	86%
A0A0R4J0F6	3.927736011	56%
M3K11	3.862699048	77%
F8WIA1	3.842171432	52%
PSF3	3.829572051	57%
PIGT	3.821267208	60%
BASI	3.747434177	74%
G3UW40	3.682472314	54%
KAD4	3.658885869	61%
SUGP1	3.630801647	51%
SLIT2	3.615743029	61%
KIF15	3.611043066	69%
Q3UMA3	3.602690664	52%
GALE	3.573333406	77%
PDCD5	3.546400758	51%
D3Z5N2	3.524383792	79%
IFI4	3.507845444	57%
MSMO1	3.483693611	57%
GNAS1	3.459294559	58%
DNM3A	3.393564932	63%
S2539	3.366041142	64%
K3W4Q8	3.360816069	73%
KI67	3.340960314	79%
IIGP1	3.330678165	84%
SBNO2	3.323354638	70%
CD2	3.296968449	50%
RNH2B	3.284702392	59%
NDRG1	3.280034785	75%
BOREA	3.274421675	53%
DOPD	3.268777389	65%
ATIF1	3.233359566	66%
4EBP1	3.2003607	62%
ERG28	3.158011445	66%
PYC	3.151068223	84%
Q9CQ43	3.142918632	78%
VPS4A	3.136851637	50%
CBX5	3.111914523	69%
SMC2	3.109652269	64%
PSMD8	3.103586189	81%
SUH	3.09436734	55%
Q9EQ08	3.09335696	61%
TPIS	3.076976336	83%
GALM	3.044242329	55%
STMN1	3.036256437	80%
E9Q1M6	3.027505613	62%
GTR1	3.023694626	64%
GPAA1	3.00142856	55%
TFR1	2.983734914	76%
TERF2	2.979699407	78%
ALDOC	2.970230418	85%
MESH1	2.960810877	51%
Q3UKN6	2.95931301	61%
DYR	2.957170294	80%
A0A0R4J0V1	2.955900269	70%
HYI	2.92141074	50%
STA5A	2.907566911	59%
TNC18	2.90099338	65%
NAMPT	2.897522132	78%
AACS	2.89308307	72%
NDUA8	2.883069609	62%
RT06	2.879233288	58%
SHCBP	2.873829522	63%
FPPS	2.871557255	80%
BAIP2	2.867752465	61%
FABP5	2.846191754	82%
E9PYG6	2.83949123	64%
E9QP00	2.81844582	58%
CK074	2.813915429	65%
E9Q718	2.805532954	51%
MCM4	2.795949987	77%

IMPA1	2.779843837	50%
J3QMA2	2.741028464	70%
NXP20	2.723781165	67%
FAS	2.719077107	79%
SPA3F	2.718619946	78%
E9Q496	2.71718032	70%
E9PWG6	2.705015172	66%
PLP2	2.704603759	88%
SETD3	2.701306318	58%
A0A0A0MQ90	2.698801123	52%
TOP2A	2.690370738	50%
CAPG	2.664927486	86%
MCM5	2.650982754	91%
Q99LB4	2.649193605	84%
ERGI1	2.64735627	52%
CATZ	2.639800205	50%
SRSF9	2.617679294	54%
A0A0U1RP59	2.601792295	61%
PCNA	2.592876153	73%
E9PYH2	2.588704044	78%
IMPA2	2.586731817	70%
PFKAP	2.573353772	57%
CDK6	2.569498455	76%
TNC18	2.558653751	71%
F6UHR6	2.551821316	53%
RL7L	2.5501486	71%
A2BHP6	2.545877507	69%
ESPL1	2.534847943	52%
ITIH4	2.519288988	76%
BRI3B	2.518654628	58%
B2RPU8	2.494793056	51%
HAUS1	2.493639948	64%
MOT4	2.479640796	70%
MCM2	2.47509484	75%
MAVS	2.467043417	82%
Q3UNN4	2.459755016	64%
E9PZ97	2.459544987	68%
SPA3G	2.449753779	78%
TM9S4	2.446434989	67%
PRI2	2.441787802	73%
3BP1	2.437179528	52%
IIGP1	2.433734268	65%
CMTD1	2.433597794	56%
E9PWG6	2.429462828	78%
MCM6	2.417444737	75%
RIR1	2.403678379	77%
XYLT1	2.399761071	71%
Q52KC3	2.393455024	72%
RFC4	2.389911836	50%
FKBP2	2.377922052	59%
MCM7	2.374954349	75%
LRC59	2.365011568	77%
MCM3	2.364746749	77%
E9Q8A3	2.36025079	60%
PGK1	2.346306028	82%
ERO1A	2.340809611	64%
HAT1	2.337554928	67%
MARK4	2.337224834	77%
ARHG2	2.333715051	80%
DMAP1	2.329111577	58%
TBA1B	2.328380379	80%
DCAF8	2.325847354	76%
GALK1	2.302387039	78%
CNDP2	2.299759994	84%
Q8C605	2.290971047	80%
E9PW39	2.290770924	61%
SYVC	2.263121452	80%
NSDHL	2.240879661	66%
MA2B1	2.232547174	55%
SERC	2.22582669	74%
AB17C	2.224911281	91%
OGFR	2.223176253	54%
T2EA	2.215388461	73%
Q9CZN7	2.212182356	78%
CAC1H	2.201571585	51%
RPB7	2.189047433	53%
DPOLA	2.183557266	58%
PA1B3	2.180860215	64%
SNTB2	2.161517247	54%
A6ZI44	2.154366575	79%

CHD6	2.147292027	56%
TLK1	2.147093049	56%
STAT2	2.1467769	64%
XYLT1	2.130477264	70%
VIME	2.130196516	81%
GLYC	2.12924806	75%
E41L2	2.126801181	70%
PPP5	2.104629423	58%
TBB5	2.089422699	86%
ICAM1	2.088484151	75%
ENDD1	2.082881302	66%
PFKAL	2.078124351	64%
RNH2B	2.074334519	54%
PDCL3	2.064675767	51%
VIME	2.061522961	83%
MP2K1	2.058629765	53%
A0A0R4J138	2.042460314	65%
A0A0B4J1E7	2.029864194	58%
NFAC2	2.02483718	82%
ALDOA	2.023559372	81%
RPAP1	2.022005589	61%
RFC4	1.987279854	62%
STML2	1.985355633	72%
SND1	1.979233844	58%
DHB7	1.971263535	71%
GBG2	1.969401634	82%
MIF	1.968522247	52%
HEM6	1.962878898	73%
CDK2	1.944907077	50%
CHD4	1.935840861	52%
KPYM	1.929475817	85%
LDHA	1.912384368	72%
A0A0R4J170	1.911242674	56%
ADA	1.905315422	53%
A1L3S7	1.899933358	62%
OAT	1.892847691	73%
PPIF	1.890645245	51%
IMP3	1.859463566	51%
Q9QXJ2	1.84545505	74%
ARMX3	1.839322416	53%
A0A0A0MQF6	1.837020016	73%
PYR1	1.824300903	72%
Q3UE92	1.823171388	52%
TBB2B	1.821006703	77%
STN1	1.809933447	69%
PUR6	1.806052401	58%
IPO5	1.78536942	67%
ADH7	1.783634354	57%
CRADD	1.775646277	67%
4F2	1.758453166	78%
G3P	1.753386046	77%
D3Z7B5	1.721490095	59%
PDIA6	1.710162624	78%
ENPL	1.708465071	69%
CA198	1.702921616	55%
COX17	1.696419399	61%
PGAM1	1.692989898	80%
ROA1	1.692583453	61%
MYH7B	1.6795756	70%
SERA	1.669320065	82%
EF1G	1.669011448	77%
PSMG3	1.662425733	71%
OGA	1.658560256	57%
GRHPR	1.648207833	80%
MBB1A	1.645840468	75%
E9PUQ5	1.635799069	52%
A0A0R4J1E2	1.61621308	78%
MICA1	1.602330095	50%
P5CS	1.599882899	75%
MIF	1.597179865	87%
Q564E2	1.589577008	84%
PHB2	1.588524208	53%
DHX36	1.572690282	54%
ENOA	1.567309792	87%
TT21B	1.549117885	53%
E9QNP0	1.539566891	56%
TRAF6	1.522846749	70%
DHX29	1.519335041	50%
PSB7	1.510012511	57%
Q31093	1.498033381	70%

SYTC	1.497845411	66%
PGAM5	1.496061672	54%
SYMC	1.487068742	67%
RNH2C	1.477285553	62%
CCS	1.475447159	60%
M3K12	1.45785279	74%
RL23A	1.456132094	53%
Q8C2Q7	1.455006766	59%
RL26	1.446939715	68%
UMPS	1.432962884	53%
AB17C	1.430177199	54%
MOT1	1.413603563	56%
IMDH2	1.406692428	78%
TBA1B	1.389955705	59%
EF1B	1.389369287	80%
IPYR	1.383526393	55%
A0A0A0MQF6	1.371039936	55%
IF5A1	1.362846451	79%
SYEP	1.361720183	59%
GAPR1	1.357117614	57%
PTMA	1.354998578	76%
DAD1	1.352988021	54%
A0A0R4J0D3	1.351135828	59%
ADT2	1.344678287	72%
TBB5	1.344341977	55%
THIO	1.338384161	81%
EF1D	1.326206589	78%
NOP2	1.318071783	54%
RIPK3	1.316804421	50%
RS29	1.316718226	56%
IF4E	1.310838381	50%
BROX	1.309723463	54%
ZER1	1.287441456	70%
ADT1	1.279918167	54%
YBOX1	1.276349573	51%
WDR3	1.26512848	51%
LS14A	1.246971579	73%
PA2G4	1.214684399	75%
URP2	1.211433566	59%
SSRG	1.197982492	52%
RPN1	1.194014959	80%
Q8C2Q7	1.192867278	57%
HNRPQ	1.190828981	52%
MIC26	1.184683507	54%
SAP3	1.17806936	52%
TBA1A	1.17655849	64%
ERF1	1.169421755	55%
HPRT	1.146563114	56%
RL19	1.143070962	82%
Q8C3V4	1.141692881	55%
AATC	1.136687692	69%
EF2	1.133266328	80%
Q8BG13	1.128935928	54%
FKBP8	1.099878377	66%
E9PW43	1.096909626	57%
STAT1	1.085843415	50%
RL6	1.083052811	52%
RL19	1.08191587	81%
ITPA	1.081146549	51%
ROAA	1.080610158	72%
GRP78	1.0785641	75%
SNRPA	1.074266366	77%
S10AD	1.070294986	59%
PGES2	1.064467865	56%
RL27	1.062161512	58%
CASP7	1.057652884	51%
GPSM1	1.057065685	52%
EF1A1	1.051480432	69%
PHB	1.048392902	73%
IF2A	1.047079248	50%
UBA1	1.044998548	79%
G6PI	1.041903443	76%
LPXN	1.036876116	54%
RL10A	1.029569179	71%
RBBP4	1.028278181	52%
HDAC1	1.018936187	59%
PPIB	1.009237446	79%
EEA1	0.996993844	50%
DEK	0.988629714	65%
GBP1	0.984194799	57%

RL14	0.98302194	59%
CTL2	0.978788974	76%
HTR5B	0.973973679	88%
PTN6	0.971391346	62%
RAB10	0.971179383	59%
HNRH1	0.958551784	83%
RFA3	0.956030258	56%
TBB4B	0.953977622	68%
RL8	0.939331155	51%
Q3TCU5	0.93724761	66%
SFXN1	0.92869476	80%
RL27A	0.924006855	51%
Q9Z1A1	0.922285588	63%
AN32E	0.919597655	56%
KRT85	0.91739029	54%
MTA2	0.915037968	54%
A0A0A6YX26	0.910082515	74%
RAB8B	0.903000604	57%
C1TM	0.895060187	53%
TCPB	0.874490194	56%
OSTC	0.874058775	53%
RL7	0.864567996	60%
HS90B	0.864227695	76%
FUBP2	0.860930387	65%
STING	0.859196742	56%
PPIA	0.85869303	81%
YIPF4	0.853214957	57%
GDIB	0.844906766	66%
RL13	0.844492705	57%
SPEE	0.840533447	67%
A0A1B0GTA4	0.832047561	58%
RS25	0.829755529	67%
AT1A1	0.82115059	54%
A0A1L1SQA8	0.818715112	56%
1433F	0.81605988	70%
PNKP	0.815304654	77%
RL27A	0.79728746	56%
TKT	0.796464877	61%
A2A547	0.780974565	80%
TCPG	0.772720765	70%
TCPD	0.767603764	83%
RS27	0.761183195	61%
CN37	0.749809186	52%
PRPS1	0.740501349	57%
UBP3	0.731093474	55%
OST48	0.729405619	52%
PR40A	0.717399317	51%
Q6ZWZ7	0.704555893	51%
PRDX1	0.703942427	59%
HS90A	0.702117829	71%
RL31	0.689031861	60%
RL17	0.683684578	63%
RS3A	0.679492804	52%
RL35	0.674449786	52%
F208B	0.672359051	56%
RS14	0.668101312	55%
Q9D8L3	0.65764416	59%
1433E	0.657499632	54%
RL10A	0.6498865	66%
DNM1L	0.645518054	55%
RL31	0.635978157	70%
PCBP1	0.626722243	75%
IRGM1	0.612352468	61%
RS18	0.606782635	52%
RL12	0.596279164	50%
PCBP1	0.595370864	63%
RS7	0.593101251	50%
HNRPK	0.579491287	58%
PSB4	0.576415746	74%
Q6ZWZ4	0.576043999	50%
VDAC2	0.574906435	79%
UBCP1	0.549088285	54%
RLA2	0.547959766	50%
G3UYV7	0.540616148	68%
CATC	0.537165283	57%
NUCL	0.531891746	56%
PSDE	0.53161895	59%
FKB1A	0.517983025	58%
U2AF2	0.506144811	57%
RL29	0.481448909	57%

RLA0	0.464531573	69%
RSSA	0.456147974	61%
PNKP	0.398051591	55%
HSP7C	0.387396585	52%
RS28	0.335413777	54%
KAPCB	-0.470857914	55%
ARP5L	-0.550545379	53%
ARC1B	-0.564631198	65%
SORCN	-0.572122368	53%
NDUA4	-0.590489965	59%
ELMO1	-0.594409681	60%
CD5	-0.638900794	63%
NSF1C	-0.660526197	62%
CLCB	-0.666952343	53%
ANXA6	-0.691317171	64%
HCD2	-0.701369887	78%
BIN2	-0.712069363	54%
Q3UDM0	-0.717330187	55%
ASC	-0.729686054	70%
A0A0N4SVT1	-0.741054097	61%
EYA4	-0.744997917	68%
LASP1	-0.755690577	77%
Q99N15	-0.768015251	53%
ANX11	-0.777212561	64%
A0A0N4SVQ1	-0.778366989	64%
GSTP1	-0.784517523	69%
Q924B0	-0.793717767	66%
CAZA1	-0.799841036	60%
BIEA	-0.799958815	57%
CCD66	-0.802007204	61%
PTCA	-0.807579386	56%
INPP	-0.809855844	68%
PROF1	-0.811687843	78%
UBC9	-0.822117613	54%
CBX3	-0.834326117	73%
UBC9	-0.84329792	64%
GIMA4	-0.85554105	87%
IF4A3	-0.872231438	69%
LSM7	-0.891835471	66%
MGDP1	-0.903836934	63%
COR1A	-0.91111261	77%
NHP2	-0.914958865	51%
CRLF3	-0.927646997	75%
COF1	-0.94141141	54%
SP100	-0.944815874	80%
SAMH1	-0.94936271	78%
RPR1B	-0.956373236	80%
TAGL2	-0.956977577	68%
RAC2	-0.957107882	84%
A0A0R4J195	-0.959440335	70%
U2AF2	-0.962340629	61%
IL16	-0.96793139	84%
Q5RKN9	-0.993656526	54%
MGN	-1.003309619	59%
K22E	-1.0042561	51%
A0A0A6YY34	-1.004623611	68%
LAMP1	-1.008024769	52%
CNN2	-1.008254104	51%
GLRX1	-1.013137832	64%
E9QMV2	-1.016496052	81%
TRNT1	-1.036411946	50%
K1671	-1.044370197	72%
A2A513	-1.048432426	78%
TPM3	-1.052414196	56%
OPA1	-1.054233874	52%
Q6GT24	-1.066243206	72%
PRDX6	-1.073920222	50%
ALDR	-1.079451626	88%
CAP1	-1.087472527	70%
THIM	-1.089823745	62%
F8WGL3	-1.0922213	56%
MOES	-1.097251805	73%
NH2L1	-1.109045467	78%
H4	-1.114510167	71%
Q91VB8	-1.1209302	64%
A0A0R4J0J8	-1.127067758	57%
PKHA1	-1.127865275	53%
ANXA5	-1.141691596	79%
NOP58	-1.146317373	53%
ARHL2	-1.154401189	69%

NCK5L	-1.156445447	66%
THUM1	-1.162496586	62%
GDPP1	-1.165136767	59%
ADK	-1.175762857	68%
E9Q0F0	-1.190445254	51%
RMXL1	-1.192946013	54%
CATE	-1.193534417	89%
B2RUG9	-1.196912924	70%
PDXK	-1.199200137	52%
D39U1	-1.200422659	59%
FUBP1	-1.2076978	52%
HBB2	-1.226541892	61%
ROA2	-1.22733264	78%
DCPS	-1.246104259	75%
PI42A	-1.25004861	77%
F193A	-1.267376018	63%
LDHB	-1.289491892	78%
HNRPD	-1.30849955	68%
Q6ZWQ9	-1.316611498	58%
FBRL	-1.323596698	62%
S10AB	-1.329855686	70%
KBTBB	-1.332604766	65%
CO6A2	-1.343800204	55%
ROA3	-1.364167377	79%
COX7C	-1.368479124	84%
FETUA	-1.386146982	68%
CBX1	-1.397948782	60%
CATE	-1.407538019	81%
LIMD2	-1.409976024	76%
GDIR2	-1.415366181	71%
B7FAU9	-1.474612825	85%
ROA0	-1.479440772	60%
IL16	-1.506103682	77%
HNRPC	-1.514502888	61%
A2A6J4	-1.515731388	82%
FLNA	-1.524219641	82%
COTL1	-1.525748345	71%
SH3L3	-1.543965971	82%
BAG5	-1.574801689	82%
CBR1	-1.583023149	77%
HPCL1	-1.616369907	82%
LSP1	-1.628432214	85%
RASF2	-1.664829713	79%
PGM2	-1.667656052	72%
BRX1	-1.679977502	54%
MDHC	-1.708506835	87%
TMCC3	-1.743178097	51%
A2AQN4	-1.749531203	53%
TBC8B	-1.751181055	59%
THRB	-1.798281139	52%
ACD10	-1.829842923	54%
NUF2	-1.860929641	57%
K1C17	-1.91191341	58%
K1C16	-1.917822773	90%
DDAH1	-1.932109472	51%
ACOT2	-1.942605378	78%
Q9Z1R9	-1.948313756	85%
RASK	-1.970742154	56%
PKHO2	-2.013592645	51%
E9QJT5	-2.035008499	80%
EYA4	-2.041675267	51%
NUD16	-2.064480502	66%
GSHR	-2.10542411	66%
K1C10	-2.112424327	86%
SET1B	-2.121980488	71%
K22E	-2.167944004	61%
PI42A	-2.173163563	61%
SPT6H	-2.252464196	80%
MRCKB	-2.254648796	52%
E9QA45	-2.264238531	83%
H12	-2.295020308	51%
G3X8X3	-2.314798048	53%
IF4A2	-2.362177239	82%
DAAM2	-2.388110879	52%
K2C1B	-2.402019745	81%
SQRD	-2.419766151	58%
A2A513	-2.426828568	95%
K2C1	-2.456747494	78%
ACYP1	-2.472075023	76%
DHB11	-2.485943772	54%

E9Q0F0	-2.505614478	69%
ACTN1	-2.528755871	83%
K2C79	-2.545858794	87%
Q9EQ08	-2.589184426	56%
E9Q0A7	-2.607330026	55%
NAA35	-2.637379099	60%
D3YX85	-2.650315559	70%
NEST	-2.662024848	61%
GBG5	-2.688484175	60%
HDDC2	-2.717227558	65%
HYEP	-2.742116187	69%
K1C14	-2.787940779	71%
Q9Z1R9	-2.807432235	75%
HMGB1	-2.819757201	64%
CCZ1	-3.024920159	57%
UBE3C	-3.045184245	62%
A0A0R4J0J3	-3.209759287	63%
PAXB1	-3.358714664	53%
PSMD8	-3.386751388	50%
Q8BH78	-3.416332776	51%
2A5A	-3.513004252	66%
FARP1	-3.517023663	52%
NOG2	-3.758666749	52%
UH1BL	-3.761452963	55%
E9QA45	-3.823835615	88%
LMNB2	-3.881132184	55%
MIO	-4.058872946	79%
K2C73	-4.200880488	74%
RUSD2	-4.4907361	65%
AP3M2	-4.535437702	57%
D3Z6X7	-4.614218008	51%
HINT2	-4.884016944	59%
Q792Z1	-5.602066665	97%
A2CG76	4.683056553	55%
A2A4U6	4.004331271	54%
IL2RA	3.989840648	86%
A0A0R4J0F6	3.927736011	56%
M3K11	3.862699048	77%
F8WIA1	3.842171432	52%
PIGT	3.821267208	60%
BASI	3.747434177	74%
KAD4	3.658885869	61%
SUGP1	3.630801647	51%
PDCD5	3.546400758	51%
MSMO1	3.483693611	57%
GNAS1	3.459294559	58%
DNM3A	3.393564932	63%
S2539	3.366041142	64%
K3W4Q8	3.360816069	73%
KI67	3.340960314	79%
CD2	3.296968449	50%
BOREA	3.274421675	53%
ATIF1	3.233359566	66%
ERG28	3.158011445	66%
PYC	3.151068223	84%
CBX5	3.111914523	69%
SMC2	3.109652269	64%
SUH	3.09436734	55%
GTR1	3.023694626	64%
GPAA1	3.00142856	55%
TFR1	2.983734914	76%
TERF2	2.979699407	78%
Q3UKN6	2.95931301	61%
A0A0R4J0V1	2.955900269	70%
TNC18	2.90099338	65%
NDUA8	2.883069609	62%
RT06	2.879233288	58%
E9QP00	2.81844582	58%
CK074	2.813915429	65%
IMPA1	2.779843837	50%
NXP20	2.723781165	67%
PLP2	2.704603759	88%
TOP2A	2.690370738	50%
ERGI1	2.64735627	52%
SRSF9	2.617679294	54%
A0A0U1RP59	2.601792295	61%
F6UHR6	2.551821316	53%
RL7L	2.5501486	71%
ESPL1	2.534847943	52%
ITIH4	2.519288988	76%

BRI3B	2.518654628	58%
MOT4	2.479640796	70%
MAVS	2.467043417	82%
Q3UNN4	2.459755016	64%
E9PZ97	2.459544987	68%
TM9S4	2.446434989	67%
IIGP1	2.433734268	65%
CMTD1	2.433597794	56%
E9PWG6	2.429462828	78%
XYLT1	2.399761071	71%
RFC4	2.389911836	50%
FKBP2	2.377922052	59%
LRC59	2.365011568	77%
ERO1A	2.340809611	64%
DMAP1	2.329111577	58%
NSDHL	2.240879661	66%
Q9CZN7	2.212182356	78%
CAC1H	2.201571585	51%
RPB7	2.189047433	53%
SNTB2	2.161517247	54%
CHD6	2.147292027	56%
E41L2	2.126801181	70%
ICAM1	2.088484151	75%
ENDD1	2.082881302	66%
RNH2B	2.074334519	54%
VIME	2.061522961	83%
STML2	1.985355633	72%
SND1	1.979233844	58%
DHB7	1.971263535	71%
GBG2	1.969401634	82%
MIF	1.968522247	52%
CHD4	1.935840861	52%
A0A0R4J170	1.911242674	56%
A1L3S7	1.899933358	62%
OAT	1.892847691	73%
PPIF	1.890645245	51%
IMP3	1.859463566	51%
ADH7	1.783634354	57%
4F2	1.758453166	78%
PDIA6	1.710162624	78%
ENPL	1.708465071	69%
CA198	1.702921616	55%
ROA1	1.692583453	61%
MYH7B	1.6795756	70%
MBB1A	1.645840468	75%
E9PUQ5	1.635799069	52%
P5CS	1.599882899	75%
PHB2	1.588524208	53%
Q31093	1.498033381	70%
PGAM5	1.496061672	54%
Q8C2Q7	1.455006766	59%
AB17C	1.430177199	54%
MOT1	1.413603563	56%
TBA1B	1.389955705	59%
A0A0A0MQF6	1.371039936	55%
GAPR1	1.357117614	57%
DAD1	1.352988021	54%
A0A0R4J0D3	1.351135828	59%
ADT2	1.344678287	72%
TBB5	1.344341977	55%
NOP2	1.318071783	54%
ADT1	1.279918167	54%
WDR3	1.26512848	51%
LS14A	1.246971579	73%
SSRG	1.197982492	52%
RPN1	1.194014959	80%
HNRPQ	1.190828981	52%
MIC26	1.184683507	54%
FKBP8	1.099878377	66%
E9PW43	1.096909626	57%
RL19	1.08191587	81%
ROAA	1.080610158	72%
GRP78	1.0785641	75%
SNRPA	1.074266366	77%
PGES2	1.064467865	56%
RL27	1.062161512	58%
PHB	1.048392902	73%
RL10A	1.029569179	71%
RBBP4	1.028278181	52%
HDAC1	1.018936187	59%

PPIB	1.009237446	79%
DEK	0.988629714	65%
CTL2	0.978788974	76%
HNRH1	0.958551784	83%
Q3TCU5	0.93724761	66%
SFXN1	0.92869476	80%
MTA2	0.915037968	54%
RAB8B	0.903000604	57%
C1TM	0.895060187	53%
OSTC	0.874058775	53%
FUBP2	0.860930387	65%
STING	0.859196742	56%
YIPF4	0.853214957	57%
AT1A1	0.82115059	54%
PNKP	0.815304654	77%
RL27A	0.79728746	56%
CN37	0.749809186	52%
UBP3	0.731093474	55%
OST48	0.729405619	52%
PR40A	0.717399317	51%
RL31	0.689031861	60%
F208B	0.672359051	56%
Q9D8L3	0.65764416	59%
IRGM1	0.612352468	61%
PCBP1	0.595370864	63%
HNRPK	0.579491287	58%
VDAC2	0.574906435	79%
U2AF2	0.506144811	57%
NDUA4	-0.590489965	59%
CD5	-0.638900794	63%
HCD2	-0.701369887	78%
Q99N15	-0.768015251	53%
A0A0N4SVQ1	-0.778366989	64%
PTCA	-0.807579386	56%
UBC9	-0.84329792	64%
LAMP1	-1.008024769	52%
THIM	-1.089823745	62%
CATE	-1.193534417	89%
B2RUG9	-1.196912924	70%
D39U1	-1.200422659	59%
Q6ZWQ9	-1.316611498	58%
COX7C	-1.368479124	84%
LIMD2	-1.409976024	76%
IL16	-1.506103682	77%
TBC8B	-1.751181055	59%
ACD10	-1.829842923	54%
K1C17	-1.91191341	58%
K1C16	-1.917822773	90%
ACOT2	-1.942605378	78%
RASK	-1.970742154	56%
EYA4	-2.041675267	51%
K1C10	-2.112424327	86%
SET1B	-2.121980488	71%
K22E	-2.167944004	61%
PI42A	-2.173163563	61%
K2C1B	-2.402019745	81%
SQRD	-2.419766151	58%
A2A513	-2.426828568	95%
K2C1	-2.456747494	78%
DHB11	-2.485943772	54%
E9Q0F0	-2.505614478	69%
K2C79	-2.545858794	87%
Q9EQ08	-2.589184426	56%
E9Q0A7	-2.607330026	55%
NAA35	-2.637379099	60%
D3YX85	-2.650315559	70%
NEST	-2.662024848	61%
GBG5	-2.688484175	60%
HDDC2	-2.717227558	65%
HYEP	-2.742116187	69%
K1C14	-2.787940779	71%
Q9Z1R9	-2.807432235	75%
HMGB1	-2.819757201	64%
A0A0R4J0J3	-3.209759287	63%
PSMD8	-3.386751388	50%
Q8BH78	-3.416332776	51%
FARP1	-3.517023663	52%
UH1BL	-3.761452963	55%
E9QA45	-3.823835615	88%
MIO	-4.058872946	79%

K2C73	-4.200880488	74%
AP3M2	-4.535437702	57%
Q792Z1	-5.602066665	97%

Table 12A. List of upregulated and downregulated proteins in iTregs vs Stim nTregs, sorted based on the fold-change and Confidence score ≥50%.

Protein	Fold-change	Confidence score
A2CG76	5.657989106	56%
M3K11	4.509704489	76%
DUS2L	4.14225154	63%
NDRG2	4.067892127	66%
A2A4U6	4.005840853	51%
PYRD	3.926554798	52%
E9Q921	3.905628453	59%
PLP2	3.577217635	63%
E9PVS1	3.475987236	57%
GCSH	3.460801273	51%
AT1B1	3.460268379	73%
LS14A	3.340596745	66%
RIPL1	3.331845328	50%
PTH2	3.132438217	52%
HSDL1	3.127337017	62%
RM01	3.114133088	69%
HEM6	2.896277154	59%
E9Q9M1	2.85709485	54%
RT33	2.856289003	59%
SYAM	2.807181974	53%
COR1C	2.777681764	52%
NEP1	2.753260303	53%
PDLI4	2.713160944	63%
ARL1	2.694743401	60%
A2AUM0	2.667574288	55%
A0A0R4J1W7	2.635598641	58%
E9Q1R5	2.591113778	57%
Q3UMA3	2.537521504	58%
HBB2	2.50627012	82%
SNX27	2.465747009	54%
XYLT1	2.427943035	60%
ITIH4	2.369637137	75%
H7BWX9	2.212874309	56%
EXOC3	2.195117169	64%
ENDD1	2.055316868	72%
WASH1	2.046238744	55%
BABA1	2.044312994	52%
Q9EQ08	1.952920588	61%
VAMP7	1.946837719	58%
A0A0N4SVQ1	1.927118411	88%
MAVS	1.918304864	76%
UB2D1	1.881335457	51%
CE162	1.814775525	64%
NEST	1.811374492	74%
RRAS	1.794195303	52%
F7DEU6	1.793348831	52%
HBB1	1.782801113	55%
WDR37	1.701282608	51%
SUV3	1.676663277	58%
NAA35	1.676247159	64%
NDUA4	1.674278579	84%
HBA	1.620753196	57%
SP110	1.615521212	81%
FOLH1	1.559357666	77%
RAB35	1.538967971	57%
F8VPV0	1.515456878	68%
A0A0R4J0V1	1.45737645	51%
RNH2B	1.449446056	56%
Q91VB8	1.448690555	54%
CASP7	1.400496445	54%
A2AP32	1.36643211	60%
K1C17	1.359503912	57%
K2C79	1.357875187	64%
H32	1.311871776	61%
MYH9	1.292172134	68%
NDUB8	1.279772753	70%
ODPX	1.275758488	66%

NDUA5	1.273114407	69%
ML12B	1.271526643	78%
CATC	1.259888432	62%
NDUS7	1.2571113	70%
MBRL	1.248397752	51%
Q91VB8	1.243211965	51%
CYH3	1.233364146	70%
H4	1.226259047	52%
CS012	1.206407294	82%
CILP1	1.195248824	64%
D3YX85	1.189041772	71%
A2A513	1.18756759	63%
AL3A1	1.145773572	79%
SUN2	1.126600942	72%
CHD9	1.117373767	68%
UBCP1	1.057800534	70%
ALBU	1.049935047	61%
ECHB	1.049259836	79%
COTL1	1.033433252	51%
ADT2	1.022358684	68%
PNKP	0.983430056	76%
P85A	0.954164446	63%
SRSF1	0.951079509	85%
E9QA45	0.93853452	86%
Q9DCC5	0.901269961	54%
COX6C	0.901238868	58%
IDHG1	0.900135772	62%
DJC16	0.892247918	52%
CX6B1	0.8876547	79%
ALDR	0.878933723	78%
SUOX	0.870261818	57%
ATPO	0.860886629	65%
A0A0A6YY34	0.859472753	78%
ECHA	0.858119424	66%
PHB2	0.856149692	67%
TT39B	0.855707346	60%
B7FAU9	0.814184712	64%
THIL	0.813818443	51%
PYRD	0.809476313	52%
MIO	0.806192253	67%
Q8C3V4	0.789784289	67%
THIM	0.786348742	52%
NDUB3	0.775349041	71%
ADA10	0.752922619	54%
NDUS3	0.748538153	85%
STAT1	0.745990076	73%
COR1C	0.735238075	59%
CD47	0.731799738	53%
NDUBB	0.726149384	50%
FBXL8	0.725992309	54%
ATP5L	0.718152104	51%
F6QFD1	0.714715698	72%
Q6PFB2	0.710464373	53%
BAG5	0.678019468	65%
IRGM1	0.673175571	70%
INPP	0.671660412	53%
PHB	0.670677363	71%
PR40A	0.66950428	52%
VIME	0.649926391	67%
ATPA	0.641585681	54%
Q99LB4	0.640830052	77%
CN37	0.636946223	52%
CAPG	0.624199188	60%
SUMO1	0.622779553	52%
ACON	0.621769615	53%
Q31093	0.615682627	54%
NDUV2	0.591516994	54%
SUMO2	0.580639633	90%
CX7A2	0.545158408	65%
A2BIE1	0.530821116	57%
DNPEP	0.521207988	53%
GRAP2	0.513419007	57%
ATPD	0.471001881	62%
UH1BL	0.452031222	56%
VDAC2	0.430430054	63%
TNC18	0.422158575	59%
D3Z6X7	-0.275591631	52%
RS28	-0.312331185	67%
A2A6J4	-0.333886045	50%
GPSM1	-0.364854389	55%

RAC2	-0.364936305	63%
NDKB	-0.383120763	50%
SERA	-0.394649492	58%
LASP1	-0.424552177	56%
PYM1	-0.425801219	61%
PROF1	-0.429147049	56%
PEBP1	-0.429930692	71%
PCBP1	-0.431425263	73%
PSDE	-0.433821968	60%
RL31	-0.43945428	54%
COR1A	-0.44297623	73%
OSTC	-0.444386077	54%
ARF1	-0.445214421	51%
DNM1L	-0.445980346	59%
PSA3	-0.45258416	52%
RSSA	-0.452784663	71%
A0A0A6YX26	-0.458164129	55%
COPD	-0.464171455	52%
HSP7C	-0.472771834	66%
GIMA4	-0.48893402	68%
GALE	-0.491554274	61%
RL31	-0.492490454	80%
CBX5	-0.494227663	53%
IMB1	-0.518859931	65%
RS27	-0.519428679	58%
Q9D8L3	-0.526091815	55%
DDX5	-0.528915817	52%
A0A1L1SQA8	-0.529029258	64%
Q9CQF7	-0.530895119	56%
ADH7	-0.531422602	53%
NHRF1	-0.531432639	51%
TM109	-0.539196937	62%
CAC1H	-0.540208906	74%
SH3L1	-0.542024535	59%
RINI	-0.542998092	54%
PABP1	-0.545263988	61%
Q9Z1A1	-0.546320164	55%
F208B	-0.546330964	54%
RLA0	-0.546458547	77%
RL10A	-0.547849831	64%
RAB5A	-0.548893103	50%
SKP1	-0.548916364	59%
GRHPR	-0.549776309	74%
TEX10	-0.551448659	55%
RL27A	-0.552095235	64%
LRRF1	-0.556288561	56%
A0A0A6YX26	-0.561687475	68%
STIP1	-0.570366259	58%
ACTG	-0.571484421	70%
EYA4	-0.572955285	63%
RBX1	-0.574720781	51%
THMS1	-0.577589186	52%
RS3A	-0.577625148	63%
Q8BFQ1	-0.579630725	62%
RL14	-0.584184048	64%
SYTC	-0.585195661	56%
Q3U2G2	-0.585381568	72%
PDIA3	-0.586011844	61%
A0A1B0GSG5	-0.590234591	59%
AIMP1	-0.59131329	63%
HS90B	-0.592805791	69%
LRC47	-0.592955984	54%
SRP14	-0.595183548	58%
BCL10	-0.5957755	57%
ARP3	-0.597094364	61%
ARBK1	-0.59980922	59%
CLIC1	-0.601378279	72%
RLA2	-0.603592092	62%
IST1	-0.60453529	56%
TES	-0.604866284	64%
NUCL	-0.605967316	69%
LY9	-0.608765695	51%
Q6ZWZ7	-0.61058827	60%
COF1	-0.615651332	77%
RL17	-0.616208346	70%
IF1A	-0.616701099	69%
UB2L3	-0.616890281	63%
Q91YS7	-0.616991017	51%
PLSL	-0.618190397	79%
THADA	-0.618289837	55%

IGSF8	-0.618601797	58%
F7A8H6	-0.619015891	57%
RL10A	-0.619138213	64%
CDK2	-0.620400764	76%
RL17	-0.621325117	54%
ELOB	-0.623792322	53%
G3X8R0	-0.626636319	58%
CAPZB	-0.627358205	75%
HINT1	-0.629229862	71%
HS90A	-0.629554533	53%
RUVB2	-0.634754435	56%
IMDH2	-0.635863747	50%
RL27A	-0.639486587	69%
CAB39	-0.640744027	62%
WDR1	-0.642126633	51%
RL35	-0.644770395	62%
CLCB	-0.64751973	56%
PDC6I	-0.6478963	67%
BAIP2	-0.648283668	65%
RUVB1	-0.649305483	76%
GLRX1	-0.654636642	69%
PFD2	-0.655957673	63%
PTPRC	-0.661490977	51%
HNRPD	-0.673618912	55%
SAHH	-0.674723684	72%
CHM4B	-0.683315722	61%
RS25	-0.685878079	70%
ITB2	-0.686032425	63%
THIO	-0.693809938	63%
FAS	-0.694085768	71%
RL34	-0.69580973	72%
SSRA	-0.696024299	73%
KBTBB	-0.696572229	71%
NAGK	-0.69723095	55%
ITAL	-0.697550865	59%
PTMA	-0.700170227	56%
SND1	-0.701821488	55%
DHPR	-0.703019135	71%
SSU72	-0.70404831	70%
RS20	-0.707340262	88%
TPM4	-0.709718354	75%
KAD2	-0.709909136	57%
A0A1C7CYV0	-0.710269825	68%
A0A1B0GTA4	-0.712791894	59%
SNX6	-0.712898929	58%
1433G	-0.720347614	74%
MIF	-0.727398692	85%
IF4G1	-0.732267177	77%
ARC1B	-0.732905424	84%
GBLP	-0.734039823	73%
IF5	-0.740944197	71%
SAP3	-0.751000506	67%
G6PI	-0.752007616	68%
Q3TJ22	-0.752525277	56%
RL4	-0.755736292	56%
PKN1	-0.75736667	59%
FKBP4	-0.759652567	53%
NUCL	-0.76058686	68%
HA1D	-0.76255854	78%
FLOT1	-0.766614316	57%
Q564E2	-0.772561477	71%
ECHD1	-0.775885802	57%
RLA1	-0.780759604	78%
E9QAZ2	-0.78121233	50%
ZER1	-0.782658854	86%
SYG	-0.782765896	67%
6PGD	-0.78782033	76%
E9PW43	-0.789526888	62%
SYSC	-0.797984072	82%
CYC	-0.801320953	59%
TCB1	-0.804109032	72%
RAB2A	-0.805456372	56%
Q3UE92	-0.814886376	60%
RS28	-0.817269249	53%
PCNA	-0.8230427	71%
LANC1	-0.824427021	51%
SERB	-0.825239365	59%
1433E	-0.82756698	63%
FKBP5	-0.8276963	63%
UGPA	-0.83133097	53%

GAPR1	-0.834927111	85%
Z4YJT3	-0.835694064	54%
LDHA	-0.838035987	71%
RL37A	-0.839208766	59%
PGP	-0.840459284	56%
A2APM5	-0.842048962	56%
PPIB	-0.842465586	51%
TAGL2	-0.844516828	68%
URP2	-0.844781463	79%
M3K12	-0.845842731	63%
SYYC	-0.85137823	82%
FXYD5	-0.856899011	52%
ACTC	-0.858333468	78%
GBB2	-0.858984432	50%
HPRT	-0.859299361	53%
SR140	-0.863493335	55%
TCPQ	-0.867847089	50%
G3UYV7	-0.871456166	53%
EXOC4	-0.872539541	53%
PFKAM	-0.872931711	59%
PP1G	-0.876932607	65%
MVD1	-0.879433116	56%
PAIRB	-0.886823037	83%
IF2G	-0.889014446	75%
NSDHL	-0.891372017	52%
ENOA	-0.892003652	83%
RS3	-0.906220052	54%
PFKAL	-0.90760953	78%
CALX	-0.909398202	53%
PIN1	-0.909785692	54%
SMYD5	-0.911940648	60%
Q8C605	-0.921885286	77%
EF1B	-0.92200363	60%
CD2	-0.925669425	83%
GMFB	-0.942273999	54%
MARK4	-0.94392759	51%
PGAM1	-0.945093052	61%
GLU2B	-0.953378083	56%
HAUS1	-0.957128854	71%
RLA2	-0.958415818	79%
FKB1A	-0.973418195	91%
RSSA	-0.97365636	61%
GLU2B	-0.977127985	52%
SYAC	-0.979640987	76%
AEDO	-0.996255392	53%
PDIA3	-0.996726904	78%
E9QA45	-0.998134649	78%
S10AB	-1.015138187	76%
KCC4	-1.01561214	62%
ALDOC	-1.017960262	82%
EP400	-1.022304838	55%
RU1C	-1.027610045	56%
CORO7	-1.027847783	84%
A0A0R4J0Z1	-1.035481646	67%
FPPS	-1.054271408	88%
GALK1	-1.055480896	72%
NDRG1	-1.05669919	96%
KIF15	-1.059074097	74%
PDIA4	-1.064455457	67%
PSB6	-1.066383133	54%
F7DEU6	-1.080333131	72%
A0A087WR57	-1.085449633	67%
HNRPM	-1.08815051	64%
GRP78	-1.101663232	80%
LEG1	-1.107036128	71%
GLGB	-1.112720906	54%
SERC	-1.114576146	72%
LRCH1	-1.117817428	65%
ZFAN6	-1.120131526	52%
PPCE	-1.145191178	84%
PGM1	-1.147275203	82%
KPYM	-1.149878496	82%
HYOU1	-1.154015606	58%
Q9Z1R9	-1.162291571	63%
RS20	-1.167131196	71%
TCOF	-1.182444543	56%
A0A0A0MQF6	-1.192087026	76%
ERP44	-1.195326517	52%
RL7A	-1.196419172	73%
RAD50	-1.196853691	54%

STRN4	-1.198396157	60%
STS	-1.211731742	53%
KI13A	-1.226718251	56%
SDF2L	-1.236848012	81%
LDHA	-1.244345608	51%
PRDX5	-1.24742343	77%
PFKAP	-1.252421892	82%
G3P	-1.263757268	62%
TXD12	-1.281562117	62%
TT21B	-1.288855879	69%
RIPK3	-1.305234238	87%
PDIA6	-1.32030456	81%
TLK1	-1.321792855	69%
PGK1	-1.329200545	77%
CCS	-1.342127574	73%
NOP2	-1.343842007	64%
A6ZI44	-1.35945162	83%
E9PUA7	-1.359682024	63%
SODC	-1.370081696	62%
G3X920	-1.37224243	54%
CH10	-1.389566424	67%
AATM	-1.393070268	85%
DHC24	-1.400358972	58%
SPA3G	-1.403353719	66%
CD48	-1.403672351	58%
ALKB5	-1.404268477	54%
TXLNA	-1.41680287	54%
MDHM	-1.418027854	77%
ALDOA	-1.421907306	86%
VATF	-1.430204492	51%
DYST	-1.436120684	55%
EGLN1	-1.436960932	60%
ENPL	-1.448612715	78%
ACLY	-1.453117773	95%
APOB	-1.457038943	67%
GAPR1	-1.463095218	53%
A0A0B4J1E7	-1.47102055	72%
SPA3F	-1.476304176	88%
MT2	-1.480386624	65%
S10AD	-1.48860292	58%
ENPL	-1.500172608	71%
MOT4	-1.501815381	82%
G3UZK1	-1.507009855	70%
RS27A	-1.520178599	54%
A0A087WSR7	-1.523329165	55%
BASI	-1.525014471	72%
HMGA1	-1.52669643	57%
GRP78	-1.536138384	78%
PACN2	-1.558201563	55%
NUCB1	-1.56056157	61%
CDK6	-1.563025965	76%
CC137	-1.564090467	51%
KPYM	-1.564594298	72%
KCRB	-1.566737076	81%
RHG22	-1.578408503	68%
F6ZHD8	-1.588726336	67%
IL2RA	-1.597345775	69%
SEC20	-1.617906244	64%
COX7C	-1.629280439	67%
AB17C	-1.630170069	91%
G5E8E4	-1.656800895	77%
PDIA1	-1.678518041	75%
AB17C	-1.681696409	81%
GTR1	-1.683074089	77%
HXK2	-1.685988105	76%
K3W4Q8	-1.696090034	72%
PDIA6	-1.71458761	85%
SATB1	-1.719932023	81%
PLCD1	-1.726917233	58%
PRAF3	-1.738156175	72%
SATB2	-1.781314873	85%
H2AV	-1.826444551	71%
CASPC	-1.854176448	52%
CASP6	-1.867183968	56%
SHOC2	-1.868214409	53%
RBX1	-1.881239854	58%
MT1	-1.886086913	74%
E9Q718	-1.891678897	62%
PDIA1	-1.915141515	75%
D3YVZ9	-1.984119994	67%

ERO1A	-2.00943444	79%
MTMR2	-2.012766593	55%
MIO	-2.023700216	67%
GNAS1	-2.040515325	58%
NUSAP	-2.044346518	52%
XPOT	-2.052410116	56%
ERO1A	-2.107882786	82%
PLXA2	-2.108376518	66%
FAAA	-2.169981388	51%
D3YXP6	-2.181594237	53%
ANGP4	-2.217129801	61%
F7AAP4	-2.220912727	68%
ZN593	-2.251456661	60%
TPIS	-2.325213721	81%
CEPT1	-2.350904385	60%
H12	-2.355236798	53%
GRIN3	-2.367632202	64%
SCAM3	-2.37872277	51%
KRT85	-2.38792522	67%
NDRG2	-2.398247871	65%
SPA3G	-2.412715444	55%
ANXA2	-2.430981498	84%
SACS	-2.446702163	52%
SATB2	-2.470664621	55%
NDST1	-2.533993197	56%
HMCS1	-2.544196329	83%
SDF2L	-2.722122677	62%
A0A140T8R5	-2.731680021	55%
SC61G	-2.731768844	79%
PROS	-2.767237838	60%
GTR3	-2.779761605	61%
CO4B	-2.810133413	57%
CS025	-2.847717997	67%
PSMD8	-2.848313712	57%
Q8CBB6	-2.870354834	84%
DDX54	-2.887269522	64%
H14	-2.905267883	78%
H2A1K	-2.953415335	69%
H15	-2.967253795	75%
E9Q9B0	-2.975441426	54%
CSTN1	-3.01698865	56%
A0A140T8R5	-3.044404268	53%
NPAT	-3.048176435	51%
ASB6	-3.074628505	70%
Q792Z1	-3.08881418	90%
ITB3	-3.10493335	67%
JUND	-3.123777656	56%
SMU1	-3.135291138	62%
CHD9	-3.149055364	50%
SSF1	-3.163700243	55%
CS043	-3.165420134	60%
NEK9	-3.167121902	60%
GRAB	-3.175869528	87%
SBP1	-3.205841696	91%
GRAB	-3.221427098	77%
KLC3	-3.290931043	54%
GTR3	-3.374183916	88%
CCD91	-3.389296045	52%
CR1L	-3.414490283	54%
SF3B5	-3.460499755	52%
ECM1	-3.511618282	77%
CLP1L	-3.550379187	57%
UBE3C	-3.571273246	65%
H4	-3.587128446	81%
Q8C253	-3.59067114	71%
SMAG2	-3.608447137	52%
F8WI14	-3.649687631	85%
ATG3	-3.697168054	52%
F8WI14	-3.712162613	92%
FOLH1	-3.755971889	78%
ECM1	-3.815754469	83%
G3X920	-3.93081132	73%
H13	-3.933487547	70%
F6V6T4	-3.960380673	54%
HPDL	-4.781425397	57%
B1ASU9	-6.288779263	67%