T and B cell collaboration in response to vaccination

Xin Gao

January /2023

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

© Copyright by Xin Gao 2023

All Rights Reserved

STATEMENT OF ORIGINALITY

The research presented in this thesis was performed in the Malaria Immunology Laboratory at the John Curtin School of Medical Research of the Australian National University under the supervision of Professor Ian Cockburn. The data herein has not been previously submitted for a degree and is the result of my own original research, with contributions from others indicated.

_____

Xin Gao

_____

Professor Ian Cockburn (supervisor)

ACKNOWLEDGMENTS

Professor Ian Cockburn

Thank you for your encouragement, trust, guidance and full support (both mentally and scientifically) throughout my PhD journey. A student cannot ask for a better supervisor like you who is both a true scientist, a kind mentor and a generous team leader. The so many things that I learned from you will keep inspiring me in the future.

Professor Di Yu

I am always grateful to you for being my guide to the field of immunology.

The Cockburn Group Members

Many thanks to all the members at the Cockburn Group, especially to Yeping, Hayley, Harry, James, Kai, Deepyan, Lauren, Hannah and Patricia for keeping the lab running efficiently, for having the mosquitos to grow happily and for maintaining the work environment to be friendly and supportive.

The Immunology Division

Thanks to all the people in the immunology division for creating such a friendly and collaborative atmosphere. Special thanks to the group leaders and the members of the Enders Group, Tscharke Group, Vinuesa Group, Cook Group and Man Group for providing suggestions and supports to my project.

Imaging and flow cytometry facility

A lot of thanks to Harpreet, Mick and Cathy for being professional and supportive to my cell sorting, flow cytometry and imaging experiments.

Australian Phenomics Facility

Many thanks to the work performed by APF, especially to Rob for looking after the animals that were used in this project, and Dominik for helping with the transgenic mouse generation.

Biomolecular Resource Facility

Many thanks to BRF, especially to Max for doing all the library construction and sequencing work to generate the high-quality RNA-seq datasets.

Center for Advanced Microscopy

Many thanks to CAM, especially to Jo and Melanie for helping with the electron microscopy.

Friends at JCSMR

Many thanks to my friends at JCSMR, especially to Martin, Jim, Anna, Kristy, Yuwei, Cheng, Shouya and Ali. Thanks for the joy and fun we experienced together. BTW, you guys never let me down when I need to pinch some reagents from you for a trial experiment.

Family

Special thanks to my mom and dad for the love and care throughout my PhD journey.

Le Xiong

Studying PhD was extremely challenging. Thanks for always giving me the comfort and encouragement when things were not right. When I presented my Zeb2 story at the ASI 2022 meeting I saw you and Ian were standing together at the back of the conference room. I guarantee that was one of the best moments in my life and this will keep encouraging me in the future. I am so thankful to have you by my side.

PUBLICATIONS

Gao X, Shen Q, Roco JA, Frith K, Munier CML, Nekrasov M, et al. ZEB2 regulates the development of CD11c+ atypical B cells. bioRxiv. 2022.
https://doi.org/10.1101/2022.09.01.506173
Represents contents in Chapter2 and Chapter4

Gao X, Luo K, Wang D, Wei Y, Yao Y, Deng J, et al. Type 17 Follicular Helper T (Tfh17) Cells are Superior for Memory Maintenance. bioRxiv. 2022.
https://doi.org/10.1101/2022.07.31.502219
Under revision by at Elife, represents contents in Chapter2 and Chapter3

Gao X, Cockburn IA. The development and function of CD11c+ atypical B cells-insights from single cell analysis. Frontiers in Immunology. 2022;13.
https://doi.org/10.3389/fimmu.2022.979060
Represents contents in Chapter1

Chen Z, **Gao X**, Yu D. Longevity of vaccine protection: Immunological mechanism, assessment methods, and improving strategy. View. 2022;3(1):20200103.

Sutton HJ, Aye R, Idris AH, Vistein R, Nduati E, Kai O, et al. Atypical B cells are part of an alternative lineage of B cells that participates in responses to vaccination and infection in humans. Cell reports. 2021;34(6):108684.

Chatterjee D, Lewis FJ, Sutton HJ, Kaczmarski JA, **Gao X**, Cai Y, et al. Avid binding by B cells to the *Plasmodium* circumsporozoite protein repeat suppresses responses to protective subdominant epitopes. Cell reports. 2021;35(2):108996.

Gao X, Wang H, Chen Z, Zhou P, Yu D. An optimized method to differentiate mouse follicular helper T cells in vitro. Cellular & molecular immunology. 2020;17(7):779-81.
https://doi.org/10.1038/s41423-019-0329-7
Represents contents in Chapter3

ABBREVIATIONS

ABC	Atypical B cells
AIM	Antigen induced marker
APC	Antigen presenting cell
AUC	Area under curve
BCR	B cell receptor
CCP	Clathrin coated pit
CDR	Complementarity-determining regions
CITE-seq	Cellular indexing of transcriptomes and epitopes by sequencing
CM	Central memory
cMBC	Classical memory B cells
CRISPR-Cas9	Clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9
cTfh	Blood circulating Tfh cells
DC	Dendritic cells
DPBS	Dulbecco's phosphate-buffered saline
DTR	Diphtheria toxin receptor
EM	Effector memory
FBS	Fetal Bovine Serum
FDC	Follicular dendritic cells
GC	Germinal center
GFP	Green fluorescent protein
GSEA	Gene set enrichment analysis
HEL	Hen egg lysozyme
iTfh1/2/17	In vitro culture induced Tfh1/2/17-like cells
KO	Knock out
LPS	Lipopolysaccharide
MWS	Mowat Wilson syndrome
OVA	Ovalbumin
PB	Plasmablast
PbCSP	*Plasmodium Berghei* circumsporozoite protein

PC	Plasma cell
PfCSP	*Plasmodium falciparum* circumsporozoite protein
RBC	Red blood cell
rCSP	Recombinant *falciparum* circumsporozoite protein
RNP	Ribonuclear protein
RT-qPCR	Reverse transcription quantitative real-time PCR
scRNA-seq	Single-cell RNA sequencing
SEM	Scanning electronic microscopy
sgRNA	Single guide RNA
SHM	Somatic hypermutation
SLE	Systemic lupus erythematosus
SPZ	*Plasmodium* sporozoite
TCR	T cell receptor
TF	Transcription factor
Tfh	Follicular helper T cells
UMAP	Uniform manifold approximation and projection
WT	Wild type

ABSTRACT

The invention of vaccines has profoundly reduced the burden of infectious disease worldwide. However, for some infectious diseases including malaria, HIV and tuberculosis we have struggled to develop vaccines that give enduring protection. Long lasting antibody responses depend on the collaboration of T cells with B cells to form memory cells and bone marrow plasma cells. Thus, my thesis aims at addressing the basic questions in T and B cell biology, hoping to provide new insights into strategies for vaccination. In my thesis I make three contributions to this area.

Firstly, I report that Tfh17 cells are superior in immunological memory maintenance. Memory Tfh cells are known to contribute to protection by enhancing humoral responses upon antigen re-exposure but how they are maintained is poorly understood. Memory Tfh are divided into Tfh1, Tfh2 and Tfh17 cells by the expression of CXCR3 and CCR6 respectively. Using a method to differentiate murine Tfh1, Tfh2 and Tfh17-like (iTfh1, iTfh2 and iTfh17) cells in vitro, I found that iTfh17 cells were superior to iTfh1 and iTfh2 cells, and have survival and proliferative advantages. Importantly, analysis of multiple human cohorts that received different vaccines for HBV, influenza virus, tetanus toxin or measles revealed that vaccine-specific Tfh17 cells outcompeted Tfh1 or Tfh2 cells for the persistence in the memory phase.

Next, I investigated the role of a poorly characterized population of B cells called atypical B cells (ABCs) in vaccination. These cells are abundant in chronic disease and autoimmune conditions, but also seen in responses to vaccination. However, the factors driving their development and maintenance are unknown. Here I show that zinc finger E-box binding homeobox 2 (*ZEB2*) is the key transcriptional factor (TF) of ABCs formation. I generated a human tonsil single-cell RNA-seq (scRNA-seq) dataset and showed that ABCs had high *ZEB2* expression. ABC formation was severely impaired among *CD23*$^{Cre/+}$*Zeb2*$^{fl/fl}$ B cells after *Plasmodium* infection or immunization in mice. In humans I found that *ZEB2* haploinsufficient Mowat Wilson syndrome patients also had low numbers of circulating ABCs. Surprisingly, *Zeb2* deficiency had no impact on other aspects of the B cell response in both mice and humans in these systems, suggesting a minimal role for ABCs in healthy immune responses.

Finally, I investigated how B cells obtain antigen from complex pathogens for presentation for T cells. While B cell responses to recombinant proteins and small viruses are well studied, the interaction with complex eukaryotic pathogens is poorly understood. Here I discovered antigen specific B cells used a unique “surface pinching” mechanism to extract antigens from the

Plasmodium sporozoites (SPZs). By visualizing the interaction of *Plasmodium falciparum* circumsporozoite protein (PfCSP) specific Igh^{g2A10} B cells with PfCSP expressing SPZs (PfCSP-SPZs), I found Igh^{g2A10} B cells directly "pinched off" surface antigens from PfCSP-SPZs. scRNA-seq analysis confirmed the repertoire of Igh^{g2A10} helping Tfh cells were mainly PfCSP specific upon SPZ immunization. I further validated this by using the SPZs strains expressing OVA on spatially different locus and showed that Igh^{g2A10} B cells could only be helped by OT-II cells when OVA was located on the surface but not interior of the SPZs.

TABLE OF CONTENTS

Chapter1- Introduction 1

1.1 The immune system 2

1.2 $CD4^+$ T cells 2

1.3 B cells 8

1.4 Tfh cells and the germinal center reaction 13

1.5 Memory B cell responses: the role of atypical memory B cells (ABCs) 18

1.6 B cell antigen uptake and presentation 28

1.7 *Plasmodium* parasites 31

1.8 Key technologies used in this thesis 37

1.9 Aims and rational for this study 39

Chapter2- Material and Methods 42

2.1 Acknowledgement of major contributions 43

2.2 Buffers/ solutions/ media, reagents and antibodies 43

2.3 Human samples and ethics statement 52

2.4 Mice strains and ethics statement 53

2.5 Parasite 55

2.6 Human sample experiments 56

2.7 Mouse experiments 57

2.8 Cellular and molecular experiments 58

2.9 Flow cytometry, imaging, bioinformatic analysis and statistics 62

Chapter3- Type 17 Follicular Helper T (Tfh17) Cells are Superior for Immunological Memory Maintenance 65

3.1 Introduction 66

3.2 Results 67

3.3 Discussion 101

Chapter4- ZEB2 Regulates the Development of CD11c+ Atypical B Cells 104

4.1 Introduction 105

4.2 Results 107

4.3 Discussion 134

Chapter5- The protective B-cell and cognate T-cell epitopes are spatially restricted in response to large parasite invasion 136

5.1 Introduction 137

5.2 Results 139

5.3 Discussion 157

Chapter6- Discussion 160

6.1 Understand the mechanism of T cell persistence and to improve the longevity of vaccine-induced protection 161

6.2 The development and functions of ABCs, and its possible role in “gate-keeping” the antibody affinity 162

6.3 Insights in improving the malaria vaccine 164

6.4 Concluding remarks 165

Chapter7- References 167

LIST OF FIGURES

Figure	Title	Page
1.1	TCR and VDJ recombination	4
1.2	The differentiation and function of effector $CD4^+$ T cell subsets	7
1.3	BCR and VDJ recombination	9
1.4	The development of naïve B cells and the differentiation of effector B cell subsets	12
1.5	Summary of the phenotype, development, and function of Tfh and cTfh cells	15
1.6	scRNA-seq revealed the conserved presence of ABCs in human	20
1.7	B cell antigen uptake and presentation through clathrin dependent endocytosis	30
1.8	The life cycle of the *Plasmodium* parasite	32
1.9	The structure of sporozoite and CSP protein	35
3.1	Optimization and characterisation for *in vitro* differentiation of mouse Tfh cells	70
3.2	The *in vitro* differentiation of induced Tfh1, Tfh2 and Tfh17-like (iTfh1, iTfh2, iTfh17) cells	73
3.3	iTfh1/2/17 cells form Tfh cells after immunization and retain polarized Th1/2/17 features	76
3.4	iTfh17 cells are superior in memory maintenance	78
3.5	Human Tfh_{CM} and Tfh_{EM} subsets phenotypically and functionally resemble T_{CM} and T_{EM} subsets respectively	82
3.6	Human Tfh_{CM} cells are enriched with the Tfh17 subset whereas Tfh_{EM} cells are enriched with the Tfh1 subset	85
3.7	AIM assays to identify antigen-specific cTfh cells	87
3.8	Early responders have detectable Tfh responses on day7	90
3.9	HBV antigen-specific Tfh17 cells are preferentially maintained in memory phase	92
3.10	The identification of influenza virus-specific cTfh cells	94
3.11	Influenza virus-specific cTfh cells show Tfh1 signatures in effector phase but Tfh17 signatures in memory phase	96

3.12	Tfh17 cells are long-lived and accumulate with aging	99
4.1	Single cell RNA-seq of human tonsils identifies TFs associated with the ABC lineage	108
4.2	Unsupervised clustering and the identification of ABCs	110
4.3	VDJ analysis of human tonsil B cells	113
4.4	Transcription factor analysis of ABC and cMBC	115
4.5	Murine ABCs induced by infection or immunization share a transcriptomic program with human ABCs	117
4.6	ABCs are not PB/PC precursors and have poorer survival and proliferation potential than cMBCs	119
4.7	The in vitro differentiation of ABC-like cells	121
4.8	An optimized Cas9-RNP based gene knock-out method in mouse B cells	123
4.9	CRISPR Cas9 screen identified Zeb2 was required for ABC formation	125
4.10	Zeb2 was required for the expression of ABC and suppression of cMBC genes	127
4.11	Zeb2 deficiency selectively impaired ABC formation during P. chabuadi infection	129
4.12	Zeb2 deficiency selectively impaired ABC formation upon PfCSP-SPZ immunization	131
4.13	ZEB2 haploinsufficient MWS patients had reduced circulating ABCs	133
5.1	Igh^{g2A10} B cells extract CSP from SPZs	140
5.2	Scanning electron microscopy to visualize the B cell-SPZ interaction	142
5.3	Unsupervised clustering of Tfh cells	144
5.4	The TCR analysis of Tfh cells	146
5.5	The generation of PfCSP-specific TCR transgenic mice (CST-II)	148
5.6	PbT-II cells were incapable to help Igh^{g2A10} B cell responses upon PfCSP-SPZ immunization	150
5.7	Anti-PbCSP humoral responses were supported by CST-II cells that were specific to the bystander surface antigen PfCSP	152
5.8	The generation and phenotyping of the transgenic SPZs	154
5.9	OT-II cells were only capable to help Igh^{g2A10} B cells when OVA was located on the SPZ surface	156

Chapter1- Introduction

1.1 The immune system

The immune system for mammals is classified into innate and adaptive immunity. The innate immunity contains the innate lymphoid cells that eradiate foreign pathogens in a non-antigen-specific manner (Netea et al. 2019). Innate immunity responds to infection in as little as a few minutes and this often results in the complete eradication for most pathogens (Netea et al. 2017), however, if the initial clearance fails, the adaptive immune system mounts a response to kill pathogens in an antigen-specific manner.

The adaptive immunity is characterized by its 'specificity', killing pathogens through humoral immunity (introduced by B cells by secreting antibodies) and cellular immunity (introduced by $CD8^+$ T cells). These processes are aided by $CD4^+$ T cells that provide "help" to the effector branches of the immune system. $CD4^+$ T cells also have effector functions of their own and can mediate the killing of pathogens via cytotoxic activity and the secretion of cytokines. The 'specific-killing' is achieved by cells using T cell receptor (TCR) and B cell receptor (BCR) expressed by recombined immunoglobulin genes allowing the specific binding to the foreign antigen epitopes. Unlike innate immunity, the establishment of adaptive immunity takes weeks, but it is crucial to protect the host from secondary infection for its ability to 'remember' and to quickly respond upon antigen restimulation (Netea et al. 2019).

1.2 $CD4^+$ T cells

TCR and VDJ recombination

The TCR is a cell-surface receptor expressed on T cells that mediates binding with the MHC-peptide complex. The TCR is composed of genetically diverse αβ chains and forms a complex with the CD3 molecule which transduces the TCR signaling through the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) (**Figure 1.1A**) (Mariuzza et al. 2020). The TCR α chain is encoded by recombined variable (V), joint (J), and constant (C) genes located on the chromosome 14 for both human and mouse. The TCR β chain is encoded by recombined variable (V), diverse (D), joint (J), and constant (C) genes located on the chromosome 7 for human and 6 for mouse (Mariuzza et al. 2020) (**Figure 1.1B**).

VDJ recombination is directed by lymphoid-specific recombinase called RAG1 and RAG2. These enzymes create double strand DNA break between the chosen pairs of segments (for

example, D2 and J1) and delete the intervening DNA, resulting in ligation of the two segments (**Figure 1.1B**). This happens in the order of D-J and then V-DJ for the β chain and V directly ligated with J in the α chain. The binding of the TCR with the MHC complex is mainly mediated by 3 complementarity-determining region (CDR) loops, 2 of which localize inside the V region and the third one called CDR3 localizes in the VDJ recombined region (VJ for alpha chain) (**Figure 1.1C**), resulting in highly diverse TCR repertoire with at least 10^7 different combinations for human TCR (Arstila et al. 1999).

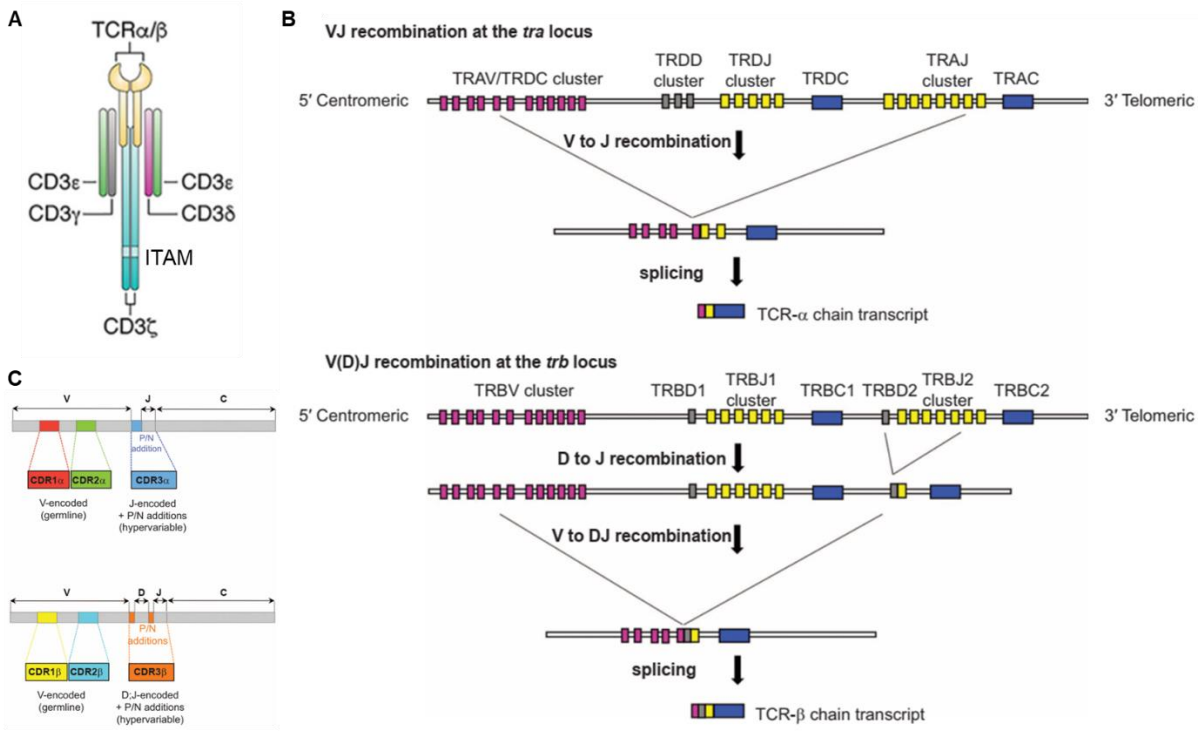


Figure 1.1 TCR and VDJ recombination
(A) Structure of TCR-CD3 complex. (B) VDJ recombination of TCR α and β chain. (C) The CDR regions in TCR α and β chain. The figures are adopted from (Attaf et al. 2015, Mariuzza et al. 2020).

The development and differentiation of CD4$^+$ T cells

CD4$^+$ T cells emerge from the lymphoid progenitor cells in the bone marrow, and complete their development in the thymus. In the thymus, the progenitor CD4 and CD8 double positive (DP) T cells use TCR to interact with cortical epithelia cells presenting self-antigen bound MHC-I and II molecules. To avoid the development of self-reactive T cells which might trigger autoimmunity, DP cells receiving too strong TCR signaling strength are eliminated and this process is called negative selection. In the meantime, to select the cells that are capable to efficiently engage with MHC molecules to mount protective immunity, the cells with too poor TCR signaling strength are also eliminated (Germain 2002, Mandl et al. 2013), while DP cells receiving the intermediate TCR signaling strength are positively selected to become CD8$^+$ or CD4$^+$ single positive naïve T cells which then emigrate to the peripheral lymphoid organs.

Upon immunization and infection, naïve CD4$^+$ T cells that are capable of recognizing foreign peptides presented on MHC-II receives activation signals from antigen presenting cells (APCs) and differentiate into effector CD4$^+$ T cells. The activation signals are classified into signal 1 which is triggered by the strong binding of TCR with an MHC-II-peptide complex, and signal 2 which is induced by the binding of co-receptor CD28 with ligands CD80 or CD86 expressed on the APCs. The concurrent signal 1 and 2 is essential for the efficient effector CD4$^+$ T cell differentiation (Bretscher 1999). The effector CD4$^+$ T cells exert a wide range of functions, ranging from coordinating the innate immune response, supporting cytotoxic CD8$^+$ T cell responses, supporting B cell responses and also inducing immune regulation (Luckheeram et al. 2012). After antigen clearance, the CD4$^+$ T cells go through an immune contraction phase during which most effector CD4$^+$ T cells undergo apoptosis but some survive into memory CD4$^+$ T cells (MacLeod et al. 2010). Memory CD4$^+$ T cells can quickly differentiate into effector cells upon antigen re-exposure with a bias to carry out their original function although this process also accompanies high degrees of plasticity (Sallusto et al. 2018).

Consistent with their multiple roles in the immune system, effector CD4$^+$ T cells are classified into regulator T (Treg), Th1, Th2, Th17, and follicular helper T (Tfh) cells with distinct phenotypes and functions (Luckheeram et al. 2012, Read et al. 2019). Treg cells (CD25^{high} FOXP3$^+$) are specialized for mediating immune suppression by expressing inhibitory molecules such as CTLA4 (Walker 2013) (**Figure 1.2A**). Th1 cells (CXCR3$^+$ T-bet$^+$) are specialized for intracellular pathogen clearance through the secretion of IFN-γ, IL-2 and IL-21 which enhance phagocytosis by macrophages and the cytotoxic effects of CD8$^+$ T cells (Hinrichs et al. 2008, Ivashkiv 2018) (**Figure 1.2B**). Th2 cells (CCR4$^+$ GATA3$^+$) are specialized to induce protective

immunity against extracellular parasites via the induction of type 2 immunity (Koyasu and Moro 2011, Luckheeram et al. 2012) (**Figure 1.2C**). Th17 cells ($CCR6^+$ $ROR\gamma t^+$) are specialized for bacteria and fungi clearance through the secretion of IL-17 family cytokines to recruit neutrophiles (Flannigan et al. 2017) (**Figure 1.2D**). Finally, Tfh cells ($CXCR5^+$ $BCL6^+$) are specialized for helping germinal center (GC) B cell formation and plasma cell (PC) differentiation by providing essential signals such as CD40L and IL-21 (Crotty 2019) (**Figure 1.2E**).

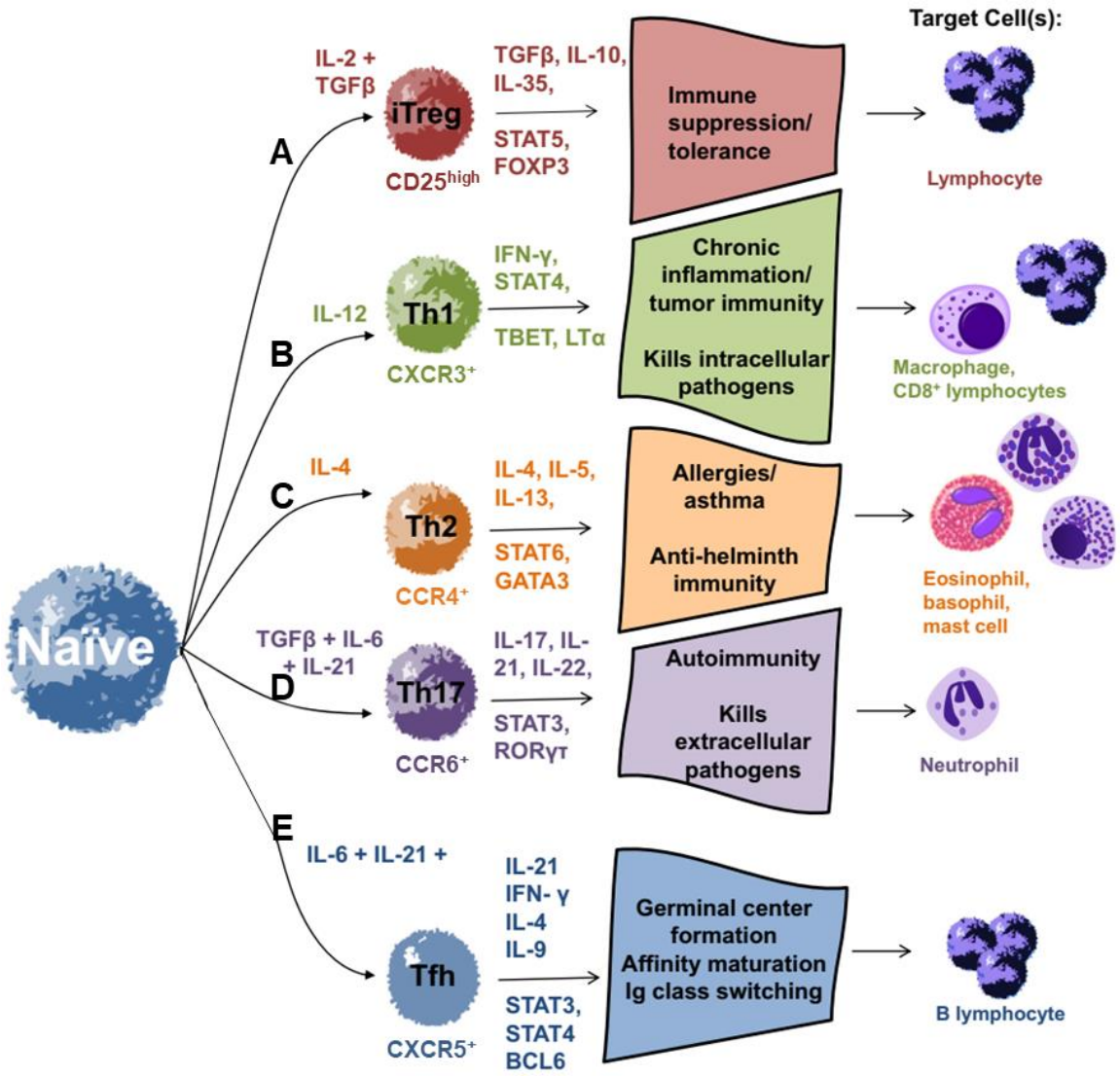


Figure 1.2 The differentiation and function of effector CD4+ T cell subsets
The cytokines that drive the differentiation of indicated effector $CD4^+$ T cell subsets, and their surface markers, master transcription factors, hallmark cytokines, and functions are summarized for Treg (A), Th1 (B), Th2 (C), Th17 (D), and Tfh (E) cells. The figure is modified based on (Knochelmann et al. 2018).

1.3 B cells

BCR and VDJ recombination

The BCR is composed of membrane bound immunoglobulin (Ig) (**Figure 1.3A**) and co-receptors CD79a and CD79b which transduce BCR signaling through the phosphorylation of immunoreceptor tyrosine-based activation motifs (ITAMs) (**Figure 1.3B**). Similar to the TCRα/β chain, the BCR is composed of heavy and light chains. The heavy chain is encoded by recombined variable (V), diverse (D), joint (J), and constant (C) genes located on the chromosome 14 for human and 12 for mouse. The light chain is encoded by recombined variable (V), joint (J), and constant (C) genes located on the chromosome 2 (Igκ) or 22 (Igλ) for human and 6 (Igκ) or 16 (Igλ) for mouse (Mourad 2020). Despite these similarities, the BCR has many major differences to the TCR. First, although the TCR only recognizes antigens in the context of MHC, the BCR can directly bind to the naïve antigens. Moreover, while TCR is only located on the cell membrane, the BCR can exist as both a membrane protein and in a secreted form as soluble antibody when the membrane bound region of the heavy chain is removed via alternative splicing (**Figure 1.3C**) (Mourad 2020).

The process of VDJ recombination in B cells is similar to that in T cells, and is directed by lymphoid-specific recombinases called RAG1 and RAG2. These enzymes create a double strand DNA break between the chosen pairs of segments and delete the intervening DNA, resulting in ligation of the two segments (**Figure 1.3D**). This happens in the order of D-J and then V-DJ for the heavy chain and V directly ligated with J in the light chain (Mourad 2020). The binding of the BCR with the antigen is mainly mediated by 3 complementarity-determining region (CDR) loops, 2 of which localize inside the V region and the third one called CDR3 localizes in the VDJ recombined region (no VJ for light chain), resulting in highly diverse BCR repertoire with a theoretical diversity of $>10^{14}$ in humans (Yaari and Kleinstein 2015). However, differently from the TCR, the germline sequences of BCR can be somatic hypermutated to further increase the diversity and affinity during the germinal center reaction.

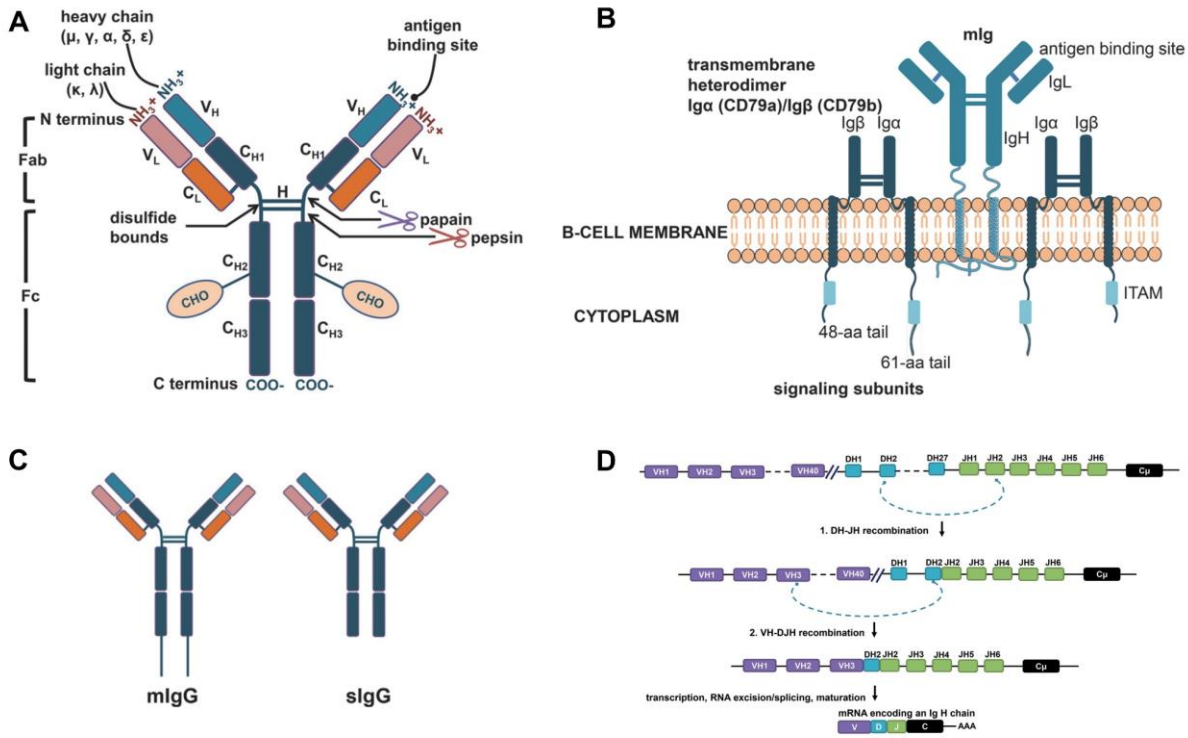


Figure 1.3 BCR and VDJ recombination

The structure of the immunoglobulin (Ig), BCR with membrane-bound and soluble form, and the process of VDJ recombination are summarized. (A) The structure of the Ig, which comprises heavy chain, light chain and constant region. (B) The structure of BCR complex, which is composed of Ig and co-receptor Igα and Igβ to trigger downstream signaling. (C) The structures of membrane-bound and the soluble form of the IgG. (D) The VDJ recombination of Ig heavy chain. The figures are adopted from (Mourad 2020).

B cell development and differentiation

B Cells are divided into B1 and B2 cells. B1 cells are innate like cells predominately originated from fetal liver and mainly reside in the peritoneal cavity instead of lymphoid tissues (Berland and Wortis 2002). B2 cells are the main drivers of most B cell driven humoral immunity and so are the main focus of this thesis.

Early B cell development occurs in the bone marrow where an ordered rearrangement of immunoglobin heavy and light chain (IgH and L) occurs (LeBien and Tedder 2008), and B cells undergoing these processes are termed Pro and Pre B cells respectively (Pieper et al. 2013). Upon completion of V(D)J rearrangement, Pre B cells become IgM^+ immature B cells and migrate out from the bone marrow into the secondary lymphoid organs such as spleen and lymph node where they develop into IgM^+IgD^+ mature B cells (or naïve B cells) including follicular and marginal zone B cells (Pieper et al. 2013) (**Figure 1.4A**). In line with the process of B cell development, the BCR receptor develops from a pre-BCR (composed of IgH+surrogate IgL) to a mature BCR (composed of IgH+IgL) and the BCR localization transits from intracellular to surface expression (Rickert 2013). The classification of different stages of early B cell development was based on flow cytometry analysis in the 1990s (Hardy et al. 1991) and this was largely confirmed by a recent study using single cell RNA-seq (scRNA-seq) analysis (Lee et al. 2021).

Naïve B cell activation is normally initiated when the BCR is bound by cognate antigen. The BCR signaling strength is determined by the affinity of BCR to the antigen but can be further modulated by many other co-receptors signals such as Toll-like receptors (TLRs), Fc receptors, CD22 and CD72 (Cyster and Allen 2019). Activated B cells upregulate the expression of antigen presentation genes like MHC-II and CD86, cell cycle genes, and change their expression of chemokine receptors including CCR7 and CXCR5 to favor their migration from the follicular area into the T-B border where they expect to interact with antigen specific $CD4^+$ T cells within 12-24h after antigen recognition otherwise they switch back to a naïve-like state (Turner et al. 2017, Cyster and Allen 2019). To be noted, B cell activation and differentiation can sometime bypass the requirement of T cell help and such process is named T-independent response. T-independent responses can be triggered by overwhelmingly strong BCR ligation or TLRs signals but such responses are transient and in most cases are not considered to be the major source of neutralizing antibody (Cyster and Allen 2019). In contrast, activated B cells during a typical T-dependent response differentiate either into GC B cells, antibody secreting cells (especially long-lived) or memory B cells. Many factors such as BCR signaling strength

and Tfh help play a role during the fate determination but this area is still not fully understood (Krautler et al. 2017). GC B cells – as their name suggests - localize in the GC, a structure in the lymphoid organs that displays high rate of cell division according to histological analysis (Klein and Dalla-Favera 2008). GC B cells are featured by performing somatic hypermutation (SHM) to improve their BCR affinity, intensive cell cycle history, and capability to differentiate into long-lived antibody secreting cells and memory B cells (**Figure 1.4B**). The germinal center is a major focus of this thesis and is discussed in more detail below.

Antibody secreting cells are terminally differentiated effector cells composed of plasmablasts (PB) and plasma cells (PC) specialized for the secretion of neutralizing antibody (**Figure 1.4C**). Although they both secret large amount of antibody, PB are proliferating and short-lived cells localizing in the extrafollicular area while PC are more long-lived and localize mainly in the bone marrow (Nutt et al. 2015). Whether all PC pass through a PB-like stage or they can originate directly from GC B cells remains an open question in the field (Nutt et al. 2015). Nevertheless, the development of antibody secreting cells requires the expression of many TFs such as Blimp-1, IRF4 and endoplasmic reticulum (ER) stress regulator Xbp1 to adapt to the secretion of enormous amounts of immunoglobulin.

Memory B cells are specialized quiescent cells favoring the quick differentiation into effector B cells upon antigen reencounter (**Figure 1.4D**). Memory B cells are conventionally deemed to form PC after antigen recall, but recent studies suggested memory B cells are highly heterogenous and it has been proposed that different subsets have different abilities to form GC B or PC based on the expression of IgM, CD73, CD80 and PD-L2 (Harms Pritchard and Pepper 2018). However different classifications have been proposed by different groups, moreover while single cell RNA-seq also uncovers heterogeneity among memory B cells in humans, the identified clusters of activated, classical memory and atypical memory B cells based on transcriptomic profiles do not obviously correlate with the populations identified in mice using different flow cytometry markers (King et al. 2021, Sutton et al. 2021).

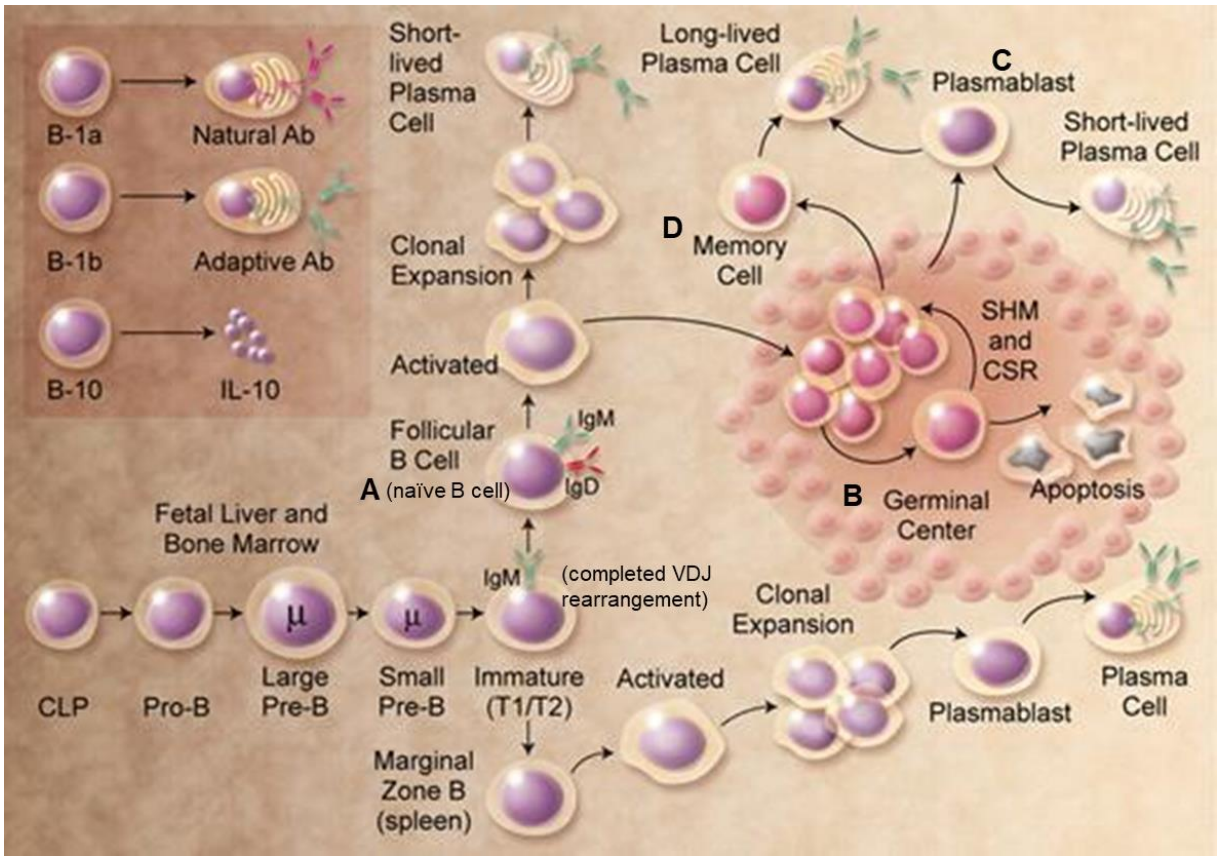


Figure 1.4 The development of naïve B cells and the differentiation of effector B cell subsets

The picture summarizes the development and differentiation of B cells. (A) B cells are originated from the fetal liver and bone marrow and developed into naïve B cells in the secondary lymphoid organs. (B) Upon antigen encounter, naïve B cells are activated, expanded, and differentiate into plasmablast or become GC B cells to perform SHM and class switch recombination. (C) Long-lived plasma cells (D) and memory B cells can be derived from GC B cells. The figure is modified based on (LeBien and Tedder 2008).

1.4 Tfh cells and the germinal center reaction

Germinal center (GC) B cells

The GC structure was first reported in the 19th century when microscopists observed the highly mitotic area in the lymphoid tissues (Cyster and Allen 2019), which was further divided into the dark zone and light zone according to the histological analysis (Klein and Dalla-Favera 2008). The dark zone mainly comprises densely packed cycling GC B cells (centroblasts) while the majority of light zone contains smaller non-dividing GC B cells (centrocytes) (Klein and Dalla-Favera 2008). The main function of the GC is to "train" the GC B cells to acquire higher affinity BCRs through somatic hypermutation (SHM) and selection, a process called affinity maturation. Affinity maturation is a Darwinian-like process in which centroblasts first undergo intensive proliferation and SHM to produce a large pool BCRs with variable affinities to the given antigen, followed by a selection process that occurs in the light zone where centrocytes with higher affinity preferentially become antibody secreting cells, whereas those with lower affinity tend to either become memory B cells, undergo apoptosis or traffic back into the dark zone to repeat these processes (Cyster and Allen 2019, Victora and Nussenzweig 2022). As a result, the average BCR affinity for the given antigen increases within the population of GC B cells over time.

The formation of GC B cells starts from naïve B cells receiving activation signals from the antigen-bound BCR. The activated B cells then migrate to the T-B border to interact with cognate CD4$^+$ T cells and some of them become precursors of GC B cells, moving back to the B cell follicle where they undergo 3-4 days of intensive proliferation to establish the GC (Klein and Dalla-Favera 2008, De Silva and Klein 2015). The transcription factor BCL6 plays a central role in regulating GC B cell formation. First, BCL6 suppresses DNA damage mediated cell cycle arrest, permitting the GC B cells to survive the genotoxic stress caused by SHM and intensive proliferation (Klein and Dalla-Favera 2008). Second, BCL6 can maintain the GC B cell identity by directly suppressing the expression of other B cell lineage defining TFs such as BLIMP1 to inhibit the differentiation of GC B cells into PB/PC (Shaffer et al. 2000). In addition to TFs, co-stimulation signals and cytokines from CD4$^+$ T cells are also important in regulating the GC B cell formation.

SHM provides the molecular basis of affinity maturation of the GC B cells. SHM is driven by activation-induced cytidine deaminase (AID), an enzyme that preferentially targets the immunoglobulin transcriptionally active region (Pilzecker and Jacobs 2019). AID replaces the

nucleotide cytosine with uracil, resulting in a nucleotide mismatch of uracil to guanine. The error-prone canonical and noncanonical DNA repair mechanisms are then triggered and mutagenesis is introduced (Maul and Gearhart 2010). It is estimated that the mutagenesis rate of AID mediated SHM is 10^{-3} base pair per generation, one million times higher than the non-AID targeted genome area (Pilzecker and Jacobs 2019), permitting the rapid diversification of the BCRs among GC B cells.

Finally, GC B cells expressing BCR with different affinities to the antigen will either form PC/memory B cells, or re-cycling to dark zone, or undergoing apoptosis. The specific mechanism that controls such fate determination is not fully understood, but it is agreed that an array of signals such as BCR, cytokines and co-receptors cooperate to orchestrate the GC B cell fate (Zotos and Tarlinton 2012). Among these factors, the signals derived from Tfh cells are one of the major determinants of GC B cell fate.

Follicular helper T (Tfh) cells

Before the discovery of Tfh cells, Th2 cells were considered to be the helper cell type to induce the effective antibody responses (Spellberg and Edwards Jr 2001). Tfh cells were first reported in 2000 when two groups discovered a unique $CXCR5^+$ $CD4^+$ T cell population localized in B cell follicle and could support B cell antibody production *in vitro* (Breitfeld et al. 2000, Schaerli et al. 2000). However it was not until 2009 when BCL6 was identified as the master transcription factor (TF) that Tfh cells become widely accepted as a distinct effector $CD4^+$ T cell lineage (Schaerli et al. 2000, Johnston et al. 2009, Nurieva et al. 2009, Yu et al. 2009, Crotty 2019). Tfh cells are specialized for supporting T-dependent B cell responses. In Bcl6 knockout mice that lack Tfh cells, immunization and infection cannot induce GC B cell formation, resulting in significantly impaired humoral responses (Johnston et al. 2009, Nurieva et al. 2009, Yu et al. 2009). To date, Tfh cells have been shown to play vital roles in a large variety of immune contexts such as infectious diseases, vaccines, autoimmunity and cancer (Crotty 2019).

The development of Tfh cells starts from naïve $CD4^+$ T cells being activated by professional APCs such as dendritic cells (DCs). The activated $CD4^+$ T cells downregulate CCR7 and upregulate CXCR5 expression to migrate to the T-B border where they receive further signals like TCR stimulation and ICOS-ICOSL interactions from activated B cells. Such T-B interaction drives the activated T cells to become $CXCR5^{int}$ precursor Tfh cells and migrate into

germinal center (GC) area in the B cell follicle where they have persistent interactions with GC B cells and differentiate into CXCR5^{hi} mature Tfh cells (Baumjohann et al. 2013). When losing the stable interaction with GC B cells, such as the knocking out SH2-domain adaptor protein (SAP) in CD4^{+} T cells, the formation of Tfh cells was abolished (Crotty 2014). In additional to the persistent interaction with cognate B cells, cytokines and TFs also play a crucial role in Tfh fate commitment. For example, IL-6 and IL-21 mediated STAT3 phosphorylation are required for Tfh differentiation while IL-2 mediated STAT5 phosphorylation strongly suppresses Tfh formation (Nurieva et al. 2008, Eto et al. 2011, Johnston et al. 2012). As for TFs, BCL6 and ASCL2 are essential for Tfh differentiation whereas BLIMP1 is an antagonist to BCL6 (Schaerli et al. 2000, Johnston et al. 2009, Nurieva et al. 2009, Yu et al. 2009, Johnston et al. 2012, Liu et al. 2014, Crotty 2019).

Since Tfh cell localize in the B cell follicle in lymphoid organs, studying human Tfh cells is obstructed by the difficulty of obtaining human lymphoid tissues. However, recent studies suggest a subset of CXCR5^{+} CD4^{+} T cells in the blood called circulating Tfh (cTfh) cells can be used as an alternative to study Tfh cells in humans. cTfh cells are early memory-like cells that originate from draining lymph nodes and can quickly differentiate into mature Tfh cells to support B cell responses upon antigen recall response (He et al. 2013). cTfh cells comprise PD-1- central memory (CM), PD-1^{+} ICOS^{-} effector memory (EM) and PD-1^{hi} ICOS^{+} activated populations with different polarizations towards bona fide Tfh cells (Herati et al. 2017). After immunizations like influenza vaccination, the percentages of CM and EM cells among cTfh cells transiently increases and the magnitude of such positively correlates with PC formation in the blood and the production of antibody (Spensieri et al. 2013, Herati et al. 2017). Consistent with this, TCR sequencing studies demonstrate clonal similarity between cTfh cells and GC residing mature Tfh cells from the tonsil, suggesting cTfh cells share the same origin as Tfh cells (Heit et al. 2017). However, the development of cTfh cells does not completely coincide with Tfh cells. For example, SAP knock-out mice with impaired T-B interactions have defects in Tfh but not cTfh cell generation (He et al. 2013).

In sum, Tfh cells are a specialized effector CD4^{+} T cell lineage for supporting GC formation and antibody secretion. cTfh cells are the memory counterpart of Tfh cells provide accelerated B cell helper function upon antigen recall response. The phenotype, origin, and function of Tfh and cTfh cells are summarized below (**Figure 1.5**)

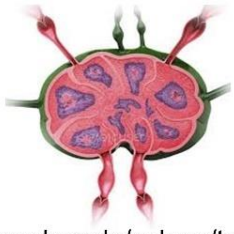
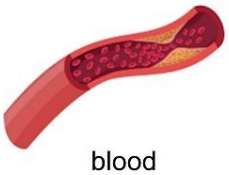
	Phenotype	Development	Function
Mature, terminal differentiated Tfh  Lymph node/spleen/tonsil	$BCL6^{+}$ $PD\text{-}1^{+}$ $CXCR5^{+}$	BCL6, ICOS and SAP dependent	Effector cell, located in germinal center and actively helping B cell response
Precursor, circulating memory Tfh (cTfh)  blood	$BCL6^{-}$ $CXCR5^{+}$ $CD45RA^{-}$ (human) $CD44^{+}$ (mouse)	BCL6 and ICOS dependent, SAP independent	Memory cell, circulating in the blood, superior than $CXCR5^{-}$ memory to support B cell response upon antigen re-stimulation

Figure 1.5 Summary of the phenotype, development, and function of Tfh and cTfh cells
The expression of phenotypic markers, the requirement of critical molecules during formation and the specialized functions are summarized for terminal differentiated Tfh and cTfh cells.

The interaction of Tfh cells and GC B cells

The interaction of Tfh cells and GC B cells has two major functional outcomes: the development/maintenance of the Tfh and GC B cells, and the selection of GC B cells.

The formation and the maintenance of Tfh cells and GC B cells are concurrent processes. On one hand, GC B cells receive complex signals from Tfh cells to survive and to proliferate. These signals include cytokines like IL-21, IL-4 and IFNγ instructing GC B cell proliferation, isotype switch and PC differentiation (Weinstein et al. 2012). Additionally, direct cell contact signals like CD40-CD40L, ICOS-ICOSL also are indispensable for B cell survival and proliferation, as antibody blockade of these interactions significantly reduces the formation of GC B cells (Weinstein et al. 2012, Baumjohann et al. 2013). On the other hand, GC B cells provide essential signals to sustain the identity of Tfh cells. One of the major signals is the presentation of cognate antigen by GC B cells resulting in the activation of the TCR, as the depletion of GC B cells by DTR system during an established GC reaction upon LCMV infection significantly reduced the number of Tfh cells, moreover a boost by MHC-II binding peptide after GC B cell depletion could rescue the loss of Tfh cells (Baumjohann et al. 2013).

Tfh cells play a critical role in the selection and fate determination of GC B cells. Enhanced interaction with Tfh cells during an ongoing GC reaction will temporarily restrict more GC B cells in the light zone for around 12h, and permit their stronger expansion in the dark zone at later timepoints, resulting in more PC differentiation and higher antibody production (Victora et al. 2010). In support of this, knockout of IL-21, which is predominately secreted by Tfh cells during the GC reaction, significantly reduces the formation of PC but accelerates the generation of memory B cells (Zotos et al. 2010). It is clear that Tfh cells affects the selection and the fate determination of GC B cells, however, since Tfh cells transduce a complex array of signals to GC B cells, and the phenotype of Tfh cells evolves over time during an GC response (Jacobsen et al. 2021), it still remains unclear what is the underlying mechanism by which GC B cells are selected and differentiated through the cognate interaction with Tfh cells (De Silva and Klein 2015).

1.5 Memory B cell responses: the role of atypical memory B cells (ABCs)

The identification of ABCs

Memory B cells form after infection or vaccination in order to provide a rapid amnestic response to reinfection. These cells in humans are characterized by the high expression of CD27 and CD21(Weisel and Shlomchik 2017). As such the markers and function of these so-called classical memory B cells (cMBCs) are well understood. However, a now seminal study in the tonsil identified a population of B cells expressing FCRL4 but not CD27. This novel population was classified as a memory cell because of negligible BCL6 and BLIMP1 expression, the hall mark TFs of GCB and plasma cells PC respectively (Ehrhardt et al. 2005). Because of their identification in the tonsils these cells were initially designated as tissue-like memory B cells, however their function was unknown. The subsequent identification of similar populations mostly in conditions of autoimmunity, chronic infection and aging but without functional characterization has led to these cells being grouped under the umbrella of atypical B cells (ABCs) which is the terminology we will use here.

The early analyses of ABC in the tonsils reveals that these cells appeared phenotypically activated being of larger size and having decreased complement receptor CD21 expression (Ehrhardt et al. 2005). Additionally, ABCs have a different integrins and chemokine receptors expression profile, including the expression of CD11c and loss of CXCR5, suggesting they had a distinct migratory pattern and tissue localization (Ehrhardt et al. 2005, Ehrhardt et al. 2008). Finally, while they were classified as memory B cells based on the lack of BLIMP1 and BCL6 expression, they also had a distinct transcription factor profile compared to cMBCs as they upregulated RUNX1, RUNX2 and SOX5 expression, suggesting they likely maintain a distinct global transcriptomic profile instead of merely under a transient stage (Ehrhardt et al. 2008).

Subsequently, ABCs were also reported in human blood (in healthy individuals and patients with infections or autoimmunity) and other tissues including spleen, lymph nodes, bone marrow, kidney, liver, synovial tissues, cerebrospinal fluid and choalveolar lavage fluid (Moir et al. 2008, Moir et al. 2008, Rakhmanov et al. 2009, Weiss et al. 2009, Isnardi et al. 2010, Charles et al. 2011, Portugal et al. 2015, Knox et al. 2017, Obeng-Adjei et al. 2017, Jenks et al. 2018, Arazi et al. 2019, Rivera-Correa et al. 2019, Sundling et al. 2019, Zhang et al. 2019, Horns et al. 2020, Johnson et al. 2020, Nehar-Belaid et al. 2020, Ramesh et al. 2020, Turner et al. 2020, Zhao et al. 2020, Holla et al. 2021, Keller et al. 2021, King et al. 2021, Stewart et al. 2021, Sutton et al. 2021, Zhang et al. 2021) **(Figure 1.6)**. Interestingly, ABCs preferentially located in the blood, spleen and bone marrow, with fewer cells detected in lymphoid systems such as lymph nodes, thoracic duct fluid and tonsil (Johnson et al. 2020, Keller et al. 2021). More recently ABC like populations have also been identified in mice facilitating experimental studies. Initially a

population of $CD11c^+CD11b^+$ or $CD21^-CD23^-$ cells was identified in older mice and classed as age-associated B cells (Hao et al. 2011, Rubtsov et al. 2011). Subsequent studies in ageing, infection and autoimmunity models had shown that $CD11c^+CD11b^+$ B cells were phenotypically and transcriptomically similar to human ABCs such as they both upregulated the expression of *Cd72, Hck, Tbx21, Zeb2* and *Zbtb32* (Hao et al. 2011, Rubtsov et al. 2011, Rubtsova et al. 2013, Jackson et al. 2014, Perez-Mazliah et al. 2018, Du et al. 2019, Kim et al. 2019, Zhang et al. 2019, Levack et al. 2020). Consistent with the preferential tissue localization of human ABCs, mouse hemagglutinin-specific $CD11c^+$T-bet^+ ABCs preferentially maintained in the spleen but not in the lymph nodes after influenza infection, but they could not migrate from the spleen of primary infected mice into the biological conjoined naïve mice (Johnson et al. 2020). Using Bcl6 deficient mice a recent study suggested that ABCs form in the extrafollicular compartment and migrate to the marginal zones of the spleen potentially explaining their preference for this compartment (Song et al. 2022).

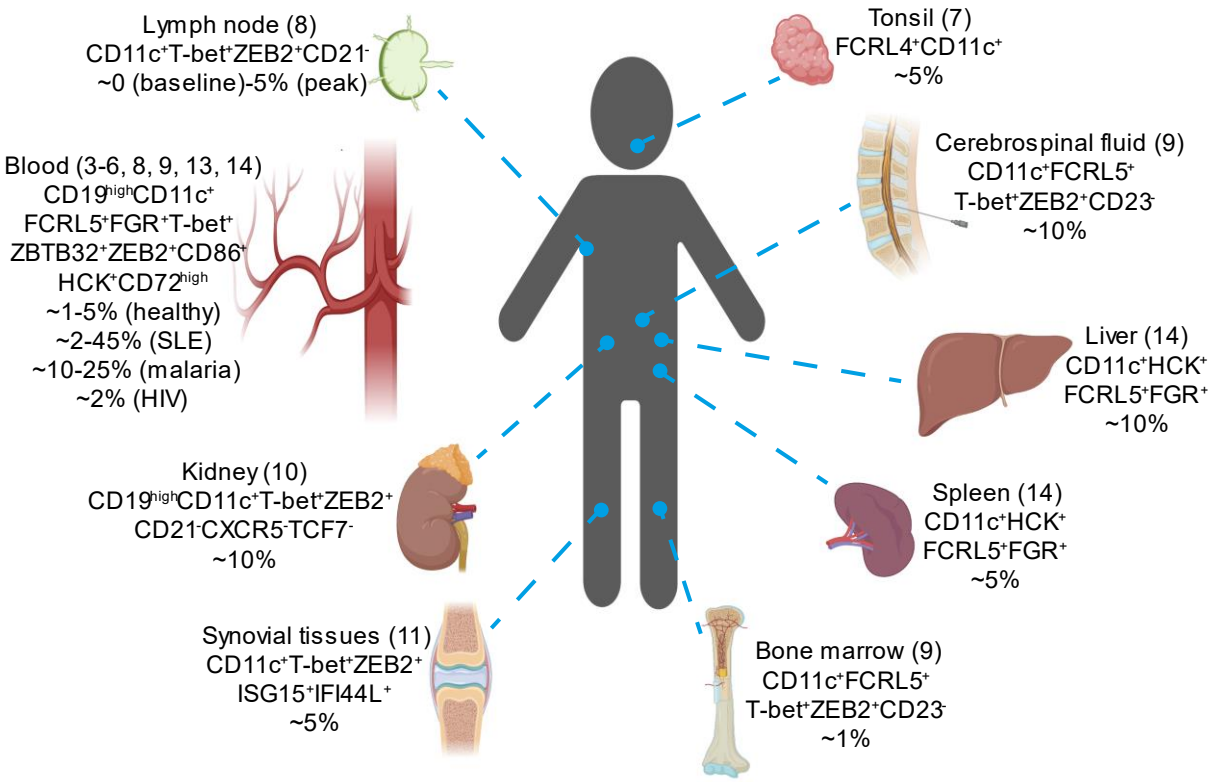


Figure 1.6 scRNA-seq revealed the conserved presence of ABCs in human
Single-cell RNA-seq studies have identified ABCs in numerous tissues by unsupervised clustering. Signature genes and estimated percentages in total B cells for ABCs cluster are shown.

Historically ABCs have been identified using flow cytometry based on a handful of key markers. $CD27^-CD21^-$ has been frequently used to identify human ABCs during chronic or repeated infection such as with malaria (Weiss et al. 2009). However, a more recent study has demonstrated that $CD21^-CD27^-$ circulating B cells could be further separated into $FCRL5^-$ and $FCRL5^+$ subsets, and only the $FCRL5^+$ subsets expressed ABCs associated signatures whereas the $FCRL5^-$ subset resembled classical memory/naive cells (Li et al. 2016). Similar concerns also raised in mouse studies. For example, $CD21^-CD23^-$ B cells were identified as ABCs in aging mice whereas these cells have been showed in surrogate light chain deficient lupus mouse model unlikely to be ABCs because of negligible CD11c expression (Aranburu et al. 2018), which is used as a canonical ABC marker in most ageing, infection and autoimmunity models (Rubtsov et al. 2011, Rubtsova et al. 2013, Jackson et al. 2014, Perez-Mazliah et al. 2018, Du et al. 2019, Kim et al. 2019, Zhang et al. 2019, Levack et al. 2020). Thus, the sensitivity and specificity for any gating strategy designed to identify ABCs is unclear. Moreover, a gating strategy that efficiently identifies ABC-like cells in one context may not identify transcriptionally similar cells in another setting.

Because single cell RNA-seq relies on the expression of thousands of genes to classify and cluster cell type, it should circumvent the problems of identification by flow cytometry which is limited to at most 10s of markers. Thus, single cell RNA seq has the potential to aid the identification of more universal markers of the ABC population and determine if cells identified in different contexts are in fact phenotypically similar to each other. Notably, nearly all datasets in a variety of tissues have detected ABCs by unsupervised clustering strategy frequently marked by genes such as $FCRL4^+$ $FCRL5^+$ $CD11c^+$ $T\text{-}bet^+$ $ZEB2^+$ and FGR^+ (Arazi et al. 2019, Zhang et al. 2019, Horns et al. 2020, Nehar-Belaid et al. 2020, Ramesh et al. 2020, Turner et al. 2020, Holla et al. 2021, King et al. 2021, Stewart et al. 2021, Sutton et al. 2021, Zhang et al. 2021). Similar observations had been reported by comparing ABCs in different disease contexts. By integrating the scRNA-seq datasets obtained side-by-side from healthy individuals, malaria and HIV infected patients, one recent study showed ABCs from these donors co-localized into one cluster suggesting they had overlapped transcriptomes (Holla et al. 2021). Further analysis comparing ABCs' signature genes derived from different datasets demonstrated the high similarity between circulating ABCs induced in infections and autoimmune diseases such as systemic lupus erythematosus (SLE), rheumatoid arthritis and common variable immunodeficiency (Holla et al. 2021). Interestingly, a recent scRNA-seq dataset on horse blood also reported a ABCs cluster marked by $FCRL4^+T\text{-}bet^+CD11c^+$ (Patel et al. 2021). These results

suggested that ABCs identified in different tissues, organisms and in healthy individuals and patients with different diseases, were highly conserved and shared a similar global transcriptomic profile.

Given that ABCs were highly conserved in different contexts, an outstanding question is the identity of the best markers to identify ABCs. The previous work from my lab attempted to address this using cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) (Stoeckius et al. 2017). In this approach antibodies conjugated to oligonucleotide sequences (barcodes) can be used to label cells in a single cell experiment and the relative levels of barcode sequence used to measure the level of surface expression. Using barcoded antibodies against CD21, CD27, CXCR3 and CD11c to phenotype ABCs in healthy individuals and in malaria infected patients, it was reported that although ABCs in malaria infected patients were predominantly $CD21^{-}CD27^{-}$, ABCs in healthy individuals appeared to express substantial levels of CD21 and CD27, suggesting these markers were not optimal for identifying general ABCs lineage. In contrast, CD11c were upregulated on ABCs clusters from both subjects, suggesting $CD11c^{+}$ was a more conserved defining feature for the general ABCs population. Nonetheless the variable expression of all ABC markers probably leads to an undercount of these cells in healthy individuals. Analysis of a more extensive panel of antibodies should in theory yield markers with even better resolution. However, the finding that CD11c is perhaps the most useful marker for ABCs is consistent with a recent cytometry by time of flight (CyTOF) study comparing 351 surface molecules on human circulating B cells and reported CD11c serves as the best marker to define $CD11c^{+}CD19^{high}$ ABCs cluster in healthy individuals (Glass et al. 2020). Overall, published scRNA-seq studies showed CD11c, FCRL4 and FCRL5 were the most frequently up-regulated surface genes for ABCs cluster, accompanied by TFs T-bet and ZEB2 (Arazi et al. 2019, Zhang et al. 2019, Horns et al. 2020, Nehar-Belaid et al. 2020, Ramesh et al. 2020, Turner et al. 2020, Zhao et al. 2020, Holla et al. 2021, King et al. 2021, Stewart et al. 2021, Sutton et al. 2021, Zhang et al. 2021). The utility of these markers in mice is unclear, CD11c is commonly used but mouse Fcrl5 is not a direct homolog of the human protein of the same name and appears to be a marker of all murine memory B cells (Perez-Mazliah et al. 2018, Kim et al. 2019).

Despite the highly conserved feature of ABCs among different immunological contexts, previous scRNA-seq analysis from my lab suggested that ABCs were heterogenous and had different degrees of polarization towards the atypical phenotype. In CITE-seq analysis $CD21^{-}$ $CD27^{-}$ ABCs which dominated in the malaria infected patients expressed highest levels of

ABCs-associated markers such as CD11c and T-bet, whereas ABCs that in healthy individuals were predominately $CD21^{+}CD27^{+}CD11c^{int}$ (Sutton et al. 2021). Longitudinal analysis of B cells specific for the malaria circumsporozoite protein in previously malaria-naïve individuals receiving a whole parasite vaccine revealed that $CD11c^{+}$ ABCs were induced in primary immunization but with repeated immunizations these cells cell increasingly adopted a $CD21^{-}$ $CD27^{-}$ phenotype (Sutton et al. 2021). Additionally, FCRL4 and FCRL5 have distinct expression patterns on ABCs under different immunological contexts. Although FCRL4 was highly expressed on circulating $CD21^{-}CD27^{-}$ ABCs in HIV infected patients, it was not expressed by circulating $CD21^{-}CD27^{-}$ ABCs in malaria infected patients nor $CD11c^{+}CXCR5^{-}$ ABCs in SLE patients (Moir et al. 2008, Portugal et al. 2015, Jenks et al. 2018).

The Development of ABCs

While single cell technologies have shown that ABC are a distinct lineage found not only in disease settings but also in healthy individuals and protective immune responses, these techniques can also help to identify the development and origin of ABCs at the cellular level, the extracellular stimulus level and the TF level.

A key controversy has been whether ABCs have a GC dependent or independent origin. BCR sequence analysis in multiple settings has shown that ABCs have comparable or slightly lower levels of somatic hypermutation (SHM) with cMBCs and such values were much higher than naïve or follicular B cells in human and mouse (Muellenbeck et al. 2013, Russell Knode et al. 2017, Austin et al. 2019, Du et al. 2019, Johnson et al. 2020, Holla et al. 2021, King et al. 2021, Sutton et al. 2021), suggesting ABCs are GC-experienced cells. Similar observations also reported on influenza-specific ABCs which acquired SHM but forming distinct clades within phylogenetic trees compared to cMBCs and plasmablast (Lau et al. 2017). Consistent with this, B cell specific knock-out of *Bcl6* abolished the development of $CD11c^+$ $T\text{-}bet^+$ ABCs in mice upon *E. muris* infection (Levack et al. 2020). However, other evidence suggests a GC-independent origin for ABCs. IgD^+ cells with the features of ABCs have been reported in SLE patients and in aged female mice (Hao et al. 2011, Rubtsov et al. 2011, Jenks et al. 2018), while *Bcl6* in B cells was not required for $CD11c^+$ $T\text{-}bet^+$ ABCs formation in Lymphocytic Choriomeningitis Virus (LCMV) and Influenza virus infection in mice (Song et al. 2022).

Regardless of the GC or non-GC origin of ABCs their formation appears to be T dependent. Several reports have showed inhibiting T-B interaction by knocking-out MHC-II, CD40 or CD40L completely abolished ABCs formation in ageing, lupus and *E.muris* infected mice (Russell Knode et al. 2017, Du et al. 2019, Levack et al. 2020). Whether follicular helper T (Tfh) cells are required for ABCs formation remains controversial since contradictory results reported in T cells specific *Bcl6* deletion LCMV and *E.muris* infection models (Levack et al. 2020, Song et al. 2022). Nevertheless, these studies suggested that ABC formation occurs at the T-B border, rather than in GC structures but this may permit somatic hypermutation to occur, explaining the BCR mutations present in ABCs. Of note, ABCs could also be GC independent, but carry high levels of SHM if they arose from GC-experienced cMBCs. This idea was supported by the observation that secondary vaccination or infection can induce stronger $CD11c^+$ ABC production than the primary response (Sundling et al. 2019, Sutton et al. 2021). Nevertheless, perhaps the best theory to generalize these observations is that the origin of ABC is not restricted to a specific type of B cell. Naïve, GC B cells and cMBC might all be the source

of ABC as long as they receive the right extracellular stimulus. This idea is supported by the observation that ABCs in *Plasmodium*-exposed Malian children could be separated into $IgD^{-}IgG^{+}$, $IgD^{+}IgM^{+}$ and $IgD^{+}IgM^{low}$ subsets with SHM rates equivalent, respectively, to cMBCs (suggesting GC and cMBC origin), naïve B cells (suggesting a naïve B cell origin) and intermediate between naïve and cMBC (suggesting T-B border origin) (Holla et al. 2021).

Extracellular factors may also play a key role in the formation of ABC. The previous study from my lab looking at ABCs in individuals with high and low levels of malaria exposure found that not only were *Plasmodium*-specific ABCs more common in people from high transmission areas, but bystander B cells specific for tetanus toxin were also more likely to be ABCs (Aye et al. 2020). Thus, inflammatory mediators may drive ABC formation. According in two pioneer studies, ABCs were highly responsive to TLR7 and TLR9 ligands stimulation *in vitro*, and chronic TLR7 but not TLR9 ligand treatment can induce the formation of $CD11c^{+}CD11b^{+}$ ABCs *in vivo* (Rubtsov et al. 2011). Consistently, the B cell specific deletion of *Tlr7* or *Myd88* abolished the development of ABCs in aged female mice (Rubtsov et al. 2011), and malaria induced ABC-generation was slightly impaired when *Tlr9* was specifically knocked-out in B cells (Rivera-Correa et al. 2017).

In addition to TLR signaling, cytokines also play a central role in regulating ABC formation. IL-21 was found to be crucial for ABC development in the SWAP-70 and DEF6 double knock-out (DKO) model of lupus as well as the *E.muris* infection model (Manni et al. 2018, Levack et al. 2020). IL-21 transgenic mice also have spontaneous ABCs formation (Naradikian et al. 2016), while IL-21 combined with TLR7 or TLR9 ligands, or with anti-CD40 plus anti-BCR antibodies, induced $CD11c^{+}T\text{-}bet^{+}$ cells differentiation *in vitro* (Naradikian et al. 2016, Jenks et al. 2018, Manni et al. 2018, Wang et al. 2018). In addition to IL-21, IFN-γ has been shown to promote the generation of $CD11c^{+}T\text{-}bet^{+}$ ABCs upon Influenza and *E.muris* infection as well as in a WAS chimera lupus model (Naradikian et al. 2016, Du et al. 2019, Levack et al. 2020). However, IFN-γ does not induce CD11c expression in human and mouse B cells by *in vitro* culture (Naradikian et al. 2016, Ambegaonkar et al. 2019). IL-4 has been showed to suppress the expression of CD11c and T-bet by *in vitro* mouse B cell culture and upon *H. polygyrus* infection *in vivo* (Naradikian et al. 2016, Jenks et al. 2018).

T-bet is often considered a key transcription factor for ABC formation. Although T-bet was one of the highest expressed TFs in ABCs, and T-bet overexpression could promote CD11c expression (Rubtsova et al. 2013), several studies have showed T-bet knock-out on B cells did

not affect the generation of CD19^{high}CD11c^{+} ABCs upon *E. muris* infection or in WAS chimera lupus mice (Du et al. 2019, Levack et al. 2020). RUNX1 and RUNX2 overexpression also does not promote FCRL4 expression although both genes are highly expressed by human ABCs (Ehrhardt et al. 2008). Irf5 has also been reported to be required for ABC formation in the SWAP-70 and DEF6 double knock-out lupus model (Manni et al. 2018). However, to date, the master regulatory TFs for ABCs' development have not been identified. Published transcriptomic datasets had reported several candidates such as ZBTB32, SOX5, BHLHE40, TOX and ZEB2 that were upregulated in human and mouse ABCs (Portugal et al. 2015, Perez-Mazliah et al. 2018), especially Zeb2 which has been proposed to control ABCs formation in lupus (Gu et al. 2021).

The Function of ABCs

Perhaps the major outstanding question in ABC biology is the role these cells play in the immune system. scRNA-seq datasets suggest that ABCs sit with the memory/naïve mega-cluster but are distinct from cMBCs, activated, and pre-GC memory B cells (King et al. 2021), Consistent with this, trajectory analysis of circulating B cells shows that ABC form a stand-alone developmental branch sharing a common progenitor with cMBCs (Holla et al. 2021). These observations suggest ABCs might exert a different function from these subsets. Previous studies have reported many functional observations on ABCs which have posited three possible roles for ABCs,1) ABCs are exhausted cells 2) ABCs are pre-plasma cells and 3) APCs are specialized for antigen presentation.

Having been initially identified in a variety of pathological conditions ABCs were assumed to be an exhausted or non-functional population. In particular early experiments showed that ABCs were unresponsive to BCR/CD40 stimulation and *in vitro* analysis reported ABCs had similar or slightly weaker capacity to differentiate into PC compared with cMBCs (Isnardi et al. 2010, Charles et al. 2011, Hao et al. 2011, Rubtsov et al. 2011, Li et al. 2016, Jenks et al. 2018, Wang et al. 2018). Using probes to track CSP-specific CD11c^{+} ABCs after vaccination, previous work from my lab found that B cells peaked around 2 weeks after malaria vaccination (~30%) in human, but this number dropped to baseline (~10%) after 6 weeks, suggesting that ABCs were not as capable as cMBCs in surviving through the memory contraction phase (Sutton et al. 2021). Consistent with this, CD19^{high} human ABCs were more prone to FasL induced apoptosis than cMBCs (Austin et al. 2019). Additionally, the numbers of circulating

ABCs dropped significantly after resolution of malaria infection in human and mice or after ART treatment in HIV patients (Moir et al. 2008, Perez-Mazliah et al. 2018, Kim et al. 2019, Sundling et al. 2019). This evidence supported the notion that ABCs were "exhausted-like" cells which may act as a non-functional memory lineage to antagonize strong humoral responses during chronic infections and autoimmunity. A variant of the "exhausted cell" hypothesis is that ABCs are formed in conditions of stress and represent a better-than-nothing emergency response. In support of this a recent study showed that B cell responses to a COVID vaccine in individuals given immunosuppressive drugs were predominantly of an ABC phenotype (Thompson et al.).

A second hypothesis for ABC function is that they are poised to develop into plasma cells. While the inability of ABCs to produce antibody in response to BCR cross may seem to contradict this, ABCs have been shown to be responsive to TLR7 ligands (Hao et al. 2011, Rubtsov et al. 2011). Moreover, previous studies suggested ABCs in SLE patients expressed higher levels of PC transcription factors IRF4 and BLIMP, suggested they were PC-precursors, capable of differentiating in to autoantibody secreting cells (Jenks et al. 2018). Consistent with this, ABCs in SLE patients were found to be highly enriched with auto-antibody clones and vaccine-induced ABCs were enriched for influenza-specific clones (Lau et al. 2017, Wang et al. 2018). In animal models *in vivo* depletion of ABCs cells by B cell specific CD11c-DTR system, significantly reduced the generation of autoantibody in TLR7-ligand chronically treated mice (Rubtsov et al. 2011). However, a serious technical issue is whether CD11c-DTR system exclusively depleted ABCs without affecting other B cell populations because this system had been used to deplete general activated B cells including GC B cells (Baumjohann et al. 2013, Hong et al. 2018). Moreover, if ABCs were poised to differentiate into PCs, we would expect to observe cells in intermediate stages of differentiation by pseudotime analysis of single cell data. However scRNA-seq trajectory analysis on tonsillar B cells does not reveal that ABCs are closer to the PC cluster than other memory B cell subsets (King et al. 2021). Nevertheless, ABCs do retain the potential to differentiate into PCs, since adoptively transferring ABCs from pre-immunized mice could potentiate antibody secretion in recipient mice upon Virus Like Particle (VLP) immunization or LCMV infection (Du et al. 2019, Song et al. 2022).

Finally, ABCs have been proposed to be specialized for antigen uptake and presentation to T cells. Both sc-RNA seq and conventional RNA seq show that ABCs upregulated antigen presentation genes including MHC-II and CD86. Interestingly, $CD19^{hi}$ $T\text{-}bet^{+}$ ABCs in healthy and HIV infected individuals preferentially localized in non-GC and extrafollicular area in LNs

(Austin et al. 2019), and $CD19^{+}B220^{+}CD11c^{+}$ ABCs in aged mouse preferentially localized in the T-B border in the spleen (Rubtsov et al. 2015), implying that ABCs were specialized to interact with T cells. Consistent with this, non-specific $CD11c^{+}$ ABCs were found to be better at presenting OVA-protein to OT-II cells than follicular B cells (Rubtsov et al. 2015). Another interesting finding was ABCs were better than follicular B cells at priming OT-II cells into $CXCR5^{+}$ follicular helper T (Tfh)-like cells by i*n vitro* co-culture, while ABC generation preceded Tfh expansion in the $CD19^{cre}$ $Ship^{fl/fl}$ lupus model (Zhang et al. 2019). Notably, a recent study showed human $CD21^{-}CD27^{-}$ ABCs had enhanced BCR signaling and antigen uptake capacity to plate bound antigen (anti-λ/κ bound to plasma membrane sheets) (Ambegaonkar et al. 2020). Further work is required to understand the role antigen presentation by ABCs– it may be that ABCs present antigen to prime a particular T cell fate (e.g. Tfh vs Th1), or conversely they may have a regulatory role sequestering antigen and limiting antigen presentation.

1.6 B cell antigen uptake and presentation

As mentioned above, activated B cells interact with antigen specific $CD4^{+}$ T cells to initiate the T-dependent humoral response and such process requires B cell to uptake the antigen and presented it to $CD4^{+}$ T cells. In support of this, $CD19^{cre}$ $MHC\text{-}II^{fl}$ KO mice whose B cells cannot present antigen to cognate T cells have significantly reduced GC B cells and PC generation (Giles et al. 2015). Similarly, Ras homolog family member G (Rhog) KO antigen specific B1-8 cells that have impaired antigen uptake function have defect in forming GC B cells after immunization with particulate antigen (Martinez-Riano et al. 2018).

B cell antigen uptake and presentation requires the B cell to first gain access to the antigen in vivo followed by internalizing, digesting and presenting the antigen. Our current knowledge on B cell access to the antigen in vivo can be grouped into two categories. For small soluble antigens < 70kDa, they can traffic through the conduits into the B cell follicle where they are directly recognized by antigen specific B cells (Hoogeboom and Tolar 2016). For larger antigens, they cannot pass through the gap between stromal cells forming the conduits and in this case, they are captured by subcapsular sinus macrophages and antigen specific B cells then acquire antigen by forming an immune synapse with these macrophages (Carrasco and Batista 2007). Another proven route for the B cell to acquire antigen is from follicular dendritic cells (FDC) where antigens are transferred by non-antigen specific B cells and deposited on the

surface of FDC and recycled in non-degradable endosome to facilitate the antigen retention (Carrasco and Batista 2007, Heesters et al. 2013).

After binding with antigens through BCR, the B cells either internalize them if they are small soluble molecules or gather them and pinch them off if they are deposited on membrane surfaces (**Figure 1.7A, B**). The mechanisms of "pinch off" involves the myosin IIa-dependent, force-mediated pulling and exocytosis of lysosomes to the immune synapse (Yuseff et al. 2011, Natkanski et al. 2013). To be noted, B cells can not only pinch off the antigen presented but also a portion of non-antigen specific molecules through forming the immune synapse (Suzuki et al. 2009). After "pinching off" antigen, clathrin-mediated endocytosis is the dominant mechanism for antigen translocation from the B cell surface to the inside of the cell (Hoogeboom and Tolar 2016) (**Figure 1.7C**), followed by clathrin-coated pits formation (**Figure 1.7D**), endosome formation and fusion with lysosome (**Figure 1.7E**), and at last, presentation of digested antigen to the surface by MHC-II (**Figure 1.7F**) (Hoogeboom and Tolar 2016). In addition to well-studied clathrin-mediated endocytosis pathway, B cells have been showed to use clathrin independent endocytosis including fast endophilin-mediated endocytosis (Malinova et al. 2021) and phagocytosis (Martinez-Riano et al. 2018) to internalize antigens. Additionally, B cells have been showed to express caveolin (McShane and Malinova 2022), therefore the caveolin-dependent endocytosis might also contribute to B cells antigen uptake. Despite our advances in understanding B cell antigen uptake in the past two decades, most studies in this field have used artificial materials such as nanoparticles and plate-bound antigens. However it remains largely unknown how B cells uptake antigen in physiological settings such as under parasite infection. To investigate this in more detail we have taken advantage of the tools available for the study of B cell responses to the circumsporozoite protein antigen from the malaria parasite *Plasmodium falciparum*. This system allows the monitoring of B cell responses in the context not only of whole protein but also when presented in the context of a complex pathogen.

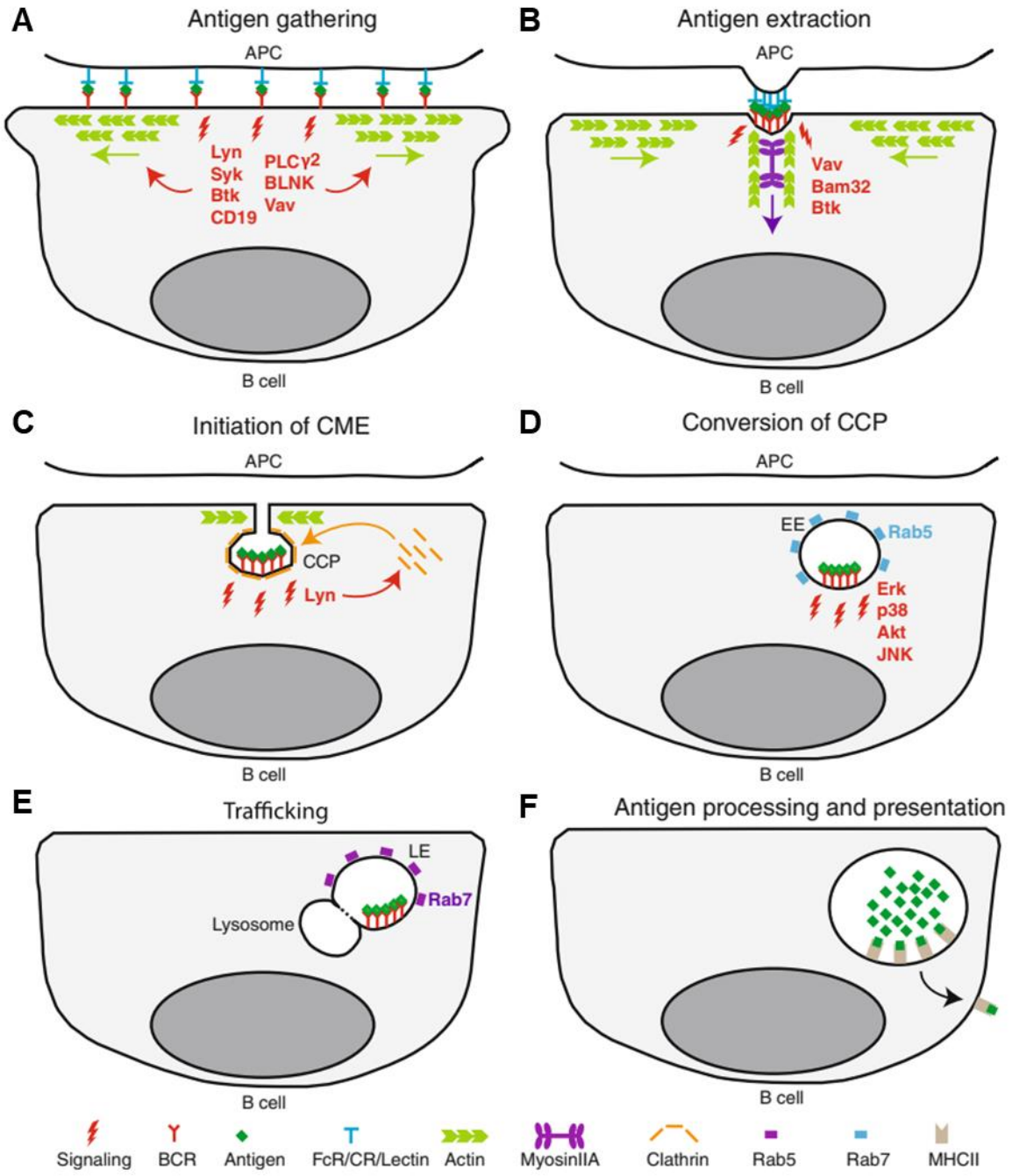


Figure 1.7 B cell antigen uptake and presentation through clathrin dependent endocytosis Critical steps involving B cell antigen uptake and presentations are summarized. (A) Upon antigen encounter, the B cell membrane spreads out and form the immune synapse with APCs. (B) B cells extract the antigens from APCs and centers them. (C-D) The antigens are wrapped by clathrin coated pits (CCP) and internalized. (E) The CCP are fused with lysosome. (F) The antigens are digested in the lysosome and loaded onto the MHC-II for antigen presentation. The figure is adopted from (Hoogeboom and Tolar 2016).

1.7 *Plasmodium* parasites

B cell responses are most often studied in the context of small molecules such as recombinant protein and haptens. However atypical B cell responses generally develop in the context of responses to pathogens, for example, in response to *Plasmodium* parasite infection in human and mouse (Gao and Cockburn 2022). Additionally, the study of B cell-antigen interactions, especially the antigen uptake by B cells, has not been well studied in the context of complex pathogens. Therefore we have turned to the *Plasmodium* parasites to address these questions.

Life cycle

The infectious disease malaria is caused by *Plasmodium* parasite infection which kills over 600000 people annually (World Health Organization 2020). In humans, *Plasmodium falciparum* and *vivax* are the most transmitted strains. For mouse models of malaria, *Plasmodium chabaudi*, *yoelii* and *berghei* are commonly used in the laboratory settings. Different strains of the *Plasmodium* parasite have similar life cycle. The infection of the mammal host begins from a mosquito taking a blood meal and injecting a few hundred *Plasmodium* parasites in the form of sporozoites into the skin. Some of the sporozoites travel into the liver through blood vessels where they develop into liver merozoites in the hepatocytes over the course of around a week in *P. falciparum* infections of humans (**Figure 1.8A**). At this stage, also known as the pre-erythrocytic stage, the infection is asymptomatic. At the conclusion of the liver stage the fully developed merozoites then egress out from the hepatocytes as merosomes (**Figure 1.8B**) (Sturm et al. 2006) and invade red blood cells initiating the erythrocytic or blood stages. At this stage the merozoites multiply themselves in infected red blood cells asexually, leading to the burst of the red blood cells and then start new rounds of infection (**Figure 1.8C**). The total number of merozoites in the blood multiply exponentially initially and usually increase from 10^1 parasites to 10^8-10^{10} parasites within ~10 days in a naive host resulting in the development of clinical symptoms like shaking chills and fever (Dondorp and Von Seidlein 2017). Ultimately malaria-naïve and many semi-immune patients will develop severe conditions such as cerebral malaria if they are left untreated. During the erythrocytic stage, some blood stage parasites will differentiate into gametocytes in response to cues that are poorly understood and if they are taken up by the mosquito in a following blood feed (**Figure 1.8D**), they mate in the mosquito's mid gut to have homologous recombination and to form zygotes (**Figure 1.8E**). After that, the zygotes develop into sporozoites again, marking the commence of a new life cycle (**Figure 1.8F**). Notably, in nearly all stages of the

life cycle (all stages in mammal host and in the form of sporozoite) *Plasmodium* parasites are haploid.

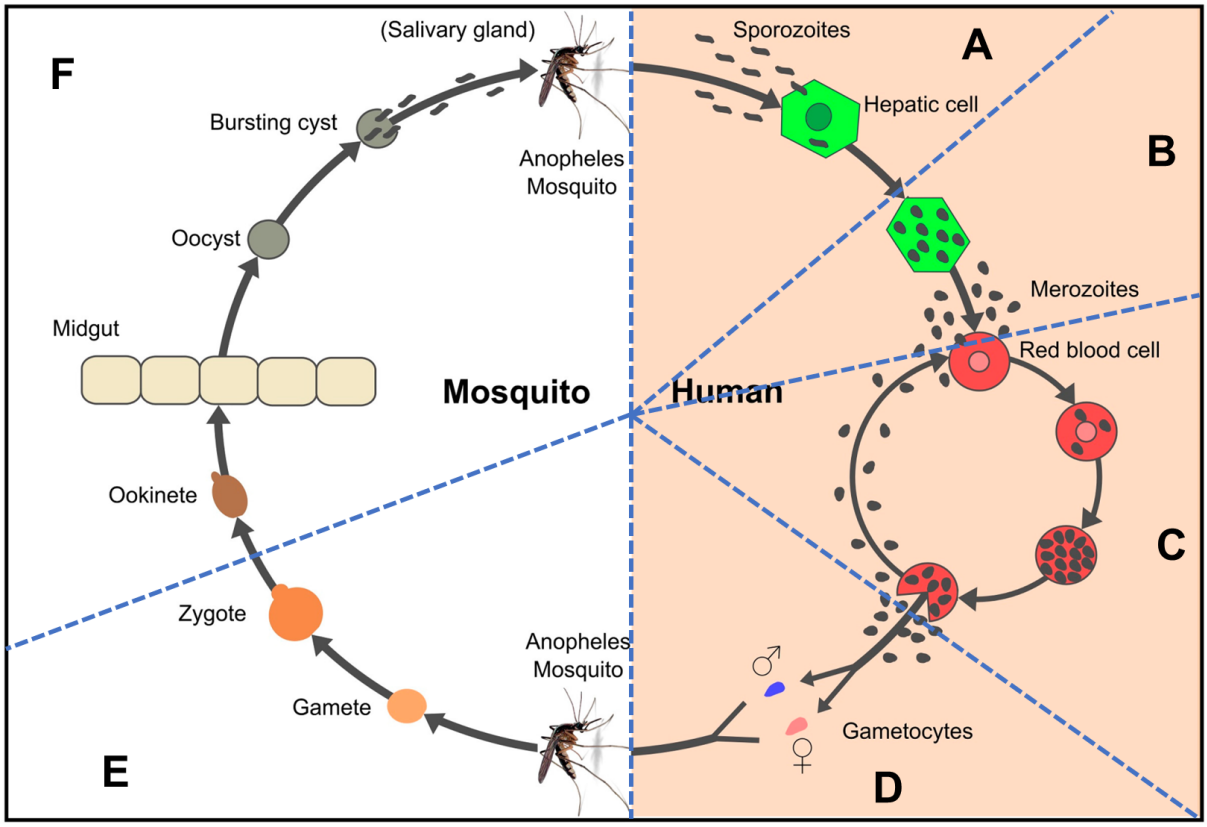


Figure 1.8 The life cycle of the *Plasmodium* parasite
The life cycle of a *Plasmodium* parasite. (A) The infection in the human starts from a mosquito having a blood meal and injecting hundreds of sporozoites into the mammal host, which invade the liver to initiate pre-erythrocytic stage and develop into merozoites. (B) The merozoites are released from the hepatocytes and infect and expand in the red blood cells to establish erythrocytic stage infection. (C) The infected red blood cells burst, release the merozoites resulting in multiple rounds of red blood cell infections. (D) During the erythrocytic stage some gametocytes are generated. (E) When the infected host is bitten by a mosquito again, some gametocytes are taken up by the mosquito, and they mate inside mosquito to form zygotes. (F) Zygotes develop into sporozoite again, marking the completion of one life cycle. The figure is adopted from (Curtidor et al. 2017).

Sporozoites

The sporozoites are morphologically crescent-shaped and about 8-14μm long (Frischknecht and Matuschewski 2017). The apical end of sporozoites is contains the apical complex, consisting of a microtubule centered polar ring and secretory vesicles termed micronemes and rhoptries (**Figure 1.9A**) (Kappe et al. 2004). The apical complex is responsible for the parasite invasion into the host cell and vacuole formation (Kappe et al. 2004). Another unique structure in sporozoites is called apicoplast, which is a non-photosynthetic plastid derived from algae involved in metabolism such as fatty acid synthesis (Sheiner et al. 2013). Apart from these unique structures, sporozoites have common organelles found in eukaryotic organisms like nucleus, golgi apparatus, mitochondria and endoplasmic reticulum (**Figure 1.9A**) (Frischknecht and Matuschewski 2017).

Individuals in malaria endemic areas do not develop sterilizing immunity to malaria. Thus other possible models for vaccine mediated immunity have long been sought. Notably in the 1960s it was shown that immunization with radiation attenuated sporozoites results in sterile protection against subsequent live parasite challenge in the mice (Nussenzweig et al. 1967). Attenuated sporozoites have been used as the model of malaria vaccines. The purpose of the attenuation is to prevent sporozoites from developing into hepatic merozoites to initiate blood stage infection, meanwhile keeping the sporozoite metabolic active and immunogenic, since killed sporozoites failed to induce adaptive immunity in mice (Nussenzweig et al. 1967). In light of this, study in human also demonstrated that X-ray treated, *Plasmodium falciparum* infected mosquito bites elicited adaptive immunity against controlled malaria infection (Hoffman et al. 2002). In addition to X-ray, chemoattenuation by pyrimethamine or chloroquine pre-treatment (Mwakingwe-Omari et al. 2021), and genetic modification that arrests the replication of blood stage parasites (Butler et al. 2011), also prevent the sporozoite to form merozoite while preserving their immunogenicity.

The sporozoites are estimated to express 3500-4200 genes according to the measurement of RNA expression in *Plasmodium falciparum* and *Plasmodium yoelii* (Lindner et al. 2019, Briquet et al. 2020). However to date, most of the *Plasmodium* genes are uncharacterized. The circumsporozoite protein (CSP) is the most well-studied gene identified in *Plasmodium* sporozoites (Frischknecht and Matuschewski 2017) as it is our only validated vaccine candidate. CSP is discovered by antigen precipitation assay as the major target of the antibody against *Plasmodium falciparum* infection in human (Clyde et al. 1973). It is abundantly expressed on the plasma membrane of the sporozoite and absent in any other stages of the *Plasmodium* life

cycle. While specific sequences vary in all species CSP is composed of (from N to C terminus) a signal peptide, region I, central tandem repeat region, region III, region II-plus, thrombospondin repeat region (TSR), and a GPI anchor (**Figure 1.9B**). Region I/II-plus/III and TSR are conserved motifs responsible for sporozoite attachment and invasion into the host cells. The central repeat region is composed of multiple repeats of strain specific amino acids (e.g. NANP for *Plasmodium falciparum;* PAPPNAND for *Plasmodium berghei*). The function of central repeats region remains undetermined, but nevertheless it is indispensable for the normal development of sporozoites, because knockout of the central repeats results in parasites that are able to complete the normal formation of zygotes in the mid gut, but have a defect in sporozoite formation (Ferguson et al. 2014). The GPI anchor contains carbohydrates, lipids, and phosphates and has been showed to induce pro-inflammatory response in the host immunity through interacting with TLR2 and TLR4 (Kurucz et al. 2013). Notably, the N terminus of the CSP molecule has been showed to be cleaved upon interaction with the heparan sulfate proteoglycan, and the inhibition of this process by blocking the cysteine protease activity significantly impairs the sporozoite infectivity in vivo (Coppi et al. 2005).

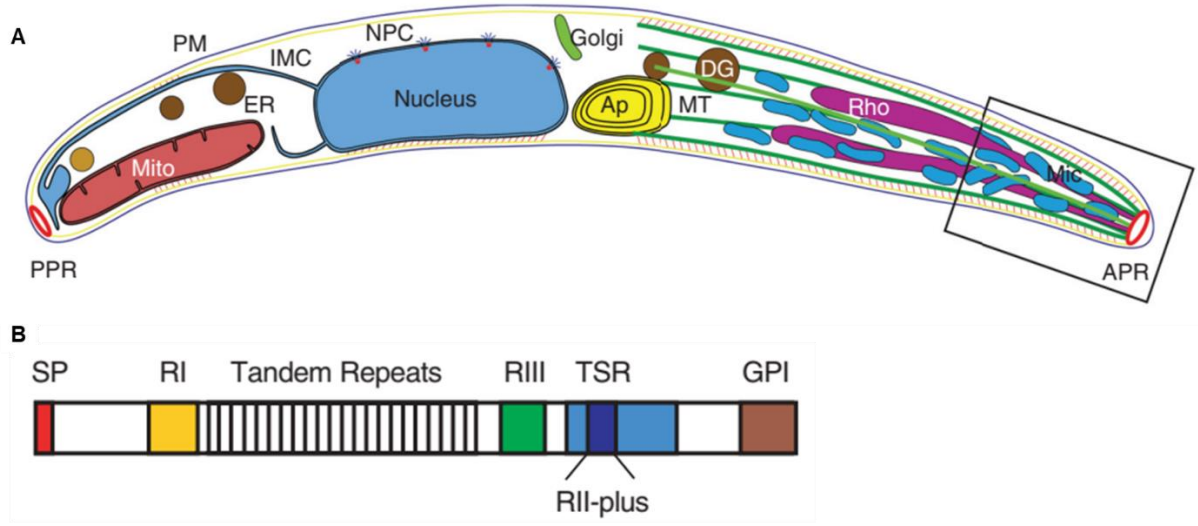


Figure 1.9 The structure of sporozoite and CSP protein
(A) Schematic of the SPZ structure and organelles. PM: plasma membrane. (From proximal to apical) PPR: proximal polar ring. Mito: mitochondria. ER: endoplasmic reticulum. IMC: inner membrane complex. NPC: nuclear pore complex. AP: apicoplast. MT: microtubules. FG: dense granules. Rho: rhoptries. Mic, micronemes. APR, apical polar ring. (B) Schematic of the CSP molecule. (From N to C terminus) The CSP contains signal peptide (SP), region one (RI), tandem repeats, region three (RIII), TSR including region II plus and GPI anchor. The pictures are adopted from (Kappe et al. 2004, Frischknecht and Matuschewski 2017).

Malaria vaccines

The RTS, S vaccine is the only approved malaria vaccine to prevent *Plasmodium falciparum* infection. RTS, S vaccine antigen is composed of CSP central repeat region, C terminus region and fused with hepatitis B antigen. The vaccine is delivered with the potent proprietary adjuvant AS01 which includes QS21 derived from saponin and MPL, a synthetic TLR4 agonist (Didierlaurent et al. 2017). Nonetheless protection induced by the RTS, S vaccine is not optimal: according to a phase III clinical trial in Africa endemic areas the efficacy against clinical malaria is 55.1% in children between 5-17 months old during 1 year after receiving the third dose, however this drops to 28% after 4 years (RTS,S Clinical Trials Partnership 2015). More recently, a CSP-based, Matrix-M adjuvanted malaria vaccine R21 showed protection in phase II clinical trial in children between 5-17 months old whose efficacy reaches 77% 1 year after the third dose (Datoo et al. 2021). Similar to AS01, matrix M is based on specific fractions of saponin, however it does not include any MPL or other TLR agonists. In addition to CSP-based vaccines, whole parasite vaccines such as those based on attenuated sporozoites have previously shown promising protection in providing sterile immunity against controlled human malaria infection in small numbers of US malaria-naïve adult recipients (Seder et al. 2013). However, in clinical settings the efficacy was found to be only 29% among Malian adults (Sissoko et al. 2017), and offers no protection in children aged 5-12 months old (Oneko et al. 2021).

For the RTS,S vaccine the titers of anti-CSP antibodies correlate broadly with the protection (RTS,S Clinical Trials Partnership 2015), suggesting that antibodies play a central role in offering the protection against *Plasmodium falciparum* infection. To be noted, sporozoites only need 15 minutes to infect hepatocytes once they enter into the blood vessel in mice, and even one sporozoite is sufficient to cause erythrocytic stage infection which the anti-CSP antibody is not protective anymore (Sidjanski and Vanderberg 1997). Therefore, not only the anti-CSP antibody titer should be maintained at high level but also the epitopes it recognizes should be carefully modulated to offer optimized protection. In support of this, studies have suggested that antibodies that target the central repeat region of the CSP, such as 2A10, prevented mice from infection upon sporozoite challenge (Zhang et al. 2017). Additionally, breath of the antibodies that binds to the C terminus region is associated with better protection after RTS, S vaccination, suggesting that anti-C terminus antibody also offers additional protection (Chaudhury et al. 2021). In contrast, the N-terminus binding antibody, 5D5 for example, failed to inhibit sporozoite development in the mosquitos (Thai et al. 2020), while CIS43 binding to the NVDP repeat region offers superior protection over 2A10 against murine malaria challenge

(Kisalu et al. 2018). In addition to antibody, tissue resident $CD8^+$ T cells in the liver has been also showed to play a central role in preventing mice from sporozoite infections (Fernandez-Ruiz et al. 2016), suggesting vaccines that prime the *Plasmodium* specific liver resident $CD8^+$ T cells might offer protection as well.

1.8 Key technologies used in this thesis

Igh^{g2A10} mice

The Igh^{g2A10} mice strain was generated by the previous work in our lab to study the antigen specific B cell response to sporozoite immunization (McNamara et al. 2020). The heavy chain germline reverted VDJ sequence of the CSP specific 2A10 antibody was inserted upstream of the IgM constant gene locus of the mouse genome. All B cells in the resulting Igh^{g2A10} mice should in theory express the 2A10 heavy chain, while the light chain is randomly recombined. As a result, around 2% of the naïve B cells in Igh^{g2A10} mice are antigen specific to the central repeat region of CSP. Adoptive transfer of Igh^{g2A10} B cells into WT recipient mice followed by sporozoite immunization successfully induces the differentiation of Igh^{g2A10} B cells into effector B cell populations such as PC, GC B cells and memory B cells (McNamara et al. 2020).

Generation of TCR transgenic mice

TCR transgenic mice are a widely used tool in immunology. Unlike wild type (WT) mice that have T cells with a diverse TCR epitope, the T cells in TCR transgenic mice uniformly recognize one specific antigen epitope. Most commonly used TCR transgenic mice were developed in the late the 20^{st} century, by co-injecting the linearized DNA constructs that encoding the TCR alpha and beta chain respectively into the embryo. The TCR sequences used to be obtained by PCR amplifying the VDJ region of the hybridoma T cell lines, but with the advance in sequencing techniques recent studies also directly synthesize the sequences, bypassing the steps to generate antigen specific hybridoma cell lines (Lee et al. 2021). One advantage of directly synthesizing over amplifying from hybridoma cell line is the VDJ sequences from the former method are more physiological, while the later method tends to select clones that have the strongest proliferation in vitro. Although the use of TCR expressing constructs differs between groups (Cho et al. 2020), the outcome is largely the same, that the T cells in transgenic mice can strongly proliferate and differentiation upon corresponding antigen stimulation. Despite the success in generating TCR transgenic strains like OT-II and SMARTA,

it remains theoretically unclear that why the alpha and beta chain constructs need to be controlled by different promoters and co-injected (Barnden et al. 1998). Also, the injected DNA will randomly integrate into the genome and it remains unclear the copy number and the locus of the transgenes in the chromosome.

Single-cell RNA-sequencing (scRNA-seq)

scRNA-seq is the technique to obtain the transcriptomic information at single cell level. A now seminal study modified the protocol of reverse transcription used for microarray assay, and identified 75% more genes in a single blastomere than the conventional microarray technique (Tang et al. 2009). Since then, many types of scRNA-seq techniques have been reported that vary in the methods for cell isolation, cDNA construction, strand specificity and transcript coverage (Chen et al. 2019). Nevertheless, these methods all aim at recovering complete transcriptomic information from individual cells. After years of improvement the current scRNA-seq technique can detect up to 10^5 to 10^6 cells per run (Natarajan et al. 2019, Jia et al. 2022). In addition to detecting transcriptomic information, scRNA-seq can be combined with other techniques to obtain other information like chromatic accessibility (e.g. single cell ATAC-seq), surface protein expression (e.g. cellular indexing of transcriptomes and epitopes by sequencing, CITE-seq), and in situ localization (e.g. Slide-seq). Notably, one feature that is commonly used in CITE-seq is called cell hash-tagging, a sample pooling technique that labels the cells from different samples with anti-CD45/MHC-I antibody ligated with different DNA barcodes.

In my thesis the scRNA-seq experiments are conducted using the commercialized drop-based method provided by 10X Genomics combined with CITE-seq and VDJ sequencing. For this method the sorted cells are labeled by DNA-barcoded antibodies and wrapped together with DNA-barcoded gel beads (each bead has a unique barcode) in oil formed vesicles to generated cDNA library and the antibody detection library. Because the unique gel beads DNA barcodes are incorporated into the DNA molecules after this stage, the information of transcriptomes, antibody binding level and VDJ sequences of a single cell can be integrated according to the unique barcode for bioinformatic analysis.

Clustered regularly interspaced short palindromic repeats associated protein 9 (CRISPR-Cas9) mediated gene knock-out

CRISPR-Cas9 system is a defense mechanism in bacteria against invasive nuclei acid and has been widely used as an efficient gene editing tool (Zhao et al. 2021). The active CRISPR-Cas9 system composes of a single guide RNA (sgRNA) and a Cas9 endonuclease (mainly derived from streptococcus pyogene). The Cas9-sgRNA complex randomly interacts with cellular DNA molecules and searches for sequences that match the protospacer adjacent motif (PAM) and sgRNA. Once the matching is specific, the Cas9 endonuclease introduces a double strand DNA break (DSB) exactly at 3 nucleotides upstream of the PAM. This results in the non-homologous end joining DNA repair mechanism which often results in a frameshift at the DSB site and thus the inactivation of the gene if the sgRNA targets the exon region (Lino et al. 2018).

The efficient delivery of CRISPR-Cas9 system into the nuclear is crucial to achieve high KO frequency. One of the popular methods to deliver CRISPR-Cas9 system to the lymphocytes is to use a viral vector that encodes the Cas9 protein and sgRNA. The viral vector is introduced by an adenovirus or lentivirus and this approach usually results in high gene KO rate. This method is also suitable for genome-wide CRISPR-Cas9 library screening. However, the continuous expression of the Cas9 protein and sgRNA in the target cell can exaggerate the off-target problem, the unintended cleavage of untargeted genomic sites showing a similar sequence compared to the target site (Modrzejewski et al. 2020), and also initiate an inflammatory response in the cell (Lino et al. 2018). Also, the KO occurs very slowly, usually ~3 days after the viral infection so it is not suitable for many prone-to-apoptosis primary cell types. Another popular and well-established method, which is used in my thesis, is conjugating recombinant Cas9 protein with sgRNA in vitro to form ribonuclear protein (RNP) followed by nucleofection delivery. Because the RNP complex is biological active, it can introduce gene KO within 24h after transfection. The effects of off-targeting and pro-inflammation are also reduced because the RNP can be degraded within 24h by the cell's endogenous protein degradation machinery (Lino et al. 2018). However, the limitations are that it is low-throughput and relatively expensive.

1.9 Aims and rational for this study

The general aim of my work is to understand T and B cell collaboration during humoral responses, and to explore the methods that could potentiate the vaccine efficacy on the grounds of these new knowledge. The general aim is targeted from three aspects: 1) Tfh cells and memory maintenance, 2) atypical B cell development and function and 3) B cell antigen uptake mechanism during large parasite invasion.

AIM1: To investigate the persistence of memory Tfh cells in vaccination

In light of the recent Covid-19 pandemic, it is impertinent to understand the mechanism behind immunological memory and how can we improve the longevity of vaccine induced protection. Despite large amount of works that have focused on vaccine induced plasma cell, memory B cells, Th1 cells and $CD8^+$ T cells, little is known about how memory Tfh cells persist. Memory Tfh cells are classified into Tfh1, Tfh2 and Tfh17 cells (Morita et al. 2011) and we specifically asked which cell type have superior persistence. To address this question, HBV vaccinated human cohorts are recruited. HBV specific Tfh1/2/17 cells are identified by antigen induced marker (AIM) assay (Dan et al. 2016) and their persistence are measured on day7 and day30 after vaccination. An in vitro culture method is also developed (Gao et al. 2020) to generate murine Tfh1/2/17-like OT-II cells to facilitate functional analysis.

AIM2: To investigate the key TFs that regulate the formation of ABCs

ABCs have been discovered nearly two decades ago but functions of this population in healthy individuals and in pathogenic conditions remains largely unknown. One of the main obstacles is the key TFs for ABCs are undetermined, thus there is no viable method to deplete this population for functional studies. To find out the key TFs of ABC development, scRNA-seq dataset on human tonsil cells are generated to find potential TF candidates. We also developed in vitro culture method and in vivo *Plasmodium* sporozoite immunization and *Plasmodium chabuadi* infection mouse models to induce ABC formation, and deployed CRISPR-Cas9 KO methods to validate the roles of these TFs in ABCs formation.

AIM3: To investigate the antigen presentation of B cells in response to large parasite invasion

Our current know on B cell antigen uptake and presentation are predominately based on virus, recombinant protein and artificial materials like nanoparticles. Although we always know the host can produce antibody against large cellular parasite infection, there is a complete lack of knowledge on how B cells uptake antigen for the parasite and present it to T cells to induce anti-parasite humoral response. We specifically ask how *Plasmodium falciparum* CSP (PfCSP) specific B cells uptake antigen and how they are helped by T cells to produce antibody. To address this question, we use the *Plasmodium berghei* sporozoites as the model and use several imaging techniques including scanning electron microscope (SEM) and image flow cytometry to visualize the Igh^{g2A10} B cell (PfCSP specific B cells generated by our lab (McNamara et al. 2020)) and sporozoite interaction. Additionally, we generated a TCR transgenic mouse strain plus several transgenic strains of *Plasmodium* parasite expressing OVA on spatially different locations of the sporozoites to study the B-T interaction upon SPZ immunization.

Chapter2- Material and Methods

2.1 Acknowledgement of major contributions

The mosquito insectary at the John Curtin School of Medical Research (JCSMR) was maintained by Yeping Cai, Lauren Abraham and Patricia Carreira. Mouse husbandry and breeding services were provided by the Australian Phenomics Facility (APF). The human tonsil dataset was generated under the collaboration with Qian Shen at Vinuesa Group at JCSMR. The library construction and sequencing for bulk and single cell RNA-seq were performed by Maxim Nekrasov at JCSMR Biomolecular Resource Facility. The embryo microinjection and transgenic mice screening were performed by Dominik Spensberger von Wiorogorsk and Rachel Milne at Phenogenomic Targeting Facility, ANU. The SEM picture was taken with the help of Jiwon Lee and Melanie Rug at the Centre for Advanced Microscopy, ANU.

2.2 Buffers/ solutions/ media, reagents and antibodies

Buffers/ media/ solutions		
Name (composition)	*Source*	*Identifier*
Dulbecco′s phosphate buffered saline	Sigma	D8537-500ML
ACK Lysis Buffer	ThermoFisher	A1049201
AccuStain Giemsa	Sigma	WG80-2.5L
Fetal bovine serum (FBS)	ThermoFisher	10099141
Flow cytometry buffer (FACS buffer, DPBS+ 2% (v/v) fetal bovine serum (FBS))	NA	NA
Complete RPMI media (10% FBS (v/v), 100 units/mL penicillin, 100 μg/mL streptomycin, 1 mM sodium pyruvate, 1% MEM nonessential amino acids (v/v), 0.055 mM β-Mercaptoethanol and 25 mM HEPES in RPMI 1640 with L-glutamine)	NA	NA
Plasmodium berghei culture media (20% FBS (v/v) in RPMI 1640 with L-glutamine)	NA	NA
ELISA blocking buffer (1% (w/w) bovine serum albumin (BSA) dissolved in PBS)	Sigma	A7030
ELISA wash buffer (DPBS+ 0.05% (v/v) Tween-20)	Sigma	P9416-50ML
ELISA developing solutions	SeraCare	5120-0047

ELISA stop solution (1% (w/w) sodium dodecyl sulfate in DPBS)	Sigma	436143-25G
Tris-acetate-EDTA (TAE) buffer	ThermoFisher	B49
Glutaraldehyde solution	Sigma	G6257
Ethanol	Sigma	1009831000
Methanol	Sigma	34860-1L-R
Dimethyl Sulfoxide	Sigma	34869

Reagents		
Name	*Source*	*Identifier*
Chromium Next GEM chip G Single cell Ki	10X Genomics	1000127
Chromium Next GEM Single 5' Library and Gel Bead Kit	10X Genomics	1000167
Chromium Single cell 5' Feature Barcode Kit	10X Genomics	1000256
Chromium Single cell 5' Library Construction Kit	10X Genomics	1000020
Chromium Single cell V(D)J Enrichment Kit Human B cell	10X Genomics	1000016
Chromium Single cell V(D)J Enrichment Kit Mouse T cell	10X Genomics	1000071
Single Index Kit N Set A	10X Genomics	1000212
Single Index Kit T Set A	10X Genomics	1000213
Primer-qPCR-Bcl6-F-CAGAGATGTGCCTCCATACTGC	IDT	NA
Primer-qPCR-Bcl6-R-CTCCTCAGAGAAACGGCAGTCA	IDT	NA
Primer-qPCR-CD40L-F-GAACTGTGAGCAGATGAGAAGGC	IDT	NA
Primer-qPCR-CD40L-R-TGGCTTCGCTTACAACGTGTGC	IDT	NA
Primer-qPCR- ICOS-F-GCAGCTTTCGTTGTGGTACTCC	IDT	NA
Primer-qPCR-ICOS-R-TGTGTTGACTGCCGCCATGAAC	IDT	NA

Primer-qPCR-TBX21 F-CACTACAGGATGTTTGTGGACGTG	IDT	NA
Primer-qPCR-TBX21-R-CCCCTTGTTGTTTGTGAGCTTTAG	IDT	NA
Primer-qPCR-GATA3-F-TGTCTGCAGCCAGGAGAGC	IDT	NA
Primer-qPCR-GATA3-R-ATGCATCAAACAACTGTGGCCA	IDT	NA
Primer-qPCR-RORC-F-TCTGGAGCTGGCCTTTCATCATCA	IDT	NA
Primer-qPCR-RORC-R-TCTGCTCACTTCCAAAGAGCTGGT	IDT	NA
Primer-qPCR-GAPDH-F-TGCACCACCAACTGCTTAG	IDT	NA
Primer-qPCR-GAPDH-R-GGATGCAGGGATGATGTTC	IDT	NA
Fgr-F (AGGCGAGAAGTTCCACATCCTG)	IDT	N/A
Fgr-R (GGATGGAATCCACAGGAGCAAC)	IDT	N/A
Itgax-F (TGCCAGGATGACCTTAGTGTCG)	IDT	N/A
Itgax-R (CAGAGTGACTGTGGTTCCGTAG)	IDT	N/A
Itgam-F (TACTTCGGGCAGTCTCTGAGTG)	IDT	N/A
Itgam-R (ATGGTTGCCTCCAGTCTCAGCA)	IDT	N/A
Cd72-F (AGCGGCTTAGAGGAGTTGCTAG)	IDT	N/A
Cd72-R (CCTGACACAATGTCGGCTTTGAG)	IDT	N/A
Cxcr5-F (ATCGTCCATGCTGTTCACGCCT)	IDT	N/A
Cxcr5-R (CAACCTTGGCAAAGAGGAGTTCC)	IDT	N/A
Cd55-F (CCAACTCCTCAGAAACCTTCCAC)	IDT	N/A
Cd55-R (CCTGTGTTAGGCTCTCCTTTGTC)	IDT	N/A
Fcer2a-F (CTAAGGAACGCCCAATCCCAGA)	IDT	N/A
Fcer2a-R (CTTTGCCACCTCTTCCTGGAGT)	IDT	N/A
Zeb1-F (ATTCAGCTACTGTGAGCCCTGC)	IDT	N/A
Zeb1-R (CATTCTGGTCCTCCACAGTGGA)	IDT	N/A
Hck-F (GCAGTTTGGAGAAGTGTGGATGG)	IDT	N/A

Hck-R (TGCAGCGACTTCATCAGGTTGG)	IDT	N/A
Tbx21-F (CCACCTGTTGTGGTCCAAGTTC)	IDT	N/A
Tbx21-R (CCACAAACATCCTGTAATGGCTTG)	IDT	N/A
Lgals1-F (GTAACACCAAGGAAGATGGGACC)	IDT	N/A
Lgals1-R (TCATGTCCGTCTGGCAGCTTGA)	IDT	N/A
Sell-F (TACATTGCCCAAAAGCCCTTAT)	IDT	N/A
Sell-R (CATCGTTCCATTTCCCAGAGTC)	IDT	N/A
Cr2-F (AGAGTGTAAGCCAGTAGGACCAC)	IDT	N/A
Cr2-R (GACACAGTTGATAAGCACTCTCAC)	IDT	N/A
Sox5-F (CGCCAGATGAAAGAGCAACTCAG)	IDT	N/A
Sox5-R (TGAGTCAGGCTCTCCAGTGTTG)	IDT	N/A
Cd24a-F (ACATCTGTTGCACCGTTTCCCG)	IDT	N/A
Cd24a-R (CAGGAGACCAGCTGTGGACTG)	IDT	N/A
Bhlhe40-F (CGGATTAACGAGTGCATTGCCC)	IDT	N/A
Bhlhe40-R (GCTTCAACGTAAGCTCCAGAACC)	IDT	N/A
Ubc-F (GCCCAGTGTTACCACCAAGA)	IDT	N/A
Ubc-R (CCCATCACACCCAAGAACA)	IDT	N/A
Sanger-Bcl6-F- TTTGGGCCTTACAGGGCAAG	IDT	NA
Sanger-Bcl6-R- GCTTGTTTTACAGCGGCCTG	IDT	NA
Sanger-Bcl6-inner- GTGTCTGCCTAGAGCATGTGA	IDT	NA
Sanger-Zeb2-F- TCCAGAAGGGCGATGGAAAG	IDT	NA
Sanger-Zeb2-R- TGAATATGTGGGGGCATTGGT	IDT	NA
Sanger-Zeb2-inner- CACTTACCTGGACCGGCTAC	IDT	NA
Chr12-integration-F1- TTCCGTATCTTATATAATGACGTAGCAATATATCA	IDT	NA
Chr12-integration-F2- CATGAGCATACAAAAATACATGCACACATATATAA AC	IDT	NA
Chr12-integration-R1- GCAACCTTATATTCCTCAATTACAACCAAC	IDT	NA
Chr12-integration-R2- ATCTTACAACAAGCCCTGATCTTCTCTT	IDT	NA

*Falciparum*CSP-ova-flag-F1-CCAGGAGGTGGAGGTATCAACTTT	IDT	NA
*Falciparum*CSP-ova-flag-F2-AAGTGTACTTACCTCGCATGAAGATGGA	IDT	NA
*Falciparum*CSP-ova-flag-R1-CATATTGTGACCTTGTCCATTACCTTGATT	IDT	NA
*Falciparum*CSP-ova-flag-R2-GCAGAGCCAGGCTTTATTCTAACTTGAATA	IDT	NA
*Berghei*CSP-ova-flag-F1-CAAATGACGGAGGTGGAGGTCATAT	IDT	NA
*Berghei*CSP-ova-flag-F2-CATGCAGCACATGCAGAAATCAATG	IDT	NA
*Berghei*CSP-ova-flag-R1-ACCATTCCTCTGTGATACTATCCCTGATC	IDT	NA
*Berghei*CSP-ova-flag-R2-CTTGAACATTTATCCATTTTACAAATTTCAGTATC	IDT	NA
SpCas9, 1000 pmol	Synthego	N/A
sgRNA-eGFP (GGGCGAGGAGCUGUUCACCG)	Synthego	N/A
sgRNA-Zeb2 (GGUGAACUAUGACAACGUAG)	Synthego	N/A
sgRNA-Bcl6 (AGAUCCUGAAAUCAGCCCUG)	Synthego	N/A
sgRNA-Tbx21 (CUCGUCACUCGGCAUCGGCU)	Synthego	N/A
sgRNA-Bhlhe40 (AAACUUACAAACUGCCGCAC)	Synthego	N/A
sgRNA-Il21r (AGGUCCAGGCAGCUCCAAGC)	Synthego	N/A
sgRNA-Maf (UCCAUGGCCAGGGGACUGGU)	Synthego	N/A
sgRNA-Mafb (UUGCCCCAUGCUCAGCUCCG)	Synthego	N/A
sgRNA-Tcf7l2 (CCGCCAUGUUAGCGGCCAAG)	Synthego	N/A
sgRNA-Bhlhe41 (UUCUUUUCUAUUAAUCUGUG)	Synthego	N/A
sgRNA-Runx2 (GUAGGUUGUAGCCCUCGGAG)	Synthego	N/A
Negative Control, Scrambled sgRNA#1	Synthego	N/A
Mouse B Cell Nucleofector™ Kit	Lonza	VVPA-1010
Recombinant Murine IL-21	Peprotech	210-21
Recombinant Murine IL-6	Peprotech	216-16
Recombinant Murine IL-12 p70	Peprotech	210-12

Recombinant Murine IL-4	Peprotech	210-14
Recombinant Human TGF-β1 (CHO derived)	Peprotech	100-21C
LPS from E. coli	Enzo	ALX-581-013-L002
LEGENDplex™ Mouse Immunoglobulin Isotyping Panel (6-plex) with Filter Plate	Biolegend	740492
Albumin from chicken egg white	Sigma	A5503-1G
NP-OVAL (Ovalbumin)	Bioseach	N-5051-100
NP-BSA (Bovine Serum Albumin), Ratio 1-4	Bioseach	N-5050XL-10
NP-BSA (Bovine Serum Albumin), Ratio > 20	Bioseach	N-5050H-10
Recombinant CSP protein	Genescript	NA
$OVA_{323-339}$	Genescript	NA
Recombinant HBV surface antigen	Beijing Bioforce	NA
HBVSA ELISA kit	Shanghai Kehua Bio-engineering Co	NA
Measles (Rubeola) Virus Antigen	Genway	GWB-636184
Tetanus toxin from Clostridium tetani	Sigma	582231-25UG
CellTrace™ CFSE Cell Proliferation Kit, for flow cytometry	ThermoFisher	C34570
Dynabeads™ Human T-Activator CD3/CD28 for T Cell Expansion and Activation	ThermoFisher	11161D
Imject™ Alum Adjuvant	ThermoFisher	77161
CD4+ T Cell Isolation Kit, mouse	Miltenyibiotec	130-104-454
Anti-PE MicroBeads	Miltenyibiotec	130-048-801
Anti-Biotin MicroBeads	Miltenyibiotec	130-090-485
LS Columns	Miltenyibiotec	130-042-401
RNeasy Micro Kit (50)	Qiagen	74004
iScript™ Reverse Transcription Supermix for RT-qPCR	Biorad	1708840

SYBR™ Green PCR Master Mix	ThermoFisher	4309155
QIAquick Gel Extraction Kit	Qiagen	28706X4
QIAprep Spin Miniprep Kit (50)	Qiagen	27104
Monarch® Genomic DNA Purification Kit	New England Biolab	T3010S
ClaI	New England Biolab	R0197S
BamHI	New England Biolab	R0136S
NotI	New England Biolab	R0189S
BglII	New England Biolab	R0144S
ApaI	New England Biolab	R0114S
PvuI	New England Biolab	R3150
NEB® 5-alpha Competent E. coli (High Efficiency)	New England Biolab	C2987U
Pyrimethamine	Sigma	46706
Saponin	Sigma	47036
7-AAD Viability Staining Solution	Biolegend	420404
Zombie Aqua™ Fixable Viability Kit	Biolegend	423102
FITC Annexin V	Biolegend	640905
CountBright™ Absolute Counting Beads, for flow cytometry	ThermoFisher	C36950

Antibodies		
Name	*Source*	*Identifier*
TotalSeq™-C0144 anti-human CD185 (CXCR5) Antibody	Biolegend	356939

TotalSeq™-C0140 anti-human CD183 (CXCR3) Antibody	Biolegend	353747
TotalSeq™-C0053 anti-human CD11c Antibody	Biolegend	37152
TotalSeq™-C0577 anti-human CD73 (Ecto-5'-nucleotidase) Antibody	Biolegend	344031
TotalSeq™-C0156 anti-human CD95 (Fas) Antibody	Biolegend	305651
TotalSeq™-C0154 anti-human CD27 Antibody	Biolegend	302853
TotalSeq™-C0389 anti-human CD38 Antibody	Biolegend	303543
TotalSeq™-C0384 anti-human IgD Antibody	Biolegend	348245
TotalSeq™-C0055 anti-human CD138 (Syndecan-1) Antibody	Biolegend	356539
TotalSeq™-C0181 anti-human CD21 Antibody	Biolegend	354923
TotalSeq™-C0829 anti-human CD307e (FcRL5) Antibody	Biolegend	340309
TotalSeq™-C0828 anti-human CD307d (FcRL4) Antibody	Biolegend	34021
TotalSeq™-C0050 anti-human CD19 Antibody	Biolegend	302265
TotalSeq™-C0394 anti-human CD71 Antibody	Biolegend	334125
TotalSeq™-C0251 anti-human Hashtag 1 Antibody	Biolegend	394661
TotalSeq™-C0252 anti-human Hashtag 2 Antibody	Biolegend	394663
TotalSeq™-C0253 anti-human Hashtag 3 Antibody	Biolegend	394665
PE/Cyanine7 anti-human CD19	Biolegend	982410
Brilliant Violet 421™ anti-human CD10 Antibody	Biolegend	312218
Brilliant Violet 785™ anti-human CD19 Antibody	Biolegend	302240
Brilliant Violet 605™ anti-human IgD Antibody	Biolegend	348232
APC anti-human CD185 (CXCR5) Antibody	Biolegend	356908
APC anti-mouse/human CD45R/B220 Antibody	Biolegend	103212
Brilliant Violet 421™ anti-human CD183 (CXCR3) Antibody	Biolegend	353716
Brilliant Violet 421™ anti-human CD196 (CCR6) Antibody	Biolegend	353408
Brilliant Violet 711™ anti-human CD279 (PD-1) Antibody	Biolegend	329928

Alexa Fluor® 700 anti-human CD4 Antibody	Biolegend	300526
PE/Cyanine7 anti-human CD197 (CCR7) Antibody	Biolegend	353226
Brilliant Violet 605™ anti-human CD45RA Antibody	Biolegend	304134
APC/Cyanine7 anti-human CD134 (OX40) Antibody	BD	350022
BB515 Mouse Anti-Human CD279 (PD-1)	BD	565936
PE-CF594 Mouse Anti-Human IgD	BD	562540
BV786 Mouse Anti-Human CD3	BD	565491
PE-CF594 Mouse Anti-Human CD27	BD	562324
PE-CF594 Mouse Anti-Human CD25	BD	562403
TruStain FcX™ (anti-mouse CD16/32) Antibody	Biolegend	101320
Brilliant Violet 421™ anti-mouse CD279 (PD-1) Antibody	Biolegend	135218
Brilliant Violet 650™ anti-mouse CD183 (CXCR3) Antibody	Biolegend	126531
PE/Cyanine7 anti-mouse CD196 (CCR6) Antibody	Biolegend	129816
APC/Cyanine7 anti-mouse/human CD44 Antibody	Biolegend	103028
Brilliant Violet 605™ anti-mouse CD69 Antibody	Biolegend	104530
APC/Cyanine7 anti-mouse CD25 Antibody	Biolegend	102026
PE/Cyanine7 anti-mouse CD23 Antibody	Biolegend	101614
PE/Cyanine7 anti-mouse CD138 (Syndecan-1) Antibody	Biolegend	142514
Brilliant Violet 510™ anti-mouse CD138 (Syndecan-1) Antibody	Biolegend	142521
Alexa Fluor® 488 anti-mouse/human GL7 Antigen (T and B cell Activation Marker) Antibody	Biolegend	144612
Brilliant Violet 605™ anti-mouse IgD Antibody	Biolegend	405727
Brilliant Violet 421™ anti-mouse CD11c Antibody	Biolegend	117330
PE/Dazzle™ 594 anti-mouse CD185 (CXCR5) Antibody	Biolegend	145522
Brilliant Violet 650™ anti-mouse CD45.1 Antibody	Biolegend	110735
PE anti-mouse CD45.1 Antibody	Biolegend	110708
PE anti-mouse/human CD45R/B220 Antibody	Biolegend	103207
PE/Cyanine7 anti-mouse CD95 (Fas) Antibody	Biolegend	152618
APC anti-mouse CD79b (Igβ) Antibody	Biolegend	132808
Brilliant Violet 605™ anti-T-bet Antibody	Biolegend	644817

Brilliant Violet 421™ anti-GATA3 Antibody	Biolegend	653814
PerCP-Cy™5.5 Mouse Anti-Mouse RORγt	BD	562683
PE-CF594 Mouse Anti-Bcl-6	BD	562401
BUV395 Rat Anti-Mouse CD19	BD	563557
BUV737 Rat Anti-Mouse CD19	BD	612781
BUV496 Mouse Anti-Mouse CD45.2	BD	741092
BUV496 Mouse Anti-Mouse CD45.1	BD	741093
BUV395 Rat Anti-Mouse CD4	BD	740208
FITC Rat Anti-Mouse CD21/CD35	BD	561769
CD38 Monoclonal Antibody (90), Alexa Fluor™ 700, eBioscience™	ThermoFisher	56-0381-82
Ax405, anti-PfCSP (clone Mab10)	NIH	NA
Ultra-LEAF™ Purified anti-mouse CD40 Antibody	Biolegend	102812
InVivoMAb anti-mouse IFNγ	Bioxcell	BE0055
InVivoMAb anti-mouse IL-4	Bioxcell	BE0045
InVivoMAb anti-mouse/human/rat/monkey/hamster/canine/bovine TGF-β	Bioxcell	BE0057

2.3 Human samples and ethics statement

For Chapter3, written informed consent was obtained from participants or the parents of children participants according to the ethics approved by human ethics committees of Renji Hospital affiliated to Shanghai Jiao Tong University School of Medicine (KY2019-161), Fourth Military Medical University (KY20163344-1), Tongji Hospital (NCT05009134), Shanghai Children's Medical Centre affiliated to Shanghai Jiao Tong University School of Medicine and Obstetrics and Gynecology Hospital of Fudan University (Kyy2018-6). Buffy coats from healthy donors for bulk RNA-seq were obtained from the blood bank of Changhai Hospital affiliated to Navy Medical University, Shanghai, China.

For Chapter4, the study on human tonsillar cells was approved by the Australian National University Human Experimentation Ethics Committee and Australian Capital Territory Government Centre for Health and Medical Research (Protocol numbers: ANU 2012/517; ACT Health ETH.4.06.268). The study on Mowat Wilson syndrome (MWS) patient and age/gender matched healthy controls was approved by the Sydney Children's Hospitals Network Human

Research Ethics Committee (MWS patients, protocol number: LNR/15/SCHN/195) and St Vincent's Hospital Sydney Human Research Ethics Committee (healthy controls, protocol numbers: HREC/13/SVH/145; 2019/ETH03336). All research involving human samples was conducted in accordance National Statement on Ethical Conduct in Human Research 2018.

Demographics

Cohort description	*Number*	*Female/ male*	*Median age (range)*
Healthy individuals for cTfh phenotyping	33	26/7	35 (21-71)
Healthy individuals received HBV vaccines	38	8/29	19 (18-20)
Healthy individuals for measles and tetanus toxin AIM assay	20	11/9	24 (18-32)
Healthy children	18	14/4	6 (0.5-12)
Cord blood	5	2/3	0 (0-0)
Recovered Covid-19 patients	13	9/4	33 (23-52)
Healthy individuals for qPCR and cytokine assay	14	9/5	42.5 (27-51)
Tonsils for scRNA-seq	2	2/0	4.5 (4-5)
MWS patients	5	1/4	12 (3-21)
Healthy controls for MWS patients	6	2/4	21.5 (21-24)

2.4 Mice strains and ethics statement

All animal procedures were approved by the Animal Experimentation Ethics Committee of the Australian National University (Protocol number: 2019/36). All research involving animals was conducted in accordance with the National Health and Medical Research Council's Australian Code for the Care and Use of Animals for Scientific Purposes and the Australian Capital Territory Animal Welfare Act 1992. All mice are B6 background.

Mouse strains

Name	*Phenotype*	*Source*
C57BL/6 (B6)	An inbred wildtype mouse strain. Carries the Ptprcb (CD45.2) congenital pan-leukocyte marker.	Bred in House, originally from the Jackson Laboratories (USA).
B6N-CD45.1	An inbred wildtype mouse strain. Carries the Ptprca (CD45.1) congenital pan-leukocyte marker.	Bred in House, originally from the Jackson Laboratories (USA).
CD28KO	CD28 is required for T cell activation. This strain has defective CD4 and CD8 T cell responses.	Gift from Prof. Carola Vinuesa Lab. Originated from (Shahinian et al. 1993).
MD4	BCR transgenic mice in which all the B cells are specific for hen egg lysozyme.	Gift from Prof. Carola Vinuesa Lab. Originated from (Goodnow et al. 1988).
Igh^{g2A10}	BCR light chain knock-in mice in which all B cells express the light chain of 2A10 antibody that binds to PfCSP.	Generated by Ozgene Pty Ltd (WA, Australia). Originated from (McNamara et al. 2020).
$Zeb2^{flox}$	Single loxP sites flank exon 7 of mouse Zeb2 gene.	Purchased from Riken institute, Japan. Originated from (Nishizaki et al. 2014).
$CD23^{Cre}$	Fcer2a (CD23) gene was replaced by Cre recombinase. All mature B cells in this mouse strain are supposed to express Cre.	Gift from Prof. Carola Vinuesa Lab.
SMARTA	TCR transgenic mice in which $CD4^+$ T cell are specific for the H2-I-Ab-restricted lymphocytic choriomeningitis virus GP61-80 epitope.	Originated from (Oxenius et al. 1998).

OT-II	TCR transgenic mice in which $CD4^+$ T cell are specific for the $Ovalbumin_{323-339}$ epitope.	Originated from (Barnden et al. 1998).
PbT-II	TCR transgenic mice in which $CD4^+$ T cell are specific for the P. *berghei* HSP90 antigen.	Gift from Prof. Bill Heath Lab. Originated from (Enders et al. 2021).
CST-II	TCR transgenic mice in which $CD4^+$ T cell are specific for the PfCSP NANP central repeat epitope.	Generated by Phenogenomic Targeting Facility.

2.5 Parasite

P. berghei parasites engineered to express various antigens were maintained by serial passage through *Anopheles stephensi* mosquitoes. *P. chabaudi* parasites (strain AS) were maintained by serial passage through mouse blood stage infection.

P. berghei parasite strains

Name	*Phenotype*	*Source*
PbCSP	Wild type *P. berghei* ANKA strain	From Leiden University.
PfCSP	Replacement of *P. berghei* CSP with PfCSP. Generated on an ANKA background.	(Espinosa et al. 2017)
PbCSP/PfCSP	Co-expressing *P. berghei* and P. *falciparum* CSP.	(Longley et al. 2015)
$PfCSP\text{-}OVA^{mcherry}$	Replacement of *P. berghei* CSP with PfCSP and express OVA protein and mCherry under same construct.	Generated by this study.
$PbCSP\text{-}PfCSPOVA^{flag}$	Genetically engineered *P. berghei* co-expresses PbCSP and $PfCSP\text{-}OVA_{260-385}$-Flag fusion protein which is under the CSP promoter integrated to chr12 non-coding locus.	Generated by this study.

PfCSP-PbCSPOVA^{flag}	Genetically engineered *P. berghei* co-expresses PfCSP and PbCSP-OVA$_{260-385}$-Flag fusion protein which is under the CSP promoter integrated to chr12 non-coding locus.	Generated by this study.

2.6 Human sample experiments

PBMC and plasma isolation

Blood from human were collected in BD Vacutainer Blood Collection Tubes. After centrifugation (400 g, 20°C, 5 min), plasma was collected and stored in -80°C for further analysis. Blood cells or buffy coats were diluted in PBS and gently loaded onto the Ficoll-Paque Plus (GE Healthcare) at the volume ratio of 1:1, followed by density gradient centrifugation (450 g, 20°C, 20 min, no brake). PBMC were then aspirated and resuspended in cold PBS for further experiment, or cryopreserved for later usage.

Antigen induced marker assay (AIM assay)

Cryopreserved PBMC were thawed, washed and counted. 5×10^5 PBMCs were resuspended in 200 µL complete RPMI media (3% FBS for AIM assay) and cultured in a 96-well flat-bottom plate for 18h in the presence of 20 µg/mL recombinant HBVSA, tetanus toxin, measles or 1 µg/mL SARS-CoV-2 Prot_S. No antigen was added to control wells. At least 6 wells were seeded (3 antigen treated wells + 3 medium only wells) for each PBMC sample. FACS was performed and the replicates for each sample were merged for downstream analysis.

In vitro survival and proliferation assays

For *in vitro* apoptosis assay, FACS purified $CD4^+$ T cells were resuspended in complete RPMI media and cultured for 3 days, followed by Annexin V and zombie aqua staining by FACS. For *in vitro* proliferation assay, FACS purified $CD4^+$ T cells were labelled by 5 µM CFSE for 5 min, washed and seeded on 96-well U-bottom plate with T cell activation Dynabeads at the ratio of 3:1 (cell number:bead number) to culture for 2.5 days, followed by FACS analysis.

T cell stimulation assay

To evaluate the effect of TCR stimulation on CXCR3 and CCR6 expression by cTfh cells, sorted cTfh cell subsets (2×10^4/well) were stimulated with plate-bound αCD3 (5 μg/mL) and αCD28 (2 μg/mL) or rested for 18h in complete RPMI media. CXCR3 and CCR6 expression by cTfh cell subsets were analysed by FACS.

Tonsiliar cell preparation for scRNA-seq

Tonsils were cut into fine pieces and smashed against a 70μm cell strainer. The tonsillar mononuclear cells were then purified by Ficoll-Paque Plus density gradient separation. After Fc blocking, tonsillar mononuclear cells were incubated in antibody cocktail on ice for 30 min followed by flow cytometry sorting for total B cells, GC B cells and memory B cells. The sorted cells were then incubated with Totalseq-C antibodies (1:50 dilution) on ice for 30 min, washed 3 times, counted and mixed at 1:1:1 ratio. 1×10^4 mixed cells were loaded onto each lane of the 10X Chromium platform. Library preparation was completed by Biomedical Research Facility (BRF) at the JCSMR following the recommended protocols. Libraries were sequenced using the NovaSeq6000 (Illumina). The 10X Cell Ranger package was used to process transcript, CITE-seq and VDJ libraries for downstream analysis.

2.7 Mouse experiments

Blood feeding

To generate sporozoites, the mice were infected by P. *berghei* infected RBC and when their parasitemia reaches 3% by Giemsa staining, they are anesthetized by IP injecting with ketamine/xylazine and placed on a mosquito cage to feed the mosquitos for 30min. The same procedures were repeated on the following day.

Recombinant protein in alum immunization

50 μg ovalbumin (OVA) or NP-OVA was emulsified in alum (volume ratio 1:1) and injected through IP for a dose of 200 μL per mouse. The spleens were collected on day 7 after immunisation or otherwise indicated in the thesis. For rCSP immunization, 10^4 Igh^{g2A10} B cells were transferred into MD4 mice followed by rCSP in alum immunization. 30ug rCSP was

emulsified in alum (1:1 volume ratio) and injected through IP for a dose of 200 μL per mouse and the spleens were analyzed on day10 post immunization.

Attenuated sporozoite immunization and Plasmodium infection

10^4 Igh^{g2A10} B cells, 10^5 OT-II cells, PbT-II cells or CST-II cells were transferred into WT, MD4 or CD28KO mice as specified in the thesis followed by irradiated sporozoites immunization. Sporozoites with indicated strains were dissected from the salivary glands of the mosquitos and resuspend in DPBS, followed by gamma irradiation using a MultiRad 225 biological irradiator. The irradiated sporozoites were counted and IV injected into the mice at the concentration of $3\text{-}5\times10^4$ per mouse per 100 μL DPBS. 100ug anti-TGFβ antibody per mouse per dose was administered by IP injection. For *P. chabaudi* infection, mice were IV injected with 1×10^5 P. *chabaudi* infected red blood cells in DPBS.

2.8 Cellular and molecular experiments

In vitro differentiation for OT-II cells

Red blood cell lysed splenocytes from WT mice were left untreated (for differentiating iTh0/1/2/17) or pre-treated by 1 μg/mL lipopolysaccharide (LPS) for 24h in the complete RPMI media. 5×10^5 per well LPS pre-treated splenocytes were co-cultured with FACS purified OT-II cells at the ratio of 50:1 in the presence of 1 μg/mL $OVA_{323\text{-}339}$ peptide and indicated cytokines for 72h to differentiate iTfh1/2/17 cells. No cytokines were added for differentiating Th0 cells. Neutralizing antibodies anti-IL-4, anti-IFN-γ and anti-TGF-β (BioxCell) were used at 10 μg/mL. Cytokines were purchased from PeproTech. For adoptive transfer of *in vitro* differentiated OT-II cells and OVA immunizations, $CD44^+$ OT-II cells cultured under iTh0 or iTfh1/2/17 conditions were FACS-purified and 5×10^4 cells were transferred into each CD28KO recipient mice, followed by OVA or NP-OVA in alum immunisations.

Cas9 RNP preparation and nucleofection

Anti-PE MACS beads (Miltenyi) enriched B220-PE^+ GFP^+ splenic mouse B cells were resuspended in complete RPMI media at 5 million cells per mL and stimulated by 10ug/mL anti-CD40 antibody for 24h. After wash once by DPBS, 1 million pre-activated B cells were resuspended in RNP solution for each nucleofection. RNP solution (for each nucleofection) was

prepared by mixing 250 nmol target gene sgRNA+ 50 nmol eGFP sgRNA+ 50 nmol SpCas9 2NLS Nuclease in 50 µL mouse B cell nucleofection buffer at RT for 30min. The nucleofection was then performed by AMAXA Nucleofector I device X-01 program. The nucleofected cells were rested in the cuvette at RT for 10min before being transferred out for downstream applications.

Mouse B cell culture

MACS enriched naïve mouse B cells or Cas9 RNP transfected B cells were cultured for 4 days in 200 µL complete RPMI media per well of 96-well plate in the presence of 3 µg/ml anti-CD40 antibody w/wo 50 ng/mL IL-21, 5ng/mL human TGFβ and 10ug/mL anti-TGFβ. Media is changed on day2 by gently aspirating 180 µL old media and adding 180 µL new media (changing media, especially the renewal of IL-21, is essential for successful ABC differentiation).

Sanger sequencing to calculate indel frequency

Genomic DNA were extracted from untreated or Cas9 RNP nucleofected cells (24h after transfection) according to manufacturer's instruction. The ~500bp DNA sequence flanking the expected cutting site were amplified by PCR and Sanger sequenced using another inner primer. The unedited and Cas9 edited sequences were then analyzed by the Synthego ICE tool for indel frequency calculation.

Quantitative RT-PCR

Total RNA was extracted from sorted T cell subsets using Qiagen RNeasy Micro Kit and reverse-transcribed to cDNA using iScript. RT-PCR was performed with SYBR Green PCR Master Mix with specific primers. The reaction of qPCR was performed by Applied Biosystem HT7900 according to the following protocol: 95 °C for 2 min, followed by 40 cycles of 95 °C for 10 sec, a specific annealing temperature for 10 sec, and 72 °C for 15 sec. Relative gene expression was calculated by 2(−Delta Delta CT) method using GAPDH or Ubc as an endogenous control (Albershardt et al. 2012).

ELISA for detecting antibody titer

For detecting anti-HBVSA antibody titer (total binding), a commercialized ELISA kit was used according to manufacturer's recommended protocol. Plasma was diluted 10 times and incubated with HBVSA pre-coated ELISA plate for 30 min under 37°C. Then HBVSA-HRP was added for another 30 min incubation, followed by 5 washes and substrate solution was used to determine the OD450 value. The antibody titer was calculated according to the standard curve generated by the standard with a known antibody titer. For detecting anti-HBVSA antibody titer (total IgG), plasma was diluted 10 times and incubated with HBVSA pre-coated ELISA plate for 30 min under 37°C. Then the plate was washed 3 times, and added by anti-human IgG-HRP (1:60000 dilution) for 30 min incubation under 37°C. The plate was then washed 5 times, and substrate solution was used to determine the OD450 value. For detecting the anti-NP antibody titer, mouse serum was diluted 2,000 times and incubated with NP2-BSA or NP23-BSA pre-coated ELISA plate for 1h at RT, followed by 3 washes and incubated with anti-mouse total IgG-HRP antibody for 1h at RT. The plate was then washed 5 times, and TMB chromogen solution was used to determine the OD405 value with 0.1% SDS as the stop solution.

RNA extraction and bulk RNA-seq

For human cTfh cell study, 0.5-1 million naïve, T_{CM}, T_{EM}, Tfh_{CM}, and Tfh_{EM} cells were sorted. For ABC study, 2×10^3 to 1×10^5 sorted B cells per sample were sorted. Cells were spin at 800 g for 8 min and total RNA was extracted using RNeasy Micro Kit following the recommended protocol. The RNA concentration and RNA integrity number (RIN) were measured by Bioanalyzer RNA 6000 pico assay. Total RNA samples with RIN ≥6.0 and concentration ≥30pmol/μL were submitted to BRF at the JCSMR for downstream process. Library preparation and sequencing was completed using NEBNext® Single Cell/Low Input RNA Library Prep Kit and NextSeq 550 (Illumina) following recommended protocol.

scRNA-seq for mouse Tfh cells

MD4 mice were left untreated or IV transferred with 10^4 Igh^{g2A10} B cells, followed by untreated, rCSP in alum or SPZs immunization. The splenocytes were collected on day10 after the immunization, and splenic $CD4^+$ T cells were enriched by MACS kit. Enriched $CD4^+$ T cells were stained by antibody cocktails containing flow cytometric and hashtag antibodies on ice for 30min, washed once and $CD44^+$ $PD\text{-}1^+$ $CXCR5^+$ Tfh cells were sorted and counted. Sorted

Tfh cells from different mice were then pooled for 10X scRNA-seq. 1×10^4 mixed cells were loaded onto each lane of the 10X Chromium platform. Library preparation was completed by Biomedical Research Facility (BRF) at the JCSMR following the recommended protocols. Libraries were sequenced using the NovaSeq6000 (Illumina). The 10X Cell Ranger package was used to process transcript, CITE-seq and VDJ libraries for downstream analysis.

Generation of CST-II mice

The method to generated CST-II TCR transgenic mice was reported previously for generating OT-II mice (Barnden et al. 1998). The original TCR alpha and beta chain sequences were recovered from scRNA-seq analysis and ligated in to the pES4 and p3A9cbTCR plasmids respectively by Genescript. The resulting plasmids were expanded by E.coli culture, extracted by miniprep and digested overnight at 37°C by ClaI-NotI (for pES4) ApaI-NotI (for p3A9cbTCR). The digested plasmids were run on a 1% TAE agarose gel (with 1:20000 GelRed nucleic acid gel stain) and the larger bands were excised for DNA purification. After determining the DNA concentrations by Nanodrop, the DNA sequences of alpha and beta chain were mixed, diluted and microinjected into the mouse embryos. The off springs were genotyped to check the integration of the TCR transgenes and one founder was produced. The expression of expected TCR beta chain was also confirmed by flow cytometry analysis and this founder can also pass on the transgene to the next generation. The microinjection and mouse transgenic mouse screening was performed by the Phenogenomic Targeting Facility, ANU.

Generation of transgenic parasite

The method to generate transgenic parasite was reported previously (Janse et al. 2006). The designed DNA sequences encoding PfCSP-OVA-flag and PbCSP-OVA-flag fusion proteins were ligated into the Pb268 plasmid (Singer and Frischknecht 2021) which contains the homologous arms of the P. *berghei* chromosome 12, CSP promoter (followed by designed sequences) and DHFR gene (pyrimethamine resistance) under eEF1 promoter. The resulting plasmids were expanded in E. coli culture, miniprepped and digested overnight at 37°C by PvuI twice. The digested DNA was cleaned up by a miniprep column for future transfection use. To transfect the *P. berghei* parasite, the mice were infected by injecting 10^5 *P. berghei* infected RBC and when parasitemia reaches >5 %, the mice were anesthetized by IP injecting with ketamine/xylazine to collect the whole blood in heparin tubes through cardiac puncture. Per 1

mL blood was then diluted into 200 mL RPMI media with 20% FBS and cultured at 37°C for 24h. Blood smear was performed to confirm the transition of ring stage parasites to schizonts. Then the schizonts were enriched by running the cultured RBC through a LS column with a 26.5-gauge needle to slow down the flow speed. After counting, schizonts were resuspended in human T cell transfection buffer containing 5ug digested DNA at 10^7 per 100ul per transfection using the U-33 program.

The transfected schizonts were immediately IV injected into WT mice fed by pyrimethamine water. On day7-10 the parasitemia was analyzed, and if it reaches >1%, the infected RBC was then cloned by diluted into 0.8 parasite per 100 μL DPBS per mouse and injected into 10-20 mice. The clones that expanded after 7-10 days were further validated by PCR (gDNA was extracted by infected RBC) to check the integration to the chromosome 12 and the presence of the expected CSP fusion protein.

2.9 Flow cytometry, imaging, bioinformatic analysis and statistics

Flow Cytometry

For phenotyping MWS patients and healthy controls, 100 μL of fresh Na heparin anti-coagulated peripheral blood was stained with monoclonal antibodies (mAb) according to the manufacturers' directions, incubated for 15 min at RT, lysed with Optilyse C (Beckman Coulter) for 10 min at RT, washed and resuspended in 250 μL 0.5% paraformaldehyde in PBS for further flow cytometry analysis. For mouse studies, ACK lysis buffer treated splenic cells, blood cells, bone marrow cells or cultured B cells were Fc blocked on ice for 10min and top up by the same volume of 2× concentrated antibody cocktail for another 1h on ice, followed by one wash and flow cytometry analysis. For staining of intracellular cytokines, cells were stimulated with PMA and ionomycin (500 ng/mL in the presence of GolgiPlus and GolgiStop for 4 h at 37°C. After surface staining, cells were permeabilised using Cytofix/Cytoperm. Antibodies specific to cytokines were incubated with cells for 30 min at 4°C. For intranuclear staining, surface staining was performed followed by fix/perm by Foxp3 kit and stained for nuclear proteins under room temperature for 60 min. Flow cytometry was performed on a FACS analyser (Fortessa X-20, BD). For imaging flow cytometry analysis, Igh^{g2A10} B cells or WT B cells were co-cultured with sorted GFP PfCSP SPZs in complete RPMI for 1h, followed by staining on ice for 30 min and analyzed by the Amnis imaging flow cytometry.

Bulk RNA-seq data analysis

The generated pair-end reads were processed online under the Galaxy project according to a standardised pipeline (Jalili et al. 2020). The count files were analysed according to a published pipeline (Law et al. 2016) for cpm normalization, MDS plot generation and differentially expressed genes calculation (low count genes were removed by filterByExpr). The heatmap was visualized by HemI. Gene set enrichment analysis (GSEA) was performed by fgsea to calculate GSEA *p* value and normalized enrichment score. Original R codes are attached.

10x single-cell RNA-seq analysis

For human cTfh cell study, the processed Seurat object was downloaded from GSE152522 and loaded into Seurat package (Hao et al. 2021). Unsupervised clustering was then performed to extract CXCR5-expression cTfh clusters. Then Tfh1 and Tfh17 signature scores were calculated for each cell by AddModuleScore function based on the signature gene sets for Tfh1 and Tfh17 derived from GSE123812. For TCR clonality analysis, cells sharing the same TCR alpha and beta chain CDR3 amino acid sequences were assigned to the same clonotype and the clonal abundance was calculated and ranked. Finally, the Tfh1 and Tfh17 signature scores for the abundant TCR clones (abundance ≥10) were extracted for statistical analysis and visualization.

For human tonsil study, transcripts and CITE-seq library outputs were loaded into Seurat package (Hao et al. 2021) for unwanted cell removal, clustering, annotation and visualization such as UMAP, violin plot, feature plot and dot plot. VDJ analysis were performed by IMGT/HighV-QUEST for SHM frequency and Ig isotype assignment. The FACS plots were made by Flowjo using the exported CITE-seq antibody expression values. Original R codes are attached.

Statistical analysis

For human result analysis, data were not assumed Gaussian distributed thus comparisons between two groups were performed by two-tailed Wilcoxon matched-pairs signed rank test and multiple comparisons were performed by Friedman test. For mouse result analysis, data were assumed Gaussian distributed thus comparisons between two groups were performed by

two-tailed unpaired *t*-test and multiple comparisons were performed by either one-way or two-way ANOVA test as specified in this paper. Statistical analysis was performed by Prism 9.0 software (GraphPad). *P* values < 0.05 were considered significant.

Chapter3- Type 17 Follicular Helper T (Tfh17) Cells are Superior for Immunological Memory Maintenance

3.1 Introduction

Follicular helper T (Tfh) cells are the specialized $CD4^+$ T cell subset that localize within B cell follicle to assist germinal center (GC) formation, plasma cell differentiation and high-affinity antibody production (Crotty 2011, Vinuesa et al. 2016). Identified in the circulation and lymphoid organs post immune response (memory phase), $CXCR5^+$ memory Tfh cells rapidly differentiate into mature effector Tfh cells and accelerate antibody response upon antigen re-exposure (MacLeod et al. 2011, Weber et al. 2012, Hale et al. 2013, He et al. 2013, Sage et al. 2014, Yu et al. 2022).

CXCR5-expressing $CD4^+$ T cells circulating in human blood (cTfh) provide important subjects to investigate memory Tfh cells since they have egressed from the site of the immune response at secondary lymphoid organs and can differentiate into effector Tfh cells upon antigen re-exposure (Tsai and Yu 2014). Noticeably, cTfh cells are heterogenous and are often classified into subsets by distinct functional markers. For example, cTfh cells are classified into Tfh1 ($CXCR3^+CCR6^-$), Tfh2 ($CXCR3^-CCR6^-$) and Tfh17 ($CXCR3^-CCR6^+$) subsets based on the expression of chemokine receptors CXCR3 and CCR6 associated with Th1 and Th17 respectively. Tfh2 and Tfh17 cells were reported to demonstrate better B cell helper function than Tfh1 cells in culture (Morita et al. 2011). The increases in Tfh2 and Tfh17 frequencies in autoimmune diseases often correlated with excessive production of pathogenic autoantibodies (Akiyama et al. 2015, Zhao et al. 2018). On the other hand, infections such as HIV and malaria mainly induce the generation of Tfh1 cells (Obeng-Adjei et al. 2015, Niessl et al. 2020). In influenza vaccination, it is also the Tfh1 subset that correlates with the titers of protective antibodies (Bentebibel et al. 2013).

Besides the classification of cTfh into Tfh1, Tfh2 and Tfh17 subsets by the features of lineage polarization, cTfh cells are composed of $CCR7^{high}PD\text{-}1^{low}$ "central memory (CM)-like" (Tfh_{CM}) and $CCR7^{low}PD\text{-}1^{high}$ "effector memory (EM)-like" (Tfh_{EM}) subsets (He et al. 2013), the latter also containing a more active $ICOS^+$ population (Bentebibel et al. 2013, Spensieri et al. 2013, Herati et al. 2017). $CCR7^{high}$ $PD\text{-}1^{low}$ Tfh_{CM} cells are dominant in human blood cTfh cells whereas circulating $CCR7^{low}PD\text{-}1^{high}$ Tfh_{EM} cells are temporarily induced in immune response and generated from the precursor stage of Tfh differentiation at secondary lymphoid organs (He et al. 2013).

There is little knowledge on the difference of memory Tfh subsets either Tfh1, Tfh2 and Tfh17 or Tfh_{EM} and Tfh_{CM} in memory responses. In this study, we developed a method to induce

antigen-specific Tfh1, Tfh2 and Tfh17-like (iTfh1, iTfh2 and iTfh17) mouse cells *in vitro*. iTfh1, iTfh2 and iTfh17 cells showed comparable B-helper function after the adoptive transfer into recipient mice followed by immediate immunization. In contrast, if transferred cells experienced an extended period of resting before re-immunization, iTfh17 cells were superior to iTfh1 and iTfh2 cells in sustaining antibody responses. Human Tfh17 cells represented ~20% Tfh_{EM} cells but accounted for >50% Tfh_{CM} cells which transcriptionally and phenotypically resemble central memory $CD4^+$ T (T_{CM}) with better survival and proliferative potential than effector memory (T_{EM}) cells (Sallusto et al. 2004). In vaccine responses to hepatitis B virus (HBV), influenza, tetanus toxin or measles, the Tfh17 subset in vaccine-specific cTfh cells is preferentially maintained into memory phase and long-lived. Complementary results from mouse and human studies thus suggest Tfh17 cells are superior for memory maintenance, the ability to persist and to support humoral response upon antigen restimulation.

3.2 Results

An optimized method to differentiate mouse follicular helper T cells in vitro

Upon priming, naïve $CD4^+$ helper T (Th) cells differentiate into distinct subsets with specialised functions. The differentiation of Th subsets is driven not only by signals from T cell receptor (TCR) and co-stimulatory receptors but also critically dependent on specific cytokine milieu. By mimicking such conditions, robust methods have been developed for the *in vitro* differentiation of type 1 and type 2 Th (Th1 and Th2) cells, and more recently IL-17-producing Th (Th17) cells and regulatory T (Treg) cells, which greatly support research and applications of these Th subsets. However, current methods for *in vitro* Tfh differentiation are not optimal. Even in the best practice, merely 20% polarised cells showed the expression of CXCR5, the key Tfh functional marker (Lu et al. 2011). The Tfh differentiation is critically dependent on co-stimulatory signals, including CD28, CD40L, ICOS and OX40 (Crotty 2011), which can be upregulated on APCs by LPS treatment (Xu et al. 2008). A conventional method for *in vitro* Tfh differentiation is to stimulate TCR transgenic naïve $CD4^+$ T cells with MHC II-specific peptides in the presence of IL-6 and IL-21 (Lu et al. 2011). We optimised this method by treating splenocytes with 1 µg/mL LPS for 24 hours before co-culturing with naïve OT-II cells with ovalbumin peptide 323-339 ($OVA_{323-339}$), IL-6 and IL-21. After the co-culture for 3 days, the pre-treatment of splenocytes as APCs with LPS significantly increased the CXCR5 expression on OT-II cells from ~5% to more than 30% (**Figure 3.1A, B**).

Our second attempt was to increase APC:T cell ratio since high-dose of antigens has been shown to promote Tfh differentiation (Baumjohann et al. 2013). By maintaining the consistent number of splenocytes and gradually reducing the number of OT-II cells, APC:T cell ratios elevated from 5:1 (the conventional method) to 50:1 and 500:1. Correspondingly, we observed more than two-fold increase of the CXCR5 expression, which showed a positive correlation with APC:T cell ratio. Using LPS-treated splenocytes and at an APC:T cell ratio of 500:1, over 70% OT-II cells displayed the $CXCR5^+$ $PD\text{-}1^+$ phenotype (**Figure 3.1A, B**). By testing the OVA peptide ranging from 0.1 to 1 μg/mL (high APC:T cell ratio at 500:1), or from 1 to 10 μg/mL (low APC:T cell ratio at 5:1), we found peptide concentrations showed negligible impact for Tfh cell differentiation (**Figure 3.1C**), suggesting that signalling other than TCR, presumably co-stimulatory signalling is the limiting factor for Tfh differentiation in the APC:T cell co-culture system.

Despite of the utility of anti-IL-4, anti-IFN-γ and anti-TGF-β antibodies in reported methods for *in vitro* Tfh differentiation (Nurieva et al. 2008, Lu et al. 2011), we found that the addition of such neutralising antibodies didn't enhance, or even inhibited Tfh polarisation (**Figure 3.1D**), which can be explained by the plasticity between Tfh and other Th subsets (Cannons et al. 2013).

Taken together, by stimulating APCs with LPS and setting the APC:T cell ratio higher than 50:1, phenotypic $CXCR5^+PD\text{-}1^+$ Tfh-like cells could be efficiently induced *in vitro*. To validate the application of this method on other TCR transgenic $CD4^+$T cells, we examined *in vitro* Tfh differentiation using naïve SMARTA cells with transgenic TCR specific recognising lymphocytic choriomeningitis virus (LCMV) glycoprotein residues 61-80 ($GP_{61\text{-}80}$), and obtained comparable results (**Figure 3.1E**).

By comparing with *in vitro* induced Tfh-like cells using with the conventional method (Conv-Tfh), we validated the optimised method (Optm-Tfh) promoted the expression of Bcl6 and CD40L but not ICOS (**Figure 3.1F, G**). Notably, several studies reported that high expression of ICOS was not always a unique marker for Tfh cells, as shown in *in vitro* culture and in mice with chronic LCMV infection (Crawford et al. 2014). Nevertheless, the Optm-Tfh method significantly enhanced the secretion of IL-21 (**Figure 3.1H**), the signature Tfh cytokine required to support GC response and plasma cell differentiation.

Finally, we assessed the B-helper function of *in vitro* generated Tfh cells by the Optm-Tfh method. We tested the Optm-Tfh methods of 5:1 and 50:1 but not the 500:1 condition due to

the low number of Tfh cells generated from the low frequency of OT-II seeding of the latter. OT-II $CD4^+$ T cells polarised in indicated conditions were treated with mitomycin-c to inhibit proliferation and then co-cultured for 6 days with isolated $B220^+$ cells that had been primed with LPS. OT-II cells from the Th0 culture condition hardly supported the survival of B cells while cells from Tfh polarisation conditions did so (**Figure 3.1I**). The Optm-Tfh method, especially APC:T cell ratio at 50:1, induced Tfh cells with superior B-helper capability, shown by several fold increase of plasma cell differentiation (**Figure 3.1J**) and antibody secretion (**Figure 3.1K**).

In conclusion, we present an optimised method to induce the *in vitro* Tfh differentiation of mouse TCR transgenic $CD4^+$ T cells. By stimulating APCs with LPS and setting the APC:T cell ratio higher than 50:1, the Optm-Tfh method efficiently induced Tfh-like cells that expressed higher levels of Tfh markers and possessed superior B-helper function. We believe this straightforward and reliable method would greatly support Tfh research and allow to test the applications that require a large amount of pure Tfh cells (e.g. testing whether increasing Tfh cell number can enhance the antibody response during vaccination).

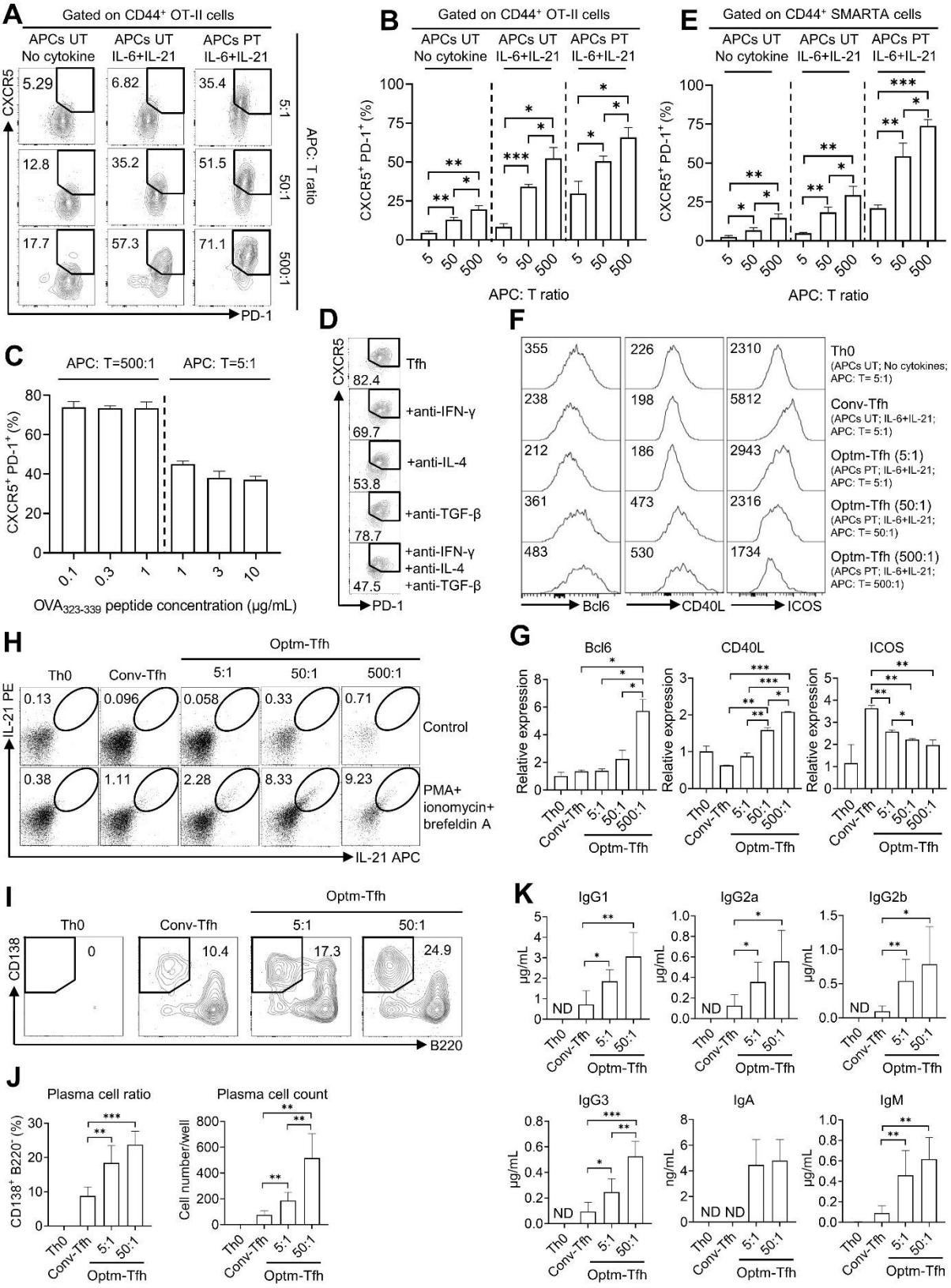

A
Gated on CD44+ OT-II cells
APCs UT No cytokine
APCs UT IL-6+IL-21
APCs PT IL-6+IL-21
CXCR5
PD-1
APC: T ratio
5:1
50:1
500:1
5.29
6.82
35.4
12.8
35.2
51.5
17.7
57.3
71.1
B
Gated on CD44+ OT-II cells
CXCR5+ PD-1+ (%)
APC: T ratio
E
Gated on CD44+ SMARTA cells
C
APC: T=500:1
APC: T=5:1
OVA323-339 peptide concentration (μg/mL)
D
CXCR5
PD-1
Tfh
+anti-IFN-γ
+anti-IL-4
+anti-TGF-β
+anti-IFN-γ +anti-IL-4 +anti-TGF-β
82.4
69.7
53.8
78.7
47.5
F
Th0 (APCs UT; No cytokines; APC: T= 5:1)
Conv-Tfh (APCs UT; IL-6+IL-21; APC: T= 5:1)
Optm-Tfh (5:1) (APCs PT; IL-6+IL-21; APC: T= 5:1)
Optm-Tfh (50:1) (APCs PT; IL-6+IL-21; APC: T= 50:1)
Optm-Tfh (500:1) (APCs PT; IL-6+IL-21; APC: T= 500:1)
355
226
2310
238
198
5812
212
186
2943
361
473
2316
483
530
1734
Bcl6
CD40L
ICOS
H
Optm-Tfh
Th0
Conv-Tfh
5:1
50:1
500:1
IL-21 PE
IL-21 APC
Control
PMA+ ionomycin+ brefeldin A
0.13
0.096
0.058
0.33
0.71
0.38
1.11
2.28
8.33
9.23
G
Bcl6
CD40L
ICOS
Relative expression
I
CD138
B220
0
10.4
17.3
24.9
J
Plasma cell ratio
CD138+ B220- (%)
Plasma cell count
Cell number/well
K
IgG1
IgG2a
IgG2b
IgG3
IgA
IgM
μg/mL
ng/mL
ND

Figure 3.1 Optimization and characterisation for in vitro differentiation of mouse Tfh cells

(A-E) APCs were either untreated (UT) or pre-treated (PT) by 1 μg/mL LPS for 24 h and co-cultured with OT-II or SMARTA cells as indicated condition for 3 days. Representative FACS plots (A) and statistics (B, E) showing the expression for CXCR5 and PD-1 on CD44+ OT-II or SMARTA cells. The effects of Tfh differentiation by antigen dose (C) or specific neutralising antibodies (D) were tested using OT-II cells. The expression of Tfh surface makers (F, G) and signature cytokine (H) in OT-II cells were quantified. (I-K) After *in vitro* differentiation for 3 days, total OT-II cells from indicated conditions were isolated by FACS purification and treated with 50 μg/mL mitomycin-c for 30 min before being co-cultured at 1:1 ratio with purified B cells that had been pre-activated with 1 μg/mL LPS for 24 h. Cells were cultured for 6 days in the presence of 1 μg/mL $OVA_{323-339}$ peptide. Representative FACS plots (I) and statistics (J) showing plasma cell differentiation. The secretion of indicated antibodies in culture supernatants were measured by bead-based ELISA (K). Results are representative of at least two independent experiments. Detailed methods are described in supplementary information. ND: not detectable. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

The in vitro differentiation of induced Tfh1, Tfh2 and Tfh17-like (iTfh1, iTfh2, iTfh17) cells

The classification of Tfh1/2/17 cells based on the expression of CXCR3 and CCR6 markers was established by characterizing human blood memory Tfh cells (Morita et al. 2011). Such memory Tfh1/2/17 subsets in $CD44^{+}CXCR5^{+}$ Tfh cells are low in numbers (**Figure 3.2A**), thus limiting functional characterization. We modified the method above to polarize OT-II cells into Tfh1/2/17 (iTfh1/2/17) *in vitro*. In addition to IL-6 and IL-21 in the iTfh differentiation method (Gao et al. 2020), Th1 (IL-12, anti-TGF-β, anti-IL-4), Th17 (TGF-β, anti-IFN-γ, anti-IL-4) and Th2 (IL-4, anti-IFN-γ, anti-TGF-β) polarization cultures were adopted for iTfh1/2/17 induction with lower IL-12, TGF-β or IL-4 concentrations for iTfh1/2/17 induction than those used for canonical iTh1, iTh17 or iTh2 induction (details in the method) (Nurieva et al. 2008, Lu et al. 2011, Read et al. 2019) (**Figure 3.2B**). iTfh1/2/17 cells expressed higher Tfh-defining markers CXCR5, PD-1 and BCL6 than iTh0/1/2/17 cells (**Figure 3.2C, D**). The BCL6 expression in iTfh1/2/17 cells was lower than $CD44^{+}CXCR5^{high}PD\text{-}1^{high}$ GC-Tfh cells in immunized mice **(Figure 3.2E, F)**. Tfh differentiation undergoes a step-by-step process, showing the generation of precursor Tfh cells expressing intermediate levels of BCL6 and subsequent maturation of GC-Tfh cells with the highest BCL6 expression (Crotty 2011, Vinuesa et al. 2016). Memory cTfh cells largely originate from precursor Tfh cells and express low levels of BCL6 (He et al. 2013). Given that iTfh1/2/17 cells expressed BCL6 lower than that in GC-Tfh cells and resemble precursor Tfh cells, iTfh1/2/17 cells are suitable to study the function of memory cTfh cells. Importantly, iTfh1/2/17 cells differentially expressed transcription factors T-bet, GATA3 and RORγt (**Figure 3.2G, H**) and chemokine receptors CXCR3 and CCR6 (**Figure 3.2I, J**), as their counterpart human Tfh1/2/17 cells (Morita et al. 2011).

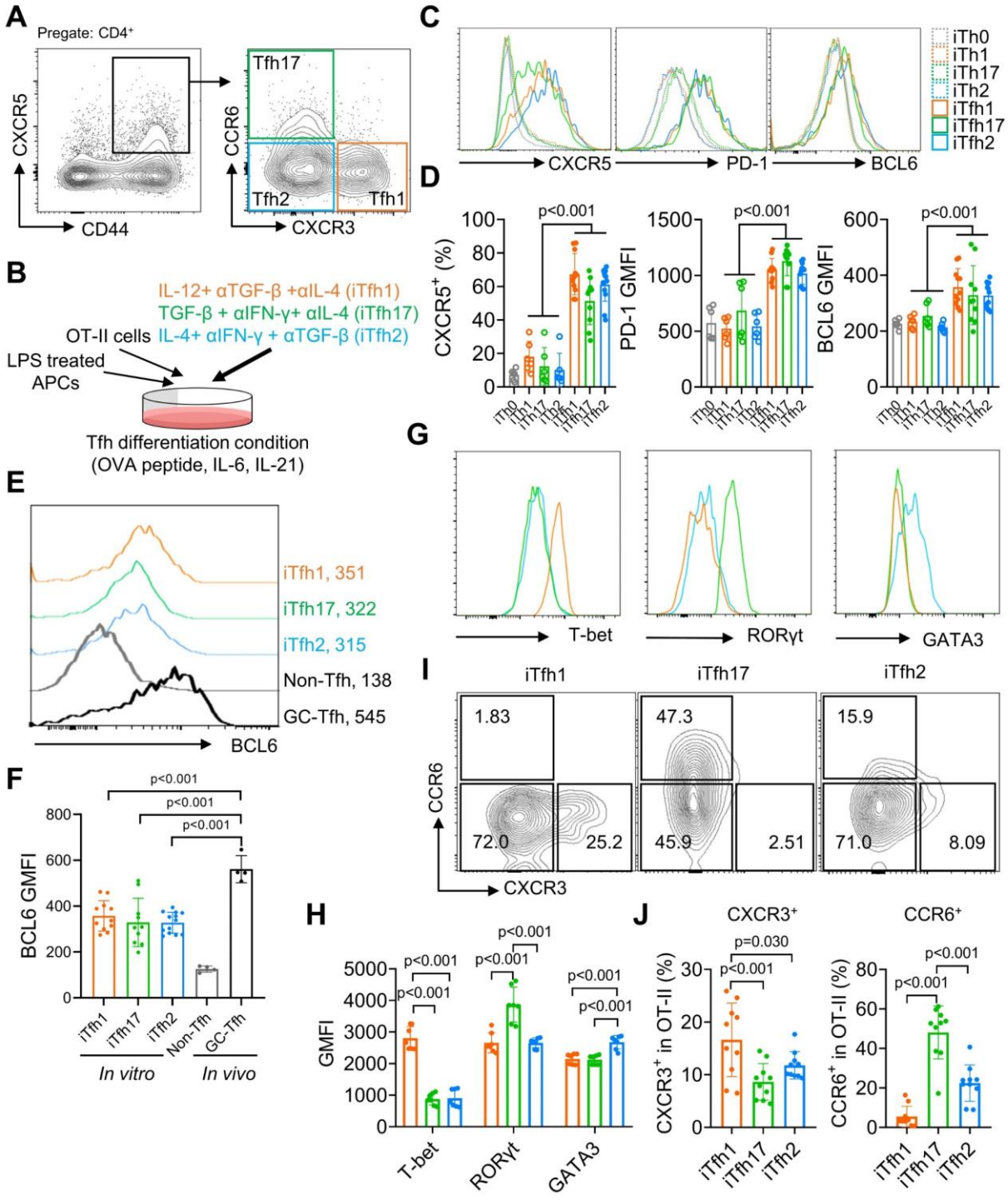

A
Pregate: CD4+
CXCR5
CD44
Tfh17
Tfh2
Tfh1
CCR6
CXCR3
B
IL-12+ αTGF-β +αIL-4 (iTfh1)
TGF-β + αIFN-γ+ αIL-4 (iTfh17)
IL-4+ αIFN-γ + αTGF-β (iTfh2)
OT-II cells
LPS treated APCs
Tfh differentiation condition
(OVA peptide, IL-6, IL-21)
C
iTh0
iTh1
iTh17
iTh2
iTfh1
iTfh17
iTfh2
CXCR5
PD-1
BCL6
D
CXCR5+ (%)
PD-1 GMFI
BCL6 GMFI
p<0.001
E
iTfh1, 351
iTfh17, 322
iTfh2, 315
Non-Tfh, 138
GC-Tfh, 545
BCL6
F
BCL6 GMFI
iTfh1
iTfh17
iTfh2
Non-Tfh
GC-Tfh
In vitro
In vivo
G
T-bet
RORγt
GATA3
I
iTfh1
iTfh17
iTfh2
1.83
47.3
15.9
72.0
25.2
45.9
2.51
71.0
8.09
CCR6
CXCR3
H
GMFI
T-bet
RORγt
GATA3
J
CXCR3+
CCR6+
CXCR3+ in OT-II (%)
CCR6+ in OT-II (%)
p=0.030

Figure 3.2 The in vitro differentiation of induced Tfh1, Tfh2 and Tfh17-like (iTfh1, iTfh2, iTfh17) cells

(A) Splenocytes from WT mice were analyzed and representative FACS plot for Tfh1, Tfh17 and Tfh2 cells was shown. (B-J) OT-II cells were co-cultured with WT splenocytes as antigen-presenting cells (APCs) in the presence of OVA peptide, indicated cytokines and blocking antibodies for three days before phenotypic analysis. For GC-Tfh and non-Tfh cells, OT-II cells were transferred into congenic WT mice followed by OVA in alum immunization. On day7, the BCL6 expressions for splenic OT-II Non-Tfh ($CD44^{+}CXCR5^{-}PD\text{-}1^{-}$) and GC-Tfh cells ($CD44^{+}CXCR5^{high}PD\text{-}1^{high}$) were analysed and compared with iTfh1/2/17 cells. Experiment design (B), representative FACS plots (C) and statistics (D) for the expression of Tfh markers CXCR5, PD-1 and BCL6. Representative FACS plots (E) and statistics (F) comparing the expressions of BCL6 of indicated cell populations. The numbers in (E) indicate GMFI values. Representative FACS plots and statistics showing the expression of transcription factors T-bet, RORγt and GATA3 (G, H) and CXCR3 *vs* CCR6 expression (I, J). The *p* values were calculated by two-way ANOVA for (D) and one-way ANOVA for (F, H, J). The results in (D, H, J) were pooled from 3 independent experiments. The results in (F) were pooled from 2 independent experiments.

iTfh1/2/17 cells form Tfh cells after immunization and retain polarized Th1/2/17 features

To examine whether iTfh1/2/17 cells retain polarized phenotypes *in vivo*, we adoptively transferred each cell type individually into congenic CD28KO recipient mice, followed by the immunization of OVA in aluminium salt (OVA-Alum) (**Figure 3.3A**). After 7 days, iTfh1/2/17 cells showed the comparable ability of effector Tfh differentiation (**Figure 3.3B**) but maintain the distinction in CXCR3 and CCR6 expression aligning with their progenitors (**Figure 3.3C, D**). These results suggest that *in vitro* generated iTfh1/2/17 cells can be used to investigate the function of Tfh1/2/17 cells.

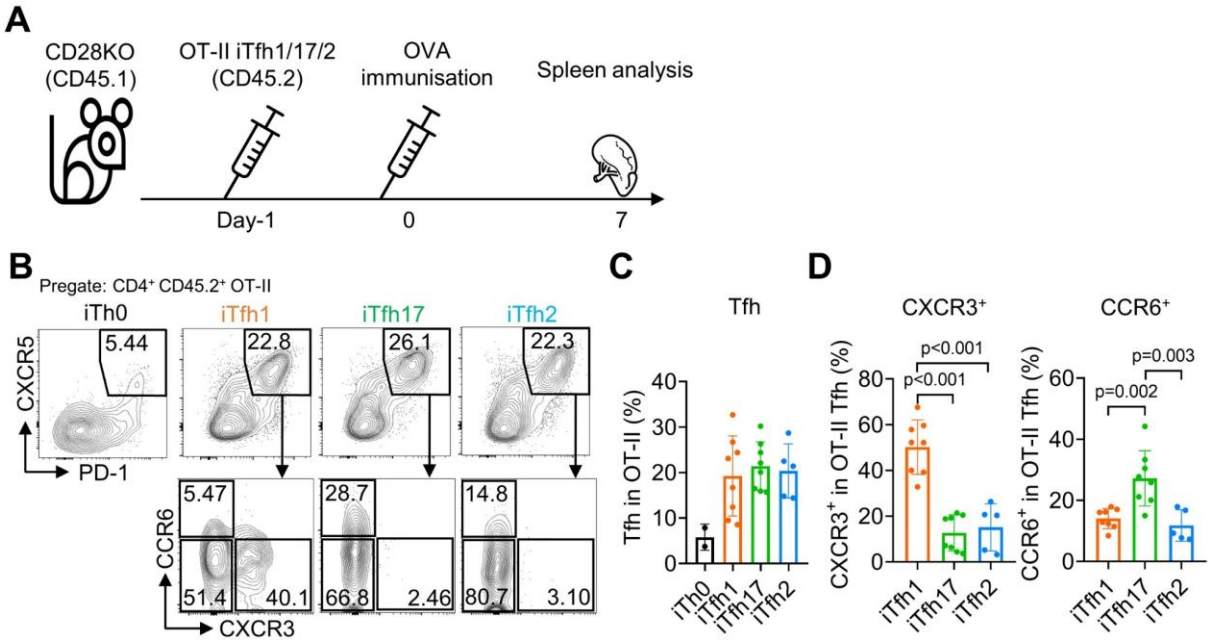


Figure 3.3 iTfh1/2/17 cells form Tfh cells after immunization and retain polarized Th1/2/17 features

(A-D) 5×104 cultured OT-II iTh0, iTfh1, iTfh2 and iTfh17 cells were FACS-purified and separately transferred into CD28KO recipients, followed by OVA immunization in alum. The spleens were collected on day7 post-immunization for FACS analysis. Experiment design (A), representative FACS plot for Tfh percentage in OT-II cells (B), statistics of Tfh percentage in OT-II cells (C) and statistics of CXCR3/CCR6^{+} percentage in OT-II Tfh cells (D). The *p* values were calculated by one-way ANOVA for (D). The results in (C, D) were pooled from 2 independent experiments.

iTfh17 cells are superior in memory maintenance after adoptive transfer

To compare the function of Tfh1, Tfh2 and Tfh17 *in vivo*, we adoptively transferred each of OT-II naïve T cell-derived iTfh1, iTfh2 or iTfh17 cells into congenic CD28KO recipient mice. T cells in CD28KO mice are defective in co-stimulation and unable to generate endogenous Tfh cells so antibody responses in CD28KO mice are dependent on transferred iTfh cells. After adoptive cell transfer, mice were immunized with OVA-Alum at day 0 (early immunization) or day 35 (late immunization) with the latter condition mimicking memory maintenance of Tfh1, Tfh2 and Tfh17 for extended *in vivo* resting **(Figure 3.4A)**.

On day 7 post the early immunization, OT-II iTfh1, iTfh2 and iTfh17 cells demonstrated largely comparable Tfh differentiation and the function in supporting the generation of germinal centre B (B_{GC}) cells and antibody-secreting B (B_{ASC}) cells (**Figure 3.4B-G**). A larger magnitude of B_{GC} cells supported by iTfh2 cells might be explained by the function of IL-4 in enhancing B_{GC} generation (Gaya et al. 2018). However, the outcome was very different in the scheme of late immunization whereby iTfh cells had experienced memory maintenance. After resting *in vivo* for 35 days, iTfh17 cells produced more than 2-fold of mature effector Tfh cells than those by iTfh1 or iTfh2 cells **(Figure 3.4B, E)**, accompanied by ~2-fold increase in B_{GC} differentiation in mice that had received iTfh17 cells than those had received iTfh1 or iTfh2 cells (**Figure 3.4C, F**). Although the trend of an increase in iTfh17-supported B_{ASC} differentiation didn't reach statistical significance (**Figure 3.4G**), iTfh17 cells were superior to iTfh1 and iTfh2 cells in supporting the production of anti-NP IgG antibodies, after resting *in vivo* for 35 days before the immunization with antigen 4-Hydroxy-3-nitrophenyl (NP)-OVA (**Figure 3.4H**). Collectively, iTfh17 cells are superior to iTfh1 or iTfh2 cells in helping B cells, but only in the scheme with extended *in vivo* resting.

The selective advantage of iTfh17 cells in supporting Tfh differentiation and humoral immunity after an extended *in vivo* resting followed by immunization suggests that Tfh17 cells may outperform Tfh1 or Tfh2 cells to sustain Tfh memory. In resting mice, Tfh17 cells expressed higher CCR7 than Tfh1 or Tfh2 cells (**Figure 3.4I**), despite the highest expression of activation marker CD44 by Tfh17 cells (**Figure 3.4J**). $CCR7^+$ marks T_{CM} cells that circulate in the blood and secondary lymphoid tissues and have longer survival and better proliferative capacity than $CCR7^-$ T_{EM} cells (Sallusto et al. 2004, Bouneaud et al. 2005). We thus hypothesized that Tfh17 cells might carry certain features of T_{CM} cells suitable for memory maintenance.

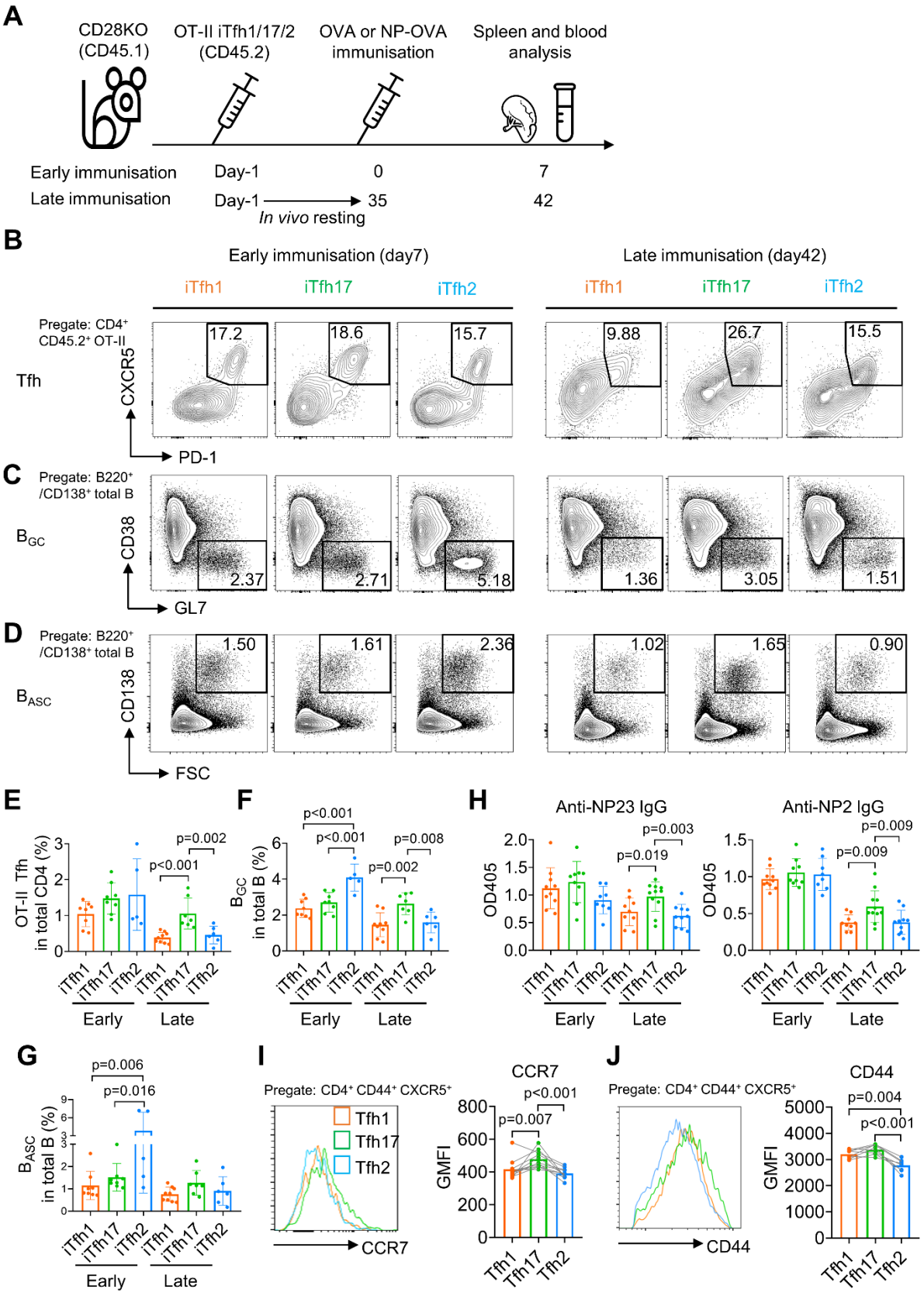

A
CD28KO (CD45.1)
OT-II iTfh1/17/2 (CD45.2)
OVA or NP-OVA immunisation
Spleen and blood analysis
Early immunisation
Day-1
0
7
Late immunisation
Day-1
35
42
In vivo resting
B
Early immunisation (day7)
Late immunisation (day42)
iTfh1
iTfh17
iTfh2
Pregate: CD4+ CD45.2+ OT-II
Tfh
CXCR5
PD-1
17.2
18.6
15.7
9.88
26.7
15.5
C
Pregate: B220+ /CD138+ total B
BGC
CD38
GL7
2.37
2.71
5.18
1.36
3.05
1.51
D
Pregate: B220+ /CD138+ total B
BASC
CD138
FSC
1.50
1.61
2.36
1.02
1.65
0.90
E
OT-II Tfh in total CD4 (%)
p=0.002
p<0.001
iTfh1
iTfh17
iTfh2
Early
Late
F
BGC in total B (%)
p<0.001
p<0.001
p=0.008
p=0.002
H
Anti-NP23 IgG
Anti-NP2 IgG
OD405
p=0.003
p=0.019
p=0.009
p=0.009
G
BASC in total B (%)
p=0.006
p=0.016
I
Pregate: CD4+ CD44+ CXCR5+
Tfh1
Tfh17
Tfh2
CCR7
GMFI
p<0.001
p=0.007
J
Pregate: CD4+ CD44+ CXCR5+
CD44
GMFI
p=0.004
p<0.001

Figure 3.4. iTfh17 cells are superior in memory maintenance after adoptive transfer
(A-H) 5×10^4 FACS-purified OT-II iTfh1, iTfh17 or iTfh2 cells were separately transferred to CD28KO recipients. The early immunization group was immunized by OVA or NP-OVA in alum one day after the adoptive cell transfer. The late immunization group was immunized by the same antigens 35 days after the adoptive cell transfer. Spleens or serum were collected on day 7 after the immunization. Experiment design (A), representative FACS plots (B, C, D) and statistics (E, F, G) showing the percentages of Tfh cells in OT-II cells, the percentages of B_{GC} in total B cells and the percentages of B_{ASC} in total B cells. For antibody titers, statistic (H) showing OD405 values of anti-NP2 and anti-NP23 total IgG. (I-J) Tfh1/2/17 cells from mouse splenocytes were analyzed for CCR7 and CD44 expression. Representative FACS plots (I). and statistics (J) showing the expressions of CCR7 and CD44. The *p* values were calculated by one-way ANOVA. The results in (E, F, G, H, I, J) were both pooled from 2 independent experiments.

Human Tfh_{CM} and Tfh_{EM} subsets phenotypically and functionally resemble T_{CM} and T_{EM} subsets respectively

Following the observation that iTfh17 cells showed a unique advantage in memory maintenance, we set to characterize the function of human Tfh17 cells in maintaining Tfh memory. We previously reported that human cTfh cells are composed of $CCR7^{high}PD\text{-}1^{low}$ T_{CM}-like and $CCR7^{low}PD\text{-}1^{high}$ T_{EM}-like subsets with the latter indicating an active Tfh differentiation (He et al. 2013), but their function has not been formally compared. We first investigated the relationship between the two cTfh subsets and corresponding $CD4^+$ T_{CM} and T_{EM} subsets by transcriptomic analysis using RNA sequencing (RNA-seq) (**Figure 3.5A**). As shown in an unsupervised multidimensional scaling (MDS) plot, Tfh_{CM} cells closely cluster with T_{CM} cells, and Tfh_{EM} cells fall between T_{CM} and T_{EM} cells on the major dimension1, implying that Tfh_{EM} cells are distinct from Tfh_{CM} and T_{CM} cells but also have effector programs different from T_{EM} cells (**Figure 3.5B**). Previous studies reported that cTfh cells predominantly show $CCR7^+$ CM phenotype (Chevalier et al. 2011). Our transcriptomic analysis indeed suggests that Tfh_{CM} cells acquire a quiescent state hardly distinguishable from T_{CM} cells. In contrast, Tfh_{EM} cells' transcriptomes are clearly separated from those of T_{EM} cells, presumably caused by the divergent effector function of Tfh cells as compared to other effector Th1, Th2 or Th17 cells. In line with this, the top 50 differentially expressed genes (DEG) indicate effector genes such as *ZEB2* and *TBX21* (Omilusik et al. 2015) were highly expressed in T_{EM} cells, intermediate levels in Tfh_{EM} cells, and lowest in Tfh_{CM} and T_{CM} cells (**Figure 3.5C**). In top 50 hallmark gene sets identified by gene set enrichment analysis (GSEA) between Tfh_{EM} *vs* Tfh_{CM} cells or T_{EM} *vs* T_{CM} cells, 37 gene sets were significantly enriched by both comparisons (NES discrepancy > 2, **Figure 3.5D**), suggesting that the transcriptomic features and regulation between Tfh_{EM} *vs* Tfh_{CM} cells are overall similar to those between T_{EM} *vs* T_{CM}. Despite T_{EM} and Tfh_{Em} cells showing distinct transcriptomes and locating separately in the MDS plot (**Figure 3.5B**), the key gene sets that are related to common effector T cell function (activation, effector differentiation and cell cycle entry) were both positively enriched in comparisons between Tfh_{EM} *vs* Tfh_{CM} cells or T_{EM} *vs* T_{CM} cells (**Figure 3.5E**). Therefore, Tfh_{CM} and Tfh_{EM} cells resemble T_{CM} and T_{EM} cells respectively at the transcriptomic levels.

We next compared Tfh_{CM} and Tfh_{EM} subsets for survival and stimulation-induced proliferation in culture, which were applied to characterize the difference between T_{CM} and T_{EM} cells (Sallusto et al. 2004). In non-stimulation culture for 3 days, T_{CM} and Tfh_{CM} cells retained ~50% viability while T_{EM} and Tfh_{EM} cells showed poorer survival of ~30% (**Figure 3.5F, G**). To

measure the proliferative potential, all subsets were labeled with carboxyfluorescein succinimidyl ester (CFSE) and stimulated by anti-CD3/CD28 for 2.5 days. While the majority of T_{EM} or Tfh_{EM} cells underwent division once, most T_{CM} or Tfh_{CM} cells reached the second or third division, indicating a better proliferative potential (**Figure 3.5H, I**). Collectively, $CCR7^{high}PD\text{-}1^{low}$ Tfh_{CM} and $CCR7^{low}PD\text{-}1^{high}$ Tfh_{EM} subsets showed not only transcriptomic profiles resembling their counterpart T_{CM} and T_{EM} cells but also functional characteristics of survival and proliferative capacity (Sallusto et al. 2004, Bouneaud et al. 2005).

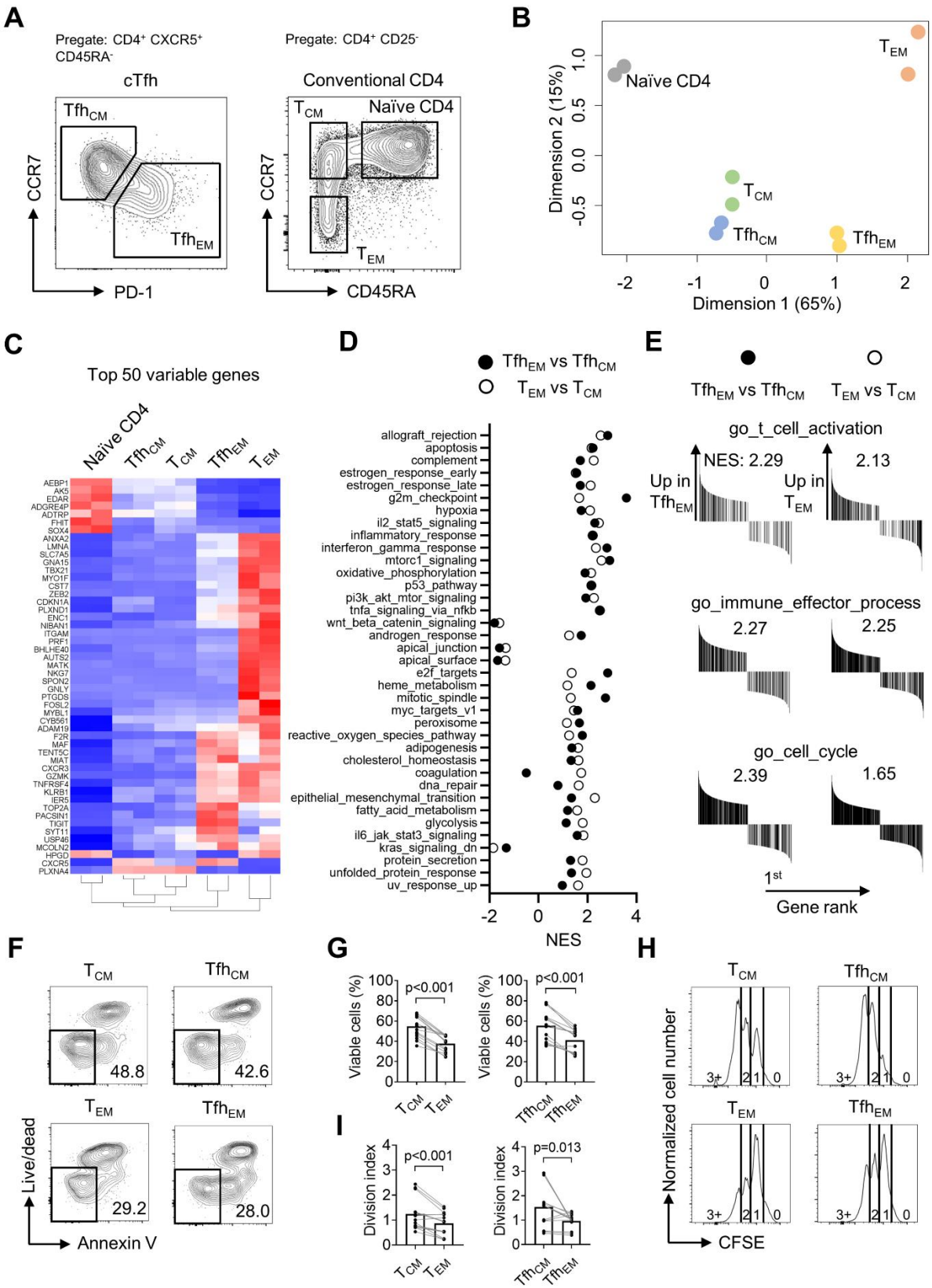

A
Pregate: CD4+ CXCR5+ CD45RA-
cTfh
TfhCM
TfhEM
CCR7
PD-1
Pregate: CD4+ CD25-
Conventional CD4
TCM
Naïve CD4
TEM
CCR7
CD45RA
B
TEM
Naïve CD4
TCM
TfhCM
TfhEM
Dimension 2 (15%)
Dimension 1 (65%)
C
Top 50 variable genes
Naïve CD4
TfhCM
TCM
TfhEM
TEM
D
TfhEM vs TfhCM
TEM vs TCM
allograft_rejection
apoptosis
complement
estrogen_response_early
estrogen_response_late
g2m_checkpoint
hypoxia
il2_stat5_signaling
inflammatory_response
interferon_gamma_response
mtorc1_signaling
oxidative_phosphorylation
p53_pathway
pi3k_akt_mtor_signaling
tnfa_signaling_via_nfkb
wnt_beta_catenin_signaling
androgen_response
apical_junction
apical_surface
e2f_targets
heme_metabolism
mitotic_spindle
myc_targets_v1
peroxisome
reactive_oxygen_species_pathway
adipogenesis
cholesterol_homeostasis
coagulation
dna_repair
epithelial_mesenchymal_transition
fatty_acid_metabolism
glycolysis
il6_jak_stat3_signaling
kras_signaling_dn
protein_secretion
unfolded_protein_response
uv_response_up
NES
E
TfhEM vs TfhCM
TEM vs TCM
go_t_cell_activation
NES: 2.29
2.13
Up in TfhEM
Up in TEM
go_immune_effector_process
2.27
2.25
go_cell_cycle
2.39
1.65
1st
Gene rank
F
TCM
TfhCM
48.8
42.6
TEM
TfhEM
29.2
28.0
Live/dead
Annexin V
G
Viable cells (%)
p<0.001
p<0.001
TCM
TEM
TfhCM
TfhEM
I
Division index
p<0.001
p=0.013
H
TCM
TfhCM
TEM
TfhEM
Normalized cell number
3+
2
1
0
CFSE

Figure 3.5 Human Tfh_{CM} and Tfh_{EM} subsets phenotypically and functionally resemble T_{CM} and T_{EM} subsets respectively

(A-E) Naïve, T_{CM}, T_{EM}, Tfh_{CM} and Tfh_{EM} cells were FACS-purified from PBMC of two healthy donors and bulk RNA-seq was performed for differentially expressed genes analysis and gene set enrichment analysis (GSEA). (A) Representative FACS plot showing the gating strategy for indicated subsets. (B) MDS plot showing sample distribution. (C) Heatmap of the top 50 variable genes normalized by z-score (scale: red: 0-2, white=0, blue: -2-0). (D) Summarized normalized enrichment score (NES) of significantly enriched (p <0.05, FDR < 0.25) hallmark gene sets by either Tfh_{EM} *vs* Tfh_{CM} or T_{EM} *vs* T_{CM}. (E) GSEA on selected gene sets were performed on Tfh_{EM} *vs* Tfh_{CM} and T_{EM} *vs* T_{CM} and the number indicates NES. (F-G) FACS-purified T_{CM}, T_{EM}, Tfh_{CM} and Tfh_{EM} cells were rested in complete RPMI for 3 days. Representative FACS plots (F) and statistics (G) showing the percentages of viable cells. (H-I) FACS-purified T_{CM}, T_{EM}, Tfh_{CM} and Tfh_{EM} cells were labelled with CFSE and stimulated by anti-CD3/CD28 for 2.5 days. Representative FACS plots (H) and statistics (I) showing the CFSE fluorescence intensity and the division index. The p values were calculated by Wilcoxon matched-pairs signed-rank test. The results in (G, I) were pooled from 5 healthy individuals with each conducted in 3 technical replicates.

Human Tfh_{CM} cells are enriched with the Tfh17 subset whereas Tfh_{EM} cells are enriched with the Tfh1 subset

From a cohort of healthy donors (N = 33), we analyzed $CCR7^{high}PD\text{-}1^{low}$ Tfh_{CM} and $CCR7^{low}PD\text{-}1^{high}$ Tfh_{EM} cells for the percentages of Tfh1/2/17 subsets based on CXCR3 and CCR6 expression. In agreement with the higher expression of CCR7 on mouse Tfh17 than that on Tfh1 or Tfh2 cells, human Tfh_{CM} cells were dominated by the Tfh17 subset (mean = 51.44%), followed by the Tfh2 subset (mean = 16.19%) and the Tfh1 subset (mean = 12.17%) (**Figure 3.6A**), whereas Tfh_{EM} cells were dominated by the Tfh1 subset (mean = 34.06%,) (**Figure 3.6B**). The population of Tfh cells that expresses both CXCR3 and CCR6 has been reported and also presented in our samples. Due to the fact that $CXCR3^{+}CCR6^{+}$ cTfh cells were fewer than Tfh1, Tfh2 or Tfh17 cells and their ontogeny remains to be fully revealed (Morita et al. 2011), we did not include this population in the following analyses. To avoid the influence of individual variation of Tfh1/2/17 polarization due to different histories of immune exposure and examine whether there is an intrinsic difference of Tfh1/2/17 frequencies between Tfh_{CM} or Tfh_{EM} cells, the percentages of Tfh1/2/17 cells in Tfh_{CM} cells were normalized to those in Tfh_{EM} cells in each individual, which demonstrated the highest Tfh_{CM}/Tfh_{EM} ratio for Tfh17 (mean = 4.18), an intermediate ratio for Tfh2 (mean = 3.38), and the lowest ratio for Tfh1 (mean = 0.75) (**Figure 3.6C**). The highest ratio (>> 1) for Tfh17 indicates Tfh_{CM} cells are highly enriched with the Tfh17 subset whereas the lowest ratio (< 1) for Tfh1 indicates that Tfh_{EM} cells are enriched with the Tfh1 subset. We also measured the expression of hallmark transcription factors TBX21, GATA3 and RORC and cytokines IFN-γ, IL-4 and IL-17A that are selectively expressed in Tfh1/2/17 cells respectively (Morita et al. 2011). These molecules for effector Th functions are abundantly expressed by T_{EM} cells but are downregulated in T_{CM} cells which enter into a resting state (Sallusto et al. 2004). Indeed, the expression of effector transcription factors and cytokines was consistently lower in Tfh_{CM} cells than those in Tfh_{EM} cells (**Figure 3.6D, F**). Notably, the ratios of expression (Tfh_{CM}/Tfh_{EM}) demonstrate modest reductions of 20-40% in Tfh17 related markers RORγt and IL-17A, in contrast to vast reductions of 70-80% in Tfh1 related markers T-bet and IFN-γ (**Figure 3.6E, G**). Such results of transcription factor and cytokine expression support the conclusion for an enrichment of the Tfh17 subset and a loss of the Tfh1 subset in Tfh_{CM} cells.

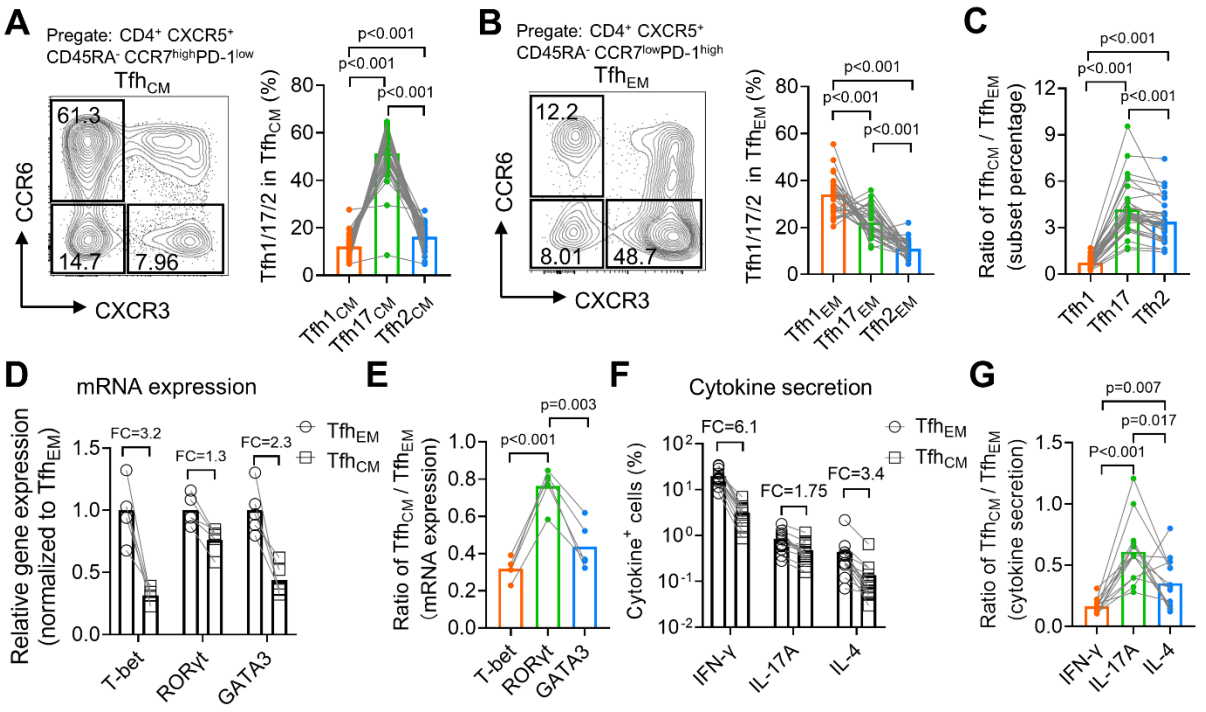


Figure 3.6. Human Tfh_{CM} cells are enriched with the Tfh17 subset whereas Tfh_{EM} cells are enriched with the Tfh1 subset

(A-C) Human PBMC samples from 33 healthy blood donors were analyzed. Representative FACS plots and statistics showing the percentages of Tfh1, Tfh2 and Tfh17 cells in Tfh_{CM} (A) or Tfh_{EM} (B) subsets. Tfh_{CM}/Tfh_{EM} ratios for Tfh1/2/17 in each individual were calculated (C). (D-E) FACS-purified Tfh_{EM} and Tfh_{CM} from 5 healthy individuals were analyzed for the expressions of indicated transcription factors by qPCR. The statistics for relative gene expression $2^{-\Delta\Delta Ct}$ (normalized to Tfh_{EM}) (D) and Tfh_{CM}/Tfh_{EM} ratios (E). (F-G) PBMC from 13 healthy individuals were analyzed for the secretions for indicated cytokines post PMA/ionomycin stimulation. The statistics for the percentages of cytokine^{+} cells (F) and the Tfh_{CM}/Tfh_{EM} ratios (G). FC: average fold change. The *p* values were calculated by Friedman test.

HBV antigen-specific Tfh17 cells are preferentially maintained in memory phase

Tfh_{EM} to Tfh_{CM} phenotype conversion occurs over the period of a few weeks when the antigen is absent (He et al. 2013). The enrichment of the Tfh17 subset in human Tfh_{CM} cells suggest the Tfh17 subset in Tfh_{EM} cells may persist longer than the Tfh1 or Tfh2 subsets. The phenomenon could also result from a biased Tfh_{CM} phenotype of Tfh17 cells generated even early in immune responses. To tease apart the cause, we next examined the phenotype of antigen-specific human cTfh cells over a period that expands both effector and memory phases after vaccination or infection.

Childhood HBV vaccination doesn't always provide life-long protection with a proportion of vaccinees with antibody titers at an undetectable level in adulthood (Bruce et al. 2016). HBV boosting vaccination is recommended for high-risk populations such as medical practitioners including medical students in China. A cohort of medical students (N = 38) with serum negative for anti-HBV surface antigen (HBVSA) antibody were recruited. Peripheral blood mononuclear cells (PBMCs) were collected at day 7 before and day 7 and 28 after the immunization (**Figure 3.7A**). Antigen-induced marker (AIM) assay was used to examine HBV vaccine-specific T cells by culturing PBMCs with HBVSA for 18h and detecting the PD-$L1^+OX40^+CD25^+$ cells as the antigen-specific population (**Figure 3.7B**). This method has been applied to characterize antigen-specific cTfh cells (Dan et al. 2016, Reiss et al. 2017). Such stimulation did not change the expression of CXCR3 and CCR6 on cTfh cells **(Figure 3.7C, D)**, indicating that AIM assays are suitable to characterize antigen-specific Tfh1/2/17 cells. As reported (Reiss et al. 2017), a background in AIM assays exists in a small proportion of samples whereby PD-$L1^+OX40^+CD25^+$ cells were detected in control cultures without antigen stimulation (**Figure 3.7E**). To specifically quantify antigen-specific response, we subtracted the value of antigen-stimulation culture by the background value from the control culture without antigen stimulation. Normalized values were then used to calculate the percentages of Tfh1, Tfh2 and Tfh17 cells (**Figure 3.7F**).

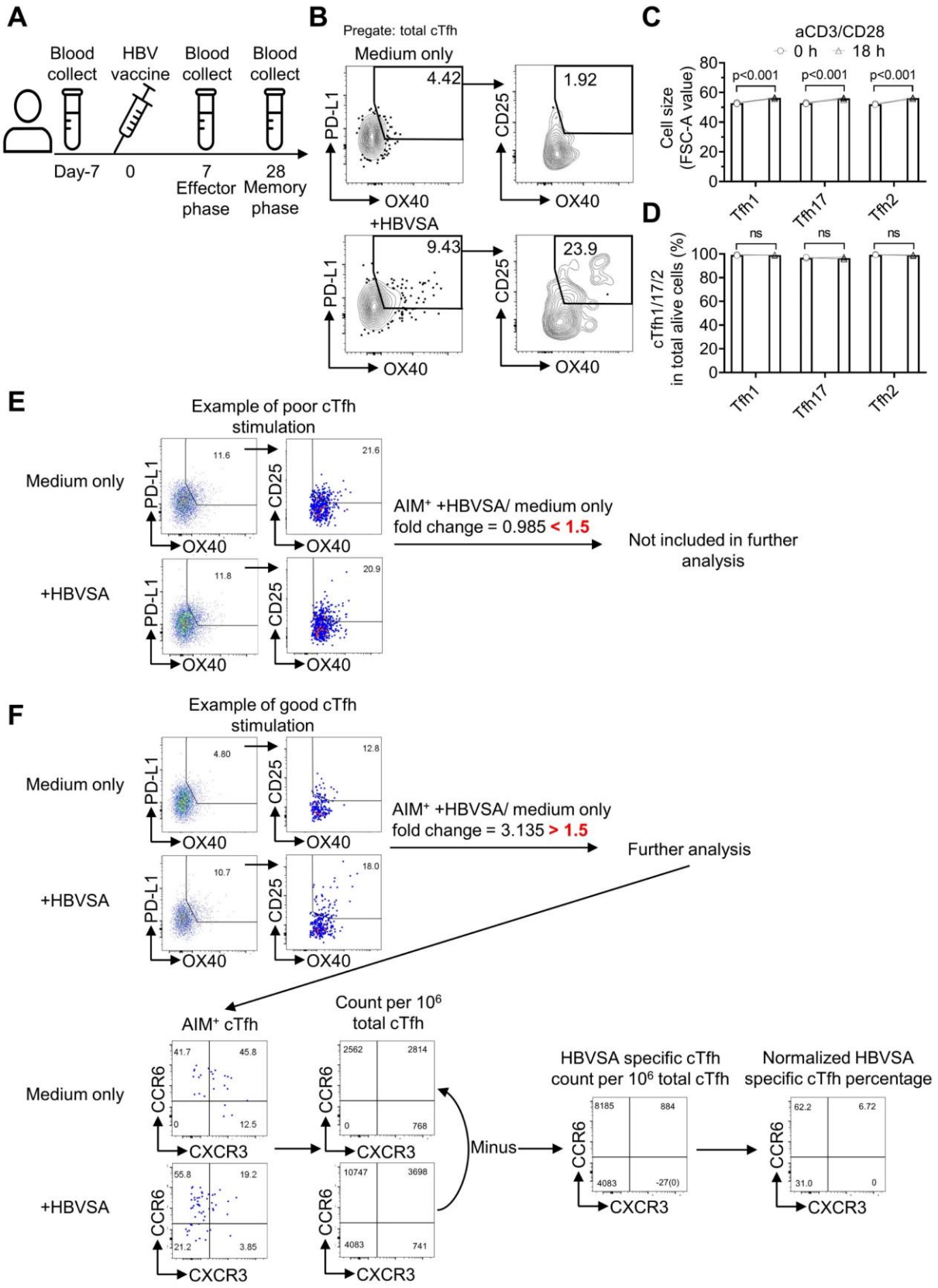

A
Blood collect
HBV vaccine
Blood collect
Blood collect
Day-7
0
7 Effector phase
28 Memory phase
B
Pregate: total cTfh
Medium only
4.42
1.92
PD-L1
CD25
OX40
+HBVSA
9.43
23.9
C
aCD3/CD28
0 h
18 h
p<0.001
Cell size (FSC-A value)
Tfh1
Tfh17
Tfh2
D
ns
cTfh1/17/2 in total alive cells (%)
E
Example of poor cTfh stimulation
11.6
21.6
11.8
20.9
AIM+ +HBVSA/ medium only
fold change = 0.985 < 1.5
Not included in further analysis
F
Example of good cTfh stimulation
4.80
12.8
10.7
18.0
AIM+ +HBVSA/ medium only
fold change = 3.135 > 1.5
Further analysis
AIM+ cTfh
Count per 10^6 total cTfh
CCR6
CXCR3
41.7
45.8
12.5
55.8
19.2
21.2
3.85
2562
2814
768
10747
3698
4083
741
Minus
HBVSA specific cTfh count per 10^6 total cTfh
8185
884
-27(0)
Normalized HBVSA specific cTfh percentage
62.2
6.72
31.0
0

Figure 3.7 AIM assays to identify antigen-specific cTfh cells

(A-B) Blood samples from HBV vaccinated healthy individuals (N = 38) were collected on indicated time points before/after HBV vaccination, PBMC were isolated and cultured with or without 20 μg/mL HBVSA for 18h, followed by FACS to analyse the phenotype of HBVSA-specific cTfh cells. Experiment design (A) and representative FACS plot (B) showing the gating strategy to detect HBVSA-specific cTfh cells by AIM assay. (C-D) FACS-purified Tfh1/2/17 cells from 5 individuals were stimulated by αCD3/CD28 for 18 hours and were analysed by FACS. Statistics (C) showing the cellular sizes before/after αCD3/CD28 stimulation. Statistics (D) showing the percentages of Tfh1/2/17 cells before/after stimulation. (E-F) Same experiment as in "(A-B)". Example (E) of one sample excluded from the analysis. Example (F) of one sample included in the analysis. The numbers of total cTfh cells were then normalized to 10^6 and the absolute counts of (AIM^+) PD-$L1^+OX40^+CD25^+$ cTfh cells were calculated, and the absolute counts of antigen-specific Tfh1, Tfh2, Tfh17 and Tfh1/17 were obtained. Any count values less than 0 were changed to 0 because cell count cannot be negative. At last, the normalized percentage was calculated based on the absolute count of antigen-specific Tfh1, Tfh2, Tfh17 and Tfh1/17. The *p* values were calculated by paired *t*-tests.

In all subjects negative for HBVSA antigen and anti-HBVSA antibody, the average percentage of the Tfh17 subset in HBVSA-specific memory cTfh cells was 49.31%, whereas the average percentage of the Tfh1 subset was 15.7% (**Figure 3.8A, B**). In alignment with the results on total Tfh_{CM} cells, HBVSA-specific Tfh_{CM} cells were also enriched with the Tfh17 subset.

To tease apart whether the Tfh17 enrichment was caused by the biased generation or better maintenance, we then analysed PBMC samples collected at day 7 and day 28 after vaccination. The cohort was divided into 4 groups based on vaccine responses measured by antibody titers: early responders (titer > 100 mIU at day7, titer = 1248$\pm$324.4 mIU at day 28, N = 11), late responders (titer < 100 mIU at day 7 and > 200 mIU at day 28, titer = 997.1$\pm$289.7 mIU/ml at day 28, N = 15), weak responders (50 mIU < titer < 200 mIU at day 28, titer = 107.9$\pm$37.2 mIU at day 28, N = 6) and non-responders (titer < 50 mIU on day 28, titer = 17.28$\pm$4.379 mIU at day 28, N =6) (**Figure 3.8C**). Tfh activation measured by a trend of increase in HBVSA-specific Tfh_{EM} cells (2.84-fold, day -7 v.s. day 7, P-value = 0.074, not reaching statistical significance) was observed only in early responders but no other groups (**Figure 3.8D, E**), suggesting individuals in this group had active Tfh response on day7 and were suitable to study the memory maintenance of Tfh1/2/17 cells.

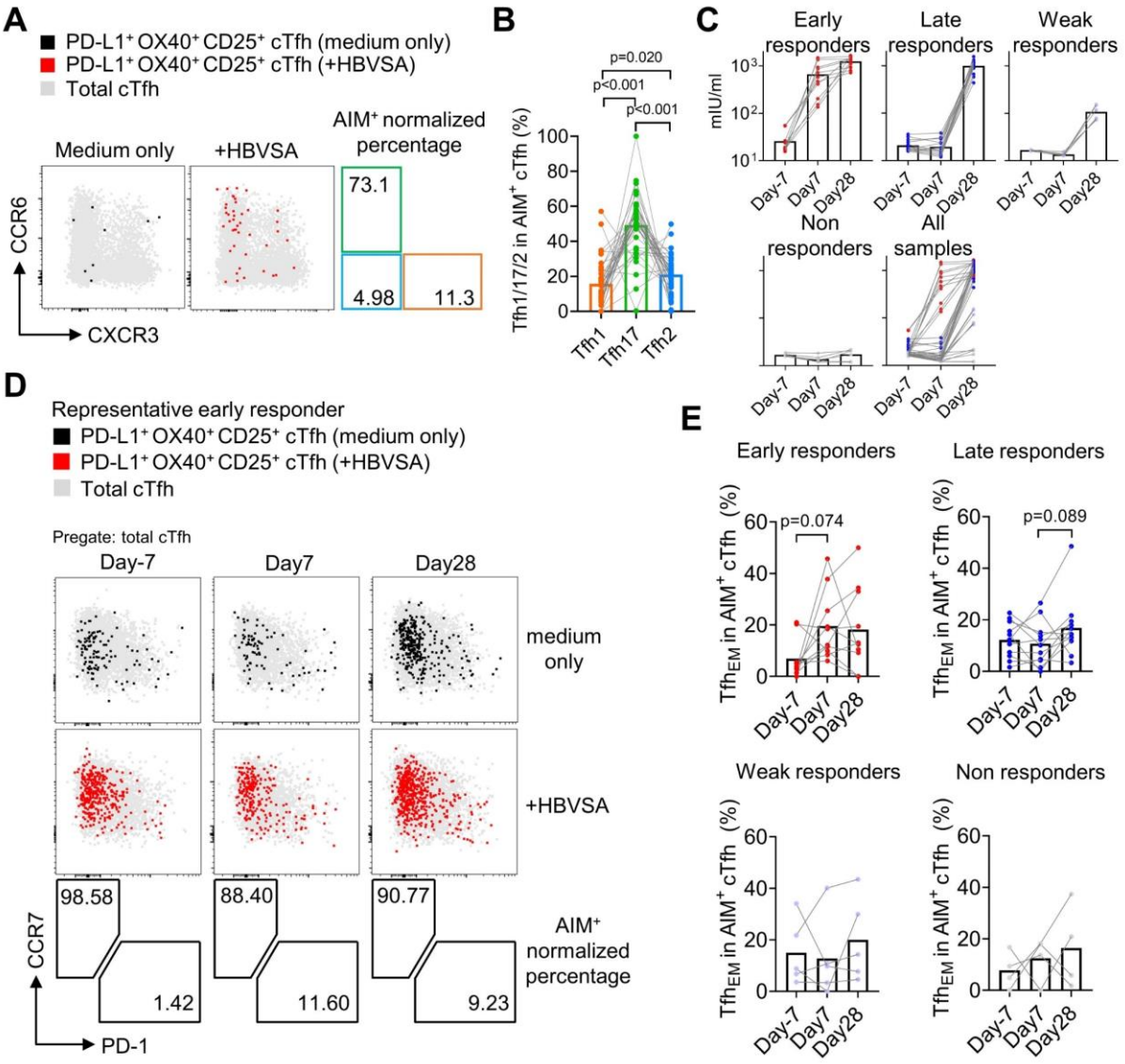


Figure 3.8 Early responders have detectable Tfh responses on day7
Blood samples from HBV vaccinated healthy individuals (N = 38) were collected on indicated time points before/after HBV vaccination, and serum was diluted 10 times to analyse the anti-HBVSA antibody titer by ELISA. PBMC were also isolated and cultured with or without 20 µg/mL HBVSA for 18h, followed by FACS to analyse the phenotype of HBVSA-specific cTfh cells. Representative FACS plot (A) and statistics (B) showing the percentage of Tfh1/2/17 cells in HBVSA-specific cTfh cells before vaccination. Classification (C) of 38 individuals into 4 groups was based on their anti-HBVSA antibody titers. Representative FACS plot of one early responder (D) showing the percentage of Tfh_{CM} and Tfh_{EM} in HBVSA-specific cTfh cells on indicated time points. The statistics (E) showing the percentages of HBVSA-specific Tfh_{EM} cells on indicated time points in 4 groups of responders. The p values were calculated by Friedman test.

We next focused on the kinetics of HBVSA-specific cTfh subsets in early responders. At day 7 post vaccination, the percentages of three subsets in HBVSA-specific cTfh cells ranged from 20% to 30% and showed no significant difference (**Figure 3.9A**). Of note, the percentages of the Tfh17 subset significantly increased from ~30% to ~50% from day 7 to 28 (*P*-value = 0.006). In contrast, the percentages of the Tfh1 subset dropped significantly from ~20% to ~10% (*P*-value = 0.020). The percentages of Tfh2 remained largely unchanged (**Figure 3.9B, C**). As a result, Tfh17 cells dominated HBVSA-specific cTfh cells in the memory phase of day 28 post vaccination **(Figure 3.9A)**. Therefore, the Tfh17 enrichment in Tfh_{CM} cells results from an advantage of Tfh17 cells in memory maintenance, rather than a biased phenotype to Tfh_{CM} cells. Significant changes in Tfh1 and Tfh17 percentages from day 7 to day 28 were selective in early responder group but not in three other groups (**Figure 3.9C**), and only observed in HBVSA-specific cTfh cells but not in total cTfh in early responders (**Figure 3.9D**), suggesting that the dynamic changes were specific to HBVSA-specific cTfh response.

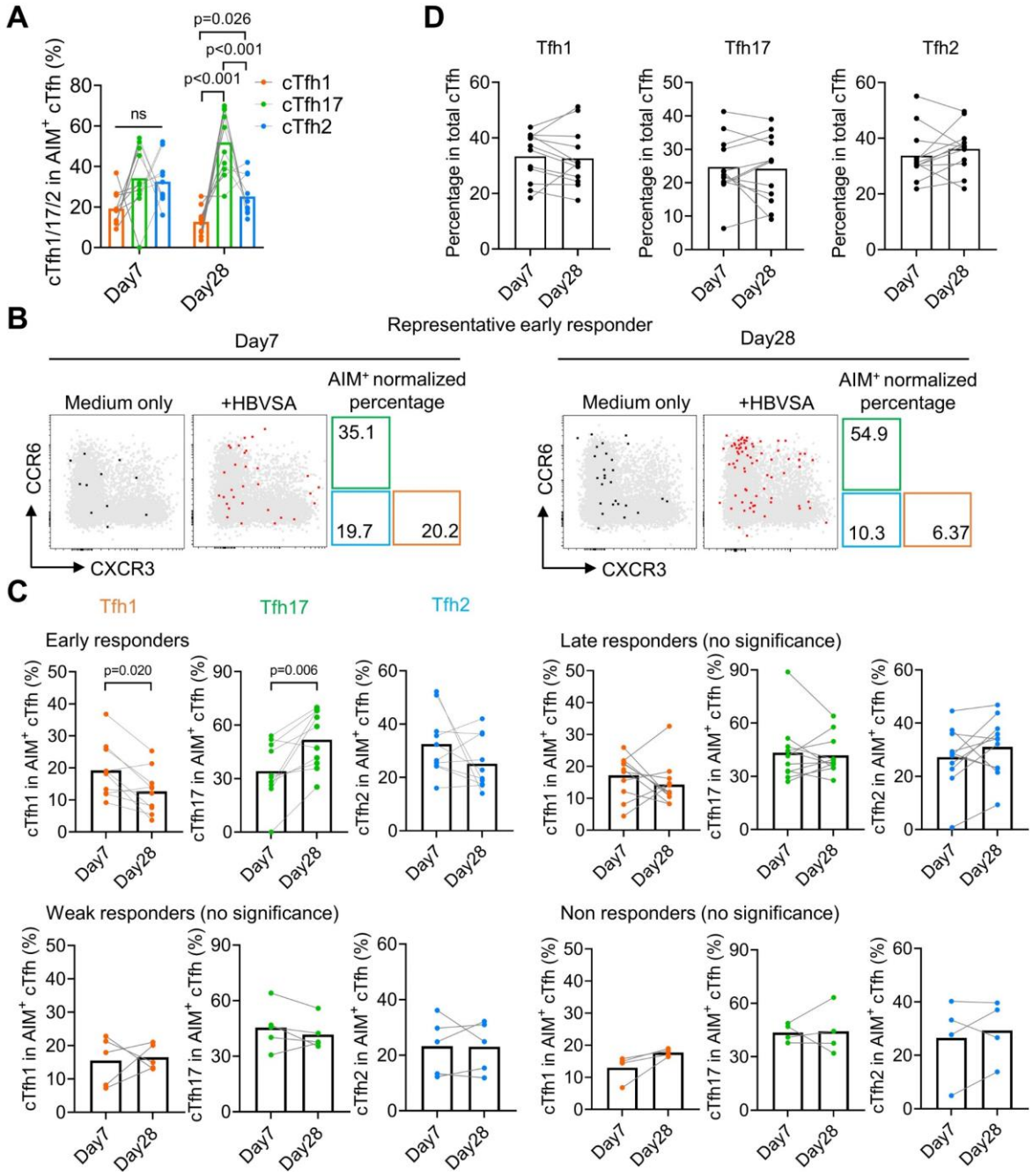


Figure 3.9 HBV antigen-specific Tfh17 cells are preferentially maintained in memory phase

Blood samples from HBV vaccinated healthy individuals (N = 38) were collected on indicated time points before/after HBV vaccination, PBMC were also isolated and cultured with or without 20 μg/mL HBVSA for 18h, followed by FACS to analyse the phenotype of HBVSA-specific cTfh cells. The statistics of early responders (A) showing the percentages of HBVSA-specific Tfh1/2/17 cells on day 7 and day 28 post HBV vaccination. Representative FACS plot (B) for an early responder showing the percentage of Tfh1/2/17 cells in HBVSA-specific cTfh cells on day7 and day28. Statistics (C) showing the percentage of Tfh1/2/17 cells in HBVSA-specific cTfh on day 7 and day 28 after the vaccination in all defined groups (N = 30, 8 samples with poor signals in AIM assay were excluded). The statistics of early responders (D) showing the percentages of Tfh1/2/17 in total cTfh on day 7 and day 28 post HBV vaccination. The p values were calculated by Friedman test for (A) and Wilcoxon matched-pairs signed-rank test for (C).

Influenza virus-specific cTfh cells show Tfh1 signatures in effector phase but Tfh17 signatures in memory phase

Single-cell RNA-seq (scRNA-seq) paired with TCR sequencing facilitates the characterization of the phenotype and function of antigen-specific T cell clones in an immune response. We took the advantage of this new technology to analyze the characteristics of influenza haemagglutinin (HA)-specific $CD4^+$ T clones in a published dataset of scRNA/TCR-seq from four healthy individuals with influenza vaccination (Meckiff et al. 2020). We compared HA-specific T cell clones before the vaccination (memory phase) and day 12 after the vaccination (effector phase) **(Figure 3.10A)**. HA-specific $CD4^+$ T cells before and after the vaccination were pooled to generate unsupervised clustering, in which CXCR5-expressing clusters 2-5 were enriched of cTfh cells, in which a total of twelve major $CD4^+$ T clones (clonal abundance $\geq$ 10) were identified **(Figure 3.10B)**.

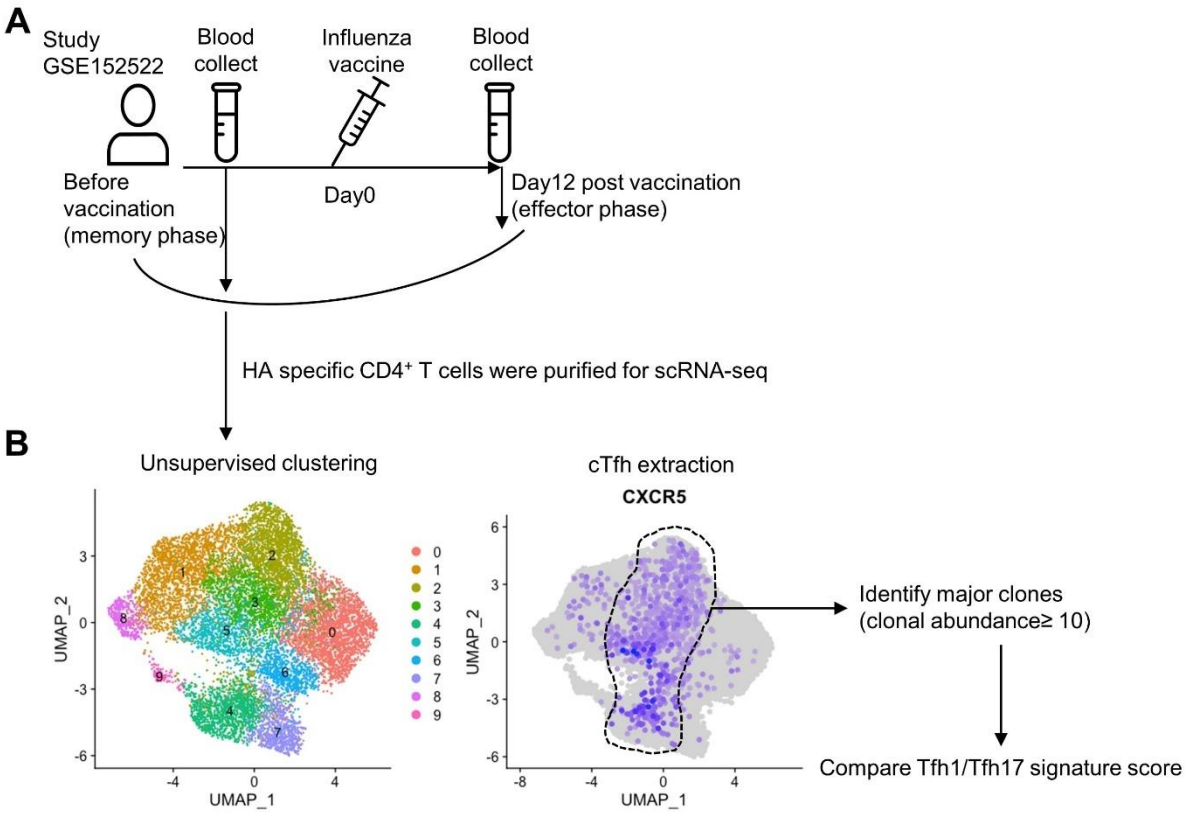


Figure 3.10 The identification of influenza virus-specific cTfh cells
The single-cell RNA-seq dataset (GSE152522, the experiment design (A)) was analyzed to identify CXCR5-expressing cTfh clusters (B), which contain 12 major clones with a total of 249 cells.

To investigate Tfh subsets-associated features in HA-specific clonal cTfh cells, we applied Tfh1 or Tfh17 signature gene sets derived from bulk RNA-seq for Tfh1/2/17 cells **(Figure 3.11A)** (Yost et al. 2019) in clonal cTfh cells. The scores of Tfh1 signature were higher in clonal cTfh cells in the effector phase than those in the memory phase (*P*-value = 0.041); by contrast, the scores of Tfh17 signature were lower in the effector phase than in the memory phase (*P*-value = 0.002) (**Figure 3.11B**). The divergence of Tfh1 and Tfh17 signatures between the effector and memory phases was consistent in individual donors **(Figure 3.11C)** and at the level measured by twelve major clones (**Figure 3.11D**). Therefore, we conclude that the advantage of Tfh17 cells in memory maintenance is consistently observed among different cohorts with different types of vaccines.

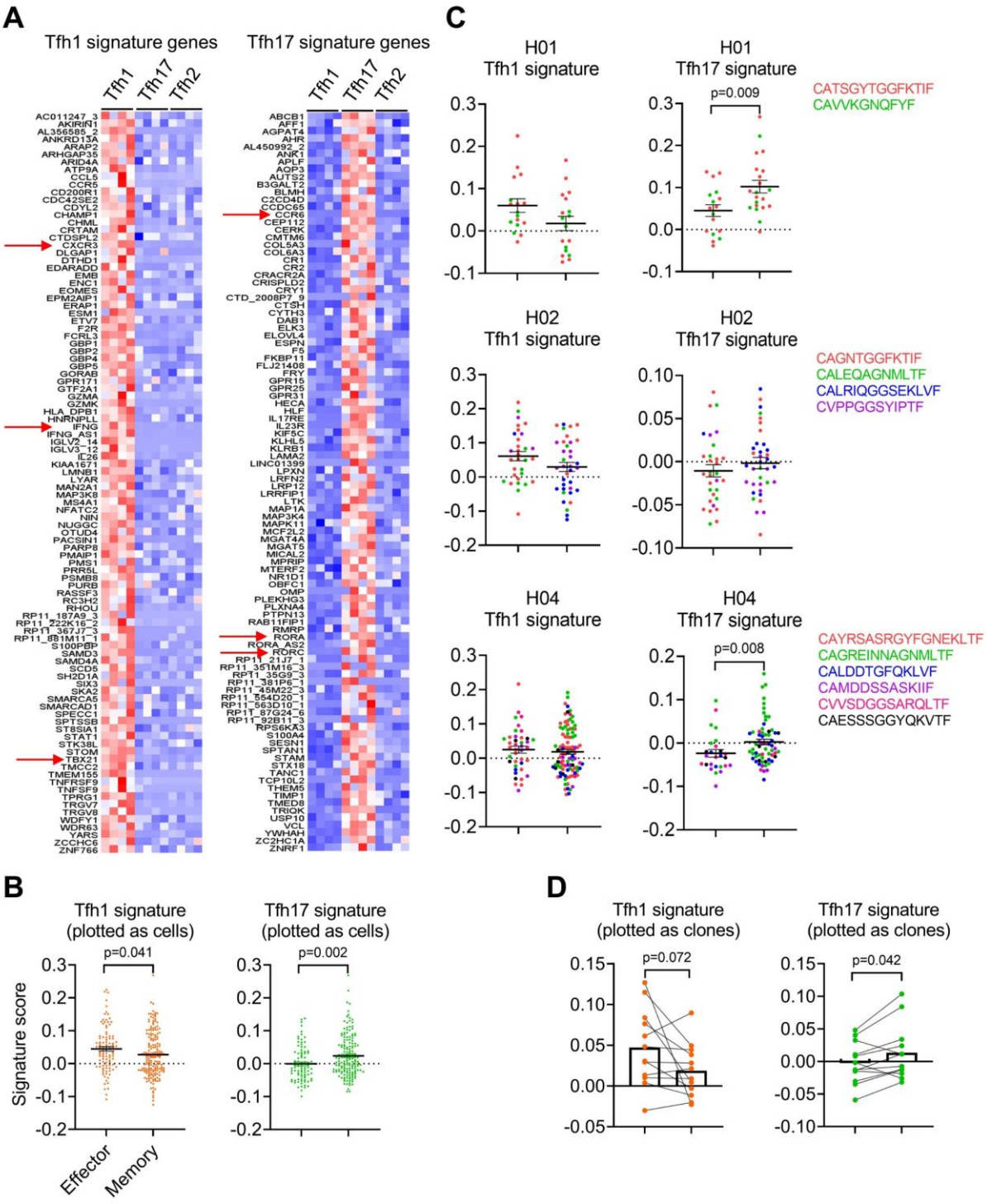

A
Tfh1 signature genes
Tfh17 signature genes
Tfh1
Tfh17
Tfh2
CXCR3
IFNG
CCR6
RORA
RORC
TBX21
B
Tfh1 signature (plotted as cells)
Tfh17 signature (plotted as cells)
p=0.041
p=0.002
Signature score
Effector
Memory
C
H01 Tfh1 signature
H01 Tfh17 signature
p=0.009
CATSGYTGGFKTIF
CAVVKGNQFYF
H02 Tfh1 signature
H02 Tfh17 signature
CAGNTGGFKTIF
CALEQAGNMLTF
CALRIQGGSEKLVF
CVPPGGSYIPTF
H04 Tfh1 signature
H04 Tfh17 signature
p=0.008
CAYRSASRGYFGNEKLTF
CAGREINNAGNMLTF
CALDDTGFQKLVF
CAMDDSSASKIIF
CVVSDGGSARQLTF
CAESSSGGYQKVTF
D
Tfh1 signature (plotted as clones)
Tfh17 signature (plotted as clones)
p=0.072
p=0.042

Figure 3.11 Influenza virus-specific cTfh cells show Tfh1 signatures in effector phase but Tfh17 signatures in memory phase

The single-cell RNA-seq dataset (GSE152522) was analyzed. (A) Top 100 signature genes for Tfh1 and Tfh17 were generated by Limma package based on bulk RNA-seq dataset GSE123812. The heatmap showing the expressions of Tfh1 and Tfh17 signature genes in Tfh12/17 samples scaled by z scores (scale: red: 0-2, white=0, blue: -2-0). (B) Comparison of Tfh1 and Tfh17 signature scores between effector and memory phase cTfh cells based on each cell. (C) Tfh1 and Tfh17 signature scores in were separated by individuals (H01, H02 and H04) and each clone was colour-coded accompanied by the CDR3 amino acid sequences. The statistics comparing the Tfh1 and Tfh17 signature scores between effector and memory phase influenza-specific cTfh cells. One individual (H03) was not included since lacking abundant clones. (D) Comparison of Tfh1 and Tfh17 signature scores between effector and memory phase cTfh cells based on TCR clone. The signature score of each clone was calculated as the mean value of the signature scores of all the cells in this clone. The p values were calculated by unpaired t-tests for (B, C) and paired t-tests for (D).

Tfh17 cells are long-lived and accumulate with aging

Our previous experiments have demonstrated that antigen-specific mouse iTfh17 cells and vaccine-specific human Tfh17 cells are superior to Tfh1 and Tfh2 cells in maintaining Tfh memory for a period of about one month (HBV) or less than one year (influenza vaccine). We next ask whether Tfh17 cells can persist for even longer periods, such as years. In a cohort of adults (N =20), we examined cTfh cells specific to vaccines for tetanus toxoid and measles, both administrated in childhood (**Figure 3.12A**). Given that community transmission of tetanus and measles is very rare (Huang et al. 2018), cTfh cells specific to tetanus toxoid and measles in adults were likely induced many years ago by childhood vaccination (Nanan et al. 2000, Locci et al. 2013, Van Damme et al. 2019). The average percentages of the Tfh17 subset in vaccine-specific memory cTfh cells were 55.22% and 45.07% for tetanus toxoid and measles respectively, which were more than twofold higher than the Tfh2 percentages and more than threefold higher than the Tfh1 percentages (**Figure 3.12B, C**). These results suggest Tfh17 cells may maintain Tfh memory for more than a decade. We also asked whether the Tfh17 dominance in memory phase was also applied to cTfh cells induced by SARS-CoV-2 infection. We examined convalescent patients with Covid-19 showing SARS-CoV-2-specific IgG antibodies (N = 13, **Figure 3.12D**). Similar to vaccine-specific cTfh cells, the Tfh17 percentages in SARS-CoV-2-specific cTfh cells were much higher than the Tfh1 or Tfh2 percentages (mean, 59.03% v.s. 12.87% or 7.73%) (**Figure 3.12E, F**).

If antigen-specific Tfh17 cells are long-lived and superior to Tfh1 and Tfh2 cells for persistence, we would expect a preferential accumulation of Tfh17 cells over Tfh1 or Tfh2 cells along with aging. By pooling results of cTfh characterization from cord blood, children, young and middle-aged adults and the elderly, we indeed observed that the Tfh17 percentages in total cTfh cells positively correlated with biological ages whereas the percentages of Tfh1 or Tfh2 subsets showed negative correlations (P-value < 0.001) **(Figure 3.12G, H)**.

In conclusion, Tfh17 cells are superior to Tfh1 and Tfh2 cells in Tfh memory maintenance, a phenomenon consistently observed in vaccination, infection and natural antigen exposure.

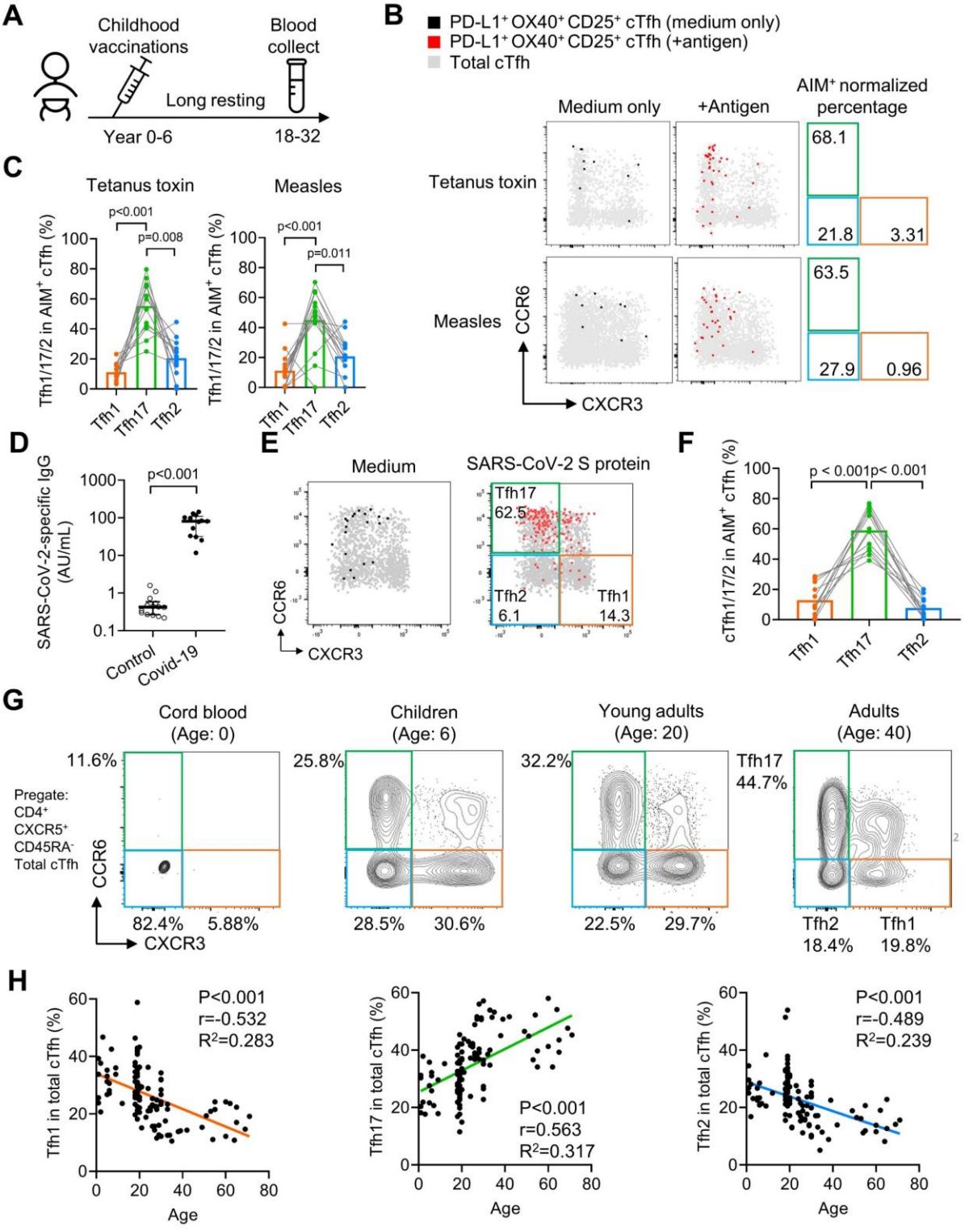

A
Childhood vaccinations
Blood collect
Long resting
Year 0-6
18-32
B
PD-L1+ OX40+ CD25+ cTfh (medium only)
PD-L1+ OX40+ CD25+ cTfh (+antigen)
Total cTfh
Medium only
+Antigen
AIM+ normalized percentage
Tetanus toxin
Measles
CCR6
CXCR3
68.1
21.8
3.31
63.5
27.9
0.96
C
Tetanus toxin
Measles
Tfh1/17/2 in AIM+ cTfh (%)
p<0.001
p=0.008
p=0.011
Tfh1
Tfh17
Tfh2
D
SARS-CoV-2-specific IgG (AU/mL)
p<0.001
Control
Covid-19
E
Medium
SARS-CoV-2 S protein
Tfh17
62.5
Tfh2
6.1
Tfh1
14.3
F
cTfh1/17/2 in AIM+ cTfh (%)
p < 0.001
p< 0.001
G
Cord blood (Age: 0)
Children (Age: 6)
Young adults (Age: 20)
Adults (Age: 40)
Pregate: CD4+ CXCR5+ CD45RA- Total cTfh
11.6%
25.8%
32.2%
Tfh17
44.7%
82.4%
5.88%
28.5%
30.6%
22.5%
29.7%
Tfh2
18.4%
Tfh1
19.8%
H
Tfh1 in total cTfh (%)
P<0.001
r=-0.532
R2=0.283
Tfh17 in total cTfh (%)
P<0.001
r=0.563
R2=0.317
Tfh2 in total cTfh (%)
P<0.001
r=-0.489
R2=0.239
Age

Figure 3.12 Tfh17 cells are long-lived and accumulate with aging
(A-C) PBMC samples from 20 healthy individuals were cultured for 18h with or without indicated antigens, followed by FACS to detect the phenotype of antigen-specific cTfh cells. Experiment design (A), representative FACS plot (B) and statistics (C) showing the percentage of Tfh1/2/17 cells in antigen-specific cTfh cells against tetanus toxin or measles. (D-F) Peripheral blood was collected from 14 healthy donors and 13 convalescent Covid-19 patients at 10-12 months post the infection, and ELISA or AIM assays were performed. The SARS-CoV-2 specific IgG antibody titers (D), representative FACS plots (E) and statistics (F) for Tfh1/2/17's percentages in AIM^+ cTfh in convalescent Covid-19 patients. One sample with poor signals in AIM assay were excluded. (G-H) PBMC samples from individuals of different ages were analysed. Representative FACS plots (G) showing the percentages of Tfh1/2/17 cells in total cTfh cells in individuals of different ages. Correlations tests between the biological age (H) with the percentages of Tfh1/2/17 cells in total cTfh cells. Cord blood samples were excluded from the correlation tests because of insufficient cTfh cell numbers. The *p* values were calculated by Friedman test for (C, F) and Pearson correlation for (H).

3.3 Discussion

Morita et al. reported circulating Tfh memory cells comprise different subsets related to Th1, Th2 and Th17 cells (Morita et al. 2011), which has advanced our understanding of Tfh subsets and helped to delineate the relationship between Tfh and other Th subsets. While the signatures of Tfh2 and Tfh17 activation were commonly reported in allergic and autoimmune diseases (Deng et al. 2019, Yao et al. 2021), the activation of Tfh1 cells was a prominent feature and associated with pathogen-specific antibody production in influenza vaccination (Bentebibel et al. 2013) and infections by HIV (Baiyegunhi et al. 2018), malaria (Obeng-Adjei et al. 2015) and more recently SARS-CoV-2 (Rydyznski Moderbacher et al. 2020, Dan et al. 2021). Beyond their distinct function in mediating isotype class switching (Hirota et al. 2013, Gowthaman et al. 2019), other difference between these Tfh subsets remains largely unknown, possibly due to the experimental hurdle of a low frequency of Tfh subsets in human blood and the lack of culture method for *in vitro* Tfh subset generation and *in vivo* functional characterization.

We optimized the method for the *in vitro* induction of antigen-specific Tfh differentiation (Gao et al. 2020) and modified the protocol by adding the conditions biased for Th1/2/17 polarization. This method corroborated the report that Tfh cells are plastic and carry positive epigenetic markings for Th1/2/17 cells (Lu et al. 2011). Notably, for the iTfh17 condition (0.1 ng/mL TGF-β + 100 ng/mL IL-6 + 50 ng/mL IL-21), TGF-β were used in a concentration of 0.1 ng/mL, much lower than those for Th17 and Treg polarization (normally 1-10 ng/mL). The condition successfully generated iTfh17 expressing both CXCR5, PD-1, BCL6 and RORγt. This phenomenon might also help to reconcile the reports that TGF-β signaling can either inhibit or support Tfh differentiation (McCarron and Marie 2014, Marshall et al. 2015), probably determined by the TGF-β signal strength.

The generation of antigen-specific iTfh1/2/17 cells in decent numbers using this modified method allowed us, for the first time, to compare Tfh1, Tfh2 and Tfh17 function *in vivo*. The results from the adoptive transfer experiment revealed that, after an extended period of *in vivo* resting to mimic memory maintenance, iTfh17 cells showed a better function than iTfh1/2 cells in supporting humoral immunity. In agreement with this, the Tfh17 subset in human cTfh cells also showed superiority over Tfh1 and Tfh2 subsets for memory maintenance. Tfh17 cells predominantly showed the Tfh_{CM} phenotype as long-lived memory cells and dominated the long-lived pool of antigen-specific memory Tfh cells for vaccines of HBV, influenza, tetanus and measles. In contrast, the human Tfh1 subset appears short-lived and accounted for the least

proportion of long-lived antigen-specific memory Tfh cells. Tfh1 is the major Tfh subset induced by influenza infection and vaccination (Bentebibel et al. 2013). The short-lived characteristics of Tfh1 cells may partially contribute to a relatively short period of humoral immunity after influenza vaccination (Young et al. 2018). In SARS-CoV-2 infection, patients with acute infection demonstrated a Tfh1-biased profile, while convalescent patients increased the proportions of virus-specific Tfh17 cells, again supporting the notion that Tfh17 cells represent the population better in memory maintenance (Rydyznski Moderbacher et al. 2020, Dan et al. 2021). In malaria infection, a recent study also reported Tfh17 and Tfh2 cells, rather than Tfh1 cells showed a Tfh_{CM} phenotype (Chan et al. 2020). All such evidence suggests the superiority of Tfh17 subset in memory maintenance appears to be a common feature for immune responses induced by both vaccination and infection. It should be noted that our results should not be misinterpreted as that Tfh17 cells are always the major subset for Tfh memory cells. In the case of SARS-CoV-2 mRNA vaccine which induces strong Th1-polarized response, Tfh17 cells are essentially not induced and the Tfh memory are maintained in the absence of Tfh17 cells (Goel et al. 2021, Tauzin et al. 2021, Wragg et al. 2022).

The mechanism underlying the advantage of Tfh17 cells in memory maintenance is an area we will focus on in the following studies. Intriguingly, Th17 cells were reported to have "stem cell-like" features and are long-lived (Kryczek et al. 2011, Muranski et al. 2011, Lindenstrom et al. 2012). The proposed mechanisms are diverse, which include a high expression of Tcf1 in Th17 cells, a key transcription factor that regulates T cell memory generation and self-renewal and favorable expression of anti-apoptosis Bcl-2 family genes to sustain the longevity (Kryczek et al. 2011, Muranski et al. 2011). The Th17's hallmark transcription factor RORγ has been shown to directly promote T cell survival by enhancing Bcl-xL expression (Sun et al. 2000). Effector and memory Tfh cells are critically regulated by specific cell death pathways of ferroptosis and pyroptosis (Yao et al. 2021, Chen et al. 2022). Future works are required to test whether these mechanisms also apply in the survival and self-renewal of Tfh17 cells. Additionally, the higher expression of CCR7 and CCR6 may direct Tfh17 cells to migrate to locations which favor cell survival, therefore future work should also investigate the distribution of Tfh1/2/17 cells.

Unveiling the superiority of Tfh17 in Tfh memory maintenance can help us to improve the rationale-based vaccine development. Many vaccines, including conventional vaccines for influenza virus and novel mRNA vaccines for SARS-CoV-2, induce Th1 responses and Tfh1-associated humoral immunity (Bentebibel et al. 2013, Lederer et al. 2020, Sahin et al. 2020). According to our results, Tfh1 cells are short-lived, which might curb the duration of vaccine-

mediated protection. New strategies might be taken to direct vaccination for more Tfh17 induction which can support a better Tfh memory formation and potentially prolong vaccine protection.

Chapter4- ZEB2 Regulates the Development of CD11c+ Atypical B Cells

4.1 Introduction

Following the resolution of infection populations of memory lymphocytes, including B cells persist. Classical memory B cells (cMBCs) express high levels of CD27 and rapidly secrete antibody in response to restimulation. However non-classical memory B cells have also been identified in a variety of different conditions. A population of FCRL4 expressing B cells was first identified in the tonsils of individuals undergoing tonsillectomy (Ehrhardt et al. 2005). Phenotypically similar $CD21^{-}$ $CD27^{-}$ cells were subsequently found at elevated frequencies in the blood of HIV and malaria infected individuals (Moir et al. 2008, Weiss et al. 2009, Portugal et al. 2015, Obeng-Adjei et al. 2017, Sundling et al. 2019, Holla et al. 2021, Sutton et al. 2021). Collectively these cells associated with chronic infection were designated "atypical" B cells (ABCs). A related population of IgD^{-} $CD27^{-}$ $CXCR5^{-}$ $CD21^{-}$ double negative 2 (DN2) B cells and $CD11c^{+}$ $CD11b^{+}$ age associated B cells has also been identified in a variety of autoimmune conditions and their number generally correlates with disease severity (Wei et al. 2007, Hao et al. 2011, Rubtsov et al. 2011, Jenks et al. 2018, Li et al. 2021). While originally associated with chronic immune conditions or infection, ABCs have also seen to respond to acute viral infections such as SARS-CoV-2 or primary vaccination with yellow fever, vaccinia virus vaccine or attenuated *Plasmodium* sporozoites (Knox et al. 2017, Woodruff et al. 2020, Sutton et al. 2021). These studies have made clear that ABCs are present in normal immune responses (Sutton et al. 2021).

While they are clearly present in responses to vaccination it is unclear whether ABCs are required for the efficient humoral responses. In autoimmunity they are associated with disease activity, enriched with autoantibody clones and have elevated plasma cell gene such as IRF4 expression thus they are also proposed to be the precursors of antibody secreting cells (Jenks et al. 2018, Wang et al. 2018). However in ABCs isolated from individuals infected by human immunodeficiency virus or hepatitis C virus were found to be unresponsive to BCR stimulation and have poor expansion in vitro (Moir et al. 2008, Charles et al. 2011), therefore they are proposed to be exhausted like cells to limit the inflammation-induced pathology (Charles et al. 2011). Additionally, ABCs have increased expression of genes associated with antigen presentation like CD86 and can efficiently prime T cells, especially follicular helper T (Tfh) cells in vitro thus they are also suggested to be antigen presenting cells, potentially capable of supporting other immune responses (Rubtsov et al. 2015, Zhang et al. 2019).

Studies of ABC function are impaired by a lack of knowledge of the key signals and transcription factors that drive their formation. IFNγ inducible T-box transcription factor 21 (Tbet) has been frequently found to be expressed on many alternative lineage cells in autoimmune and infectious disease conditions (Obeng-Adjei et al. 2017, Jenks et al. 2018, Austin et al. 2019). Although the number of circulating $CD11c^+$ $CD21^{low}$ ABCs were found low in one patient with inherited Tbet deficiency (Yang et al. 2022), in animal models of systemic lupus erythematosus (SLE) Tbet has been found to be dispensable for disease progression and the formation of $CD11c^+$ $CD11b^+$ ABCs (Du et al. 2019, Ricker et al. 2021). Moreover, in an *E. muris* model of infection $CD19^{hi}$ $CD11c^+$ ABCs formed normally in Tbet knockout (KO) animals (Levack et al. 2020). Therefore, it is unlikely that Tbet is the lineage defining TF for ABCs. Other TFs that have been associated with the ABC phenotype include ZBTB32, SOX5, BHLHE40, TOX and ZEB2 (Gao and Cockburn 2022).

Here we found that the transcription factor Zeb2 was upregulated in human ABCs and an analogous population in mice. Functional characterization of murine ABCs showed they were exhausted-like cells with poor survival and recall responses. Zeb2 KO resulted in an almost complete loss of ABC formation in mice, however other aspects of the humoral response proceeded normally in response to vaccination and infection. In humans ZEB2 haploinsufficient MWS patients also had reduced ABCs but normal development of other B cell population. Thus we show that ZEB2 is selectively required for the formation of ABCs, but ABCs themselves appear dispensable for efficient humoral immunity.

4.2 Results

Single cell RNA-seq of human tonsils identifies TFs associated with the ABC lineage

To extend our understanding of ABCs in lymphoid tissues we analysed B cells from the tonsils of two donors by single cell-RNA seq. To correlate transcriptomic data with surface marker expression we included a panel of 14 barcoded antibodies for cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) analysis. These antibodies were specific for surface markers that have previously been suggested to delineate different memory B cell subsets (Glass et al. 2020). Because tonsillar B cells are dominated by IgD^+ naïve cells that were of limited interest to us, we sorted three samples from each donor (i) total tonsil B cells (ii) $CD10^+$ IgD^- germinal center (GC) tonsillar B cells and (iii) CD10- IgD^- memory B cells (**Figure 4.1**). This allowed us to focus on populations of interest, without losing information about the relative frequencies of each population. In addition to our extended CITE-seq analysis we performed VDJ sequencing on the B cells to determine mutational frequencies.

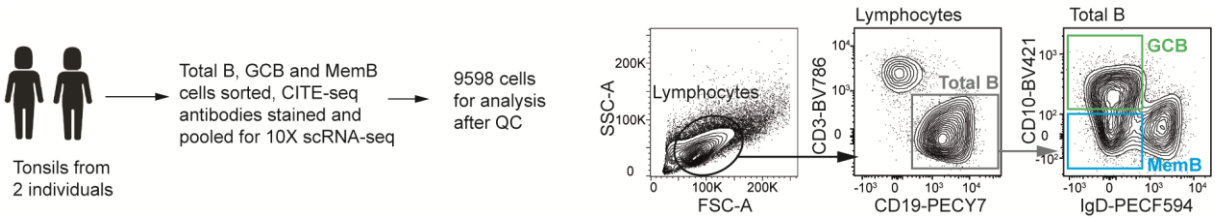


Figure 4.1 Single cell RNA-seq of human tonsil B cells
Human tonsil total B cells, GC B cells and memory B cells were sorted, labelled by CITE-seq antibodies and mixed at 1:1:1 ratio for 10X scRNA-seq. The data was processed through Cellranger pipeline and analyzed by Seurat package. Experiment design and gating strategy for sorting.

Consistent with previous scRNA-seq analyses of lymphoid tissues (Turner et al. 2020, King et al. 2021), tonsillar B cells formed three distinct major clusters (**Figure 4.2A**) – a population of antibody secreting cells (plasmablast/plasmacell, PB/PC), a GC supercluster divided into light zone (LZ), dark zone (DZ), proliferating cells (ProGC) and unassigned GC1, GC2, and finally a supercluster of more quiescent cells that were a mix of naïve and memory cells (**Figure 4.2B**). In addition to naïve cells and classical memory B cells (cMBC) this latter cluster included IgD^+ memory cells, and some Pre-GC memory cells that are enriched with LZ genes such as CD83 (**Figure 4.2C**). A population of cells expressing *FOS, JUN* and *CD69* was observed, akin to the activated memory B cells in blood observed by us and others (**Figure 4.2C**) (Sutton et al. 2021, Zhang et al. 2021). ABCs could be identified as a distinct peninsular within the memory/naïve cluster that was enriched for the expression of previously identified ABC genes and down regulation of cMBC genes **(Figure 4.2D)** (Sullivan et al. 2015). Cells in this ABC cluster also had the highest surface expression of FCRL4, CD11c and CD19 by CITE-seq analysis**,** while GC B cells had no CD11c expression **(Figure 4.2E, F).** As expected, combining FCRL4 vs CD11c or commonly used CD27 vs CD21 can efficiently separate ABC from cMBC (**Figure 4.2G**).

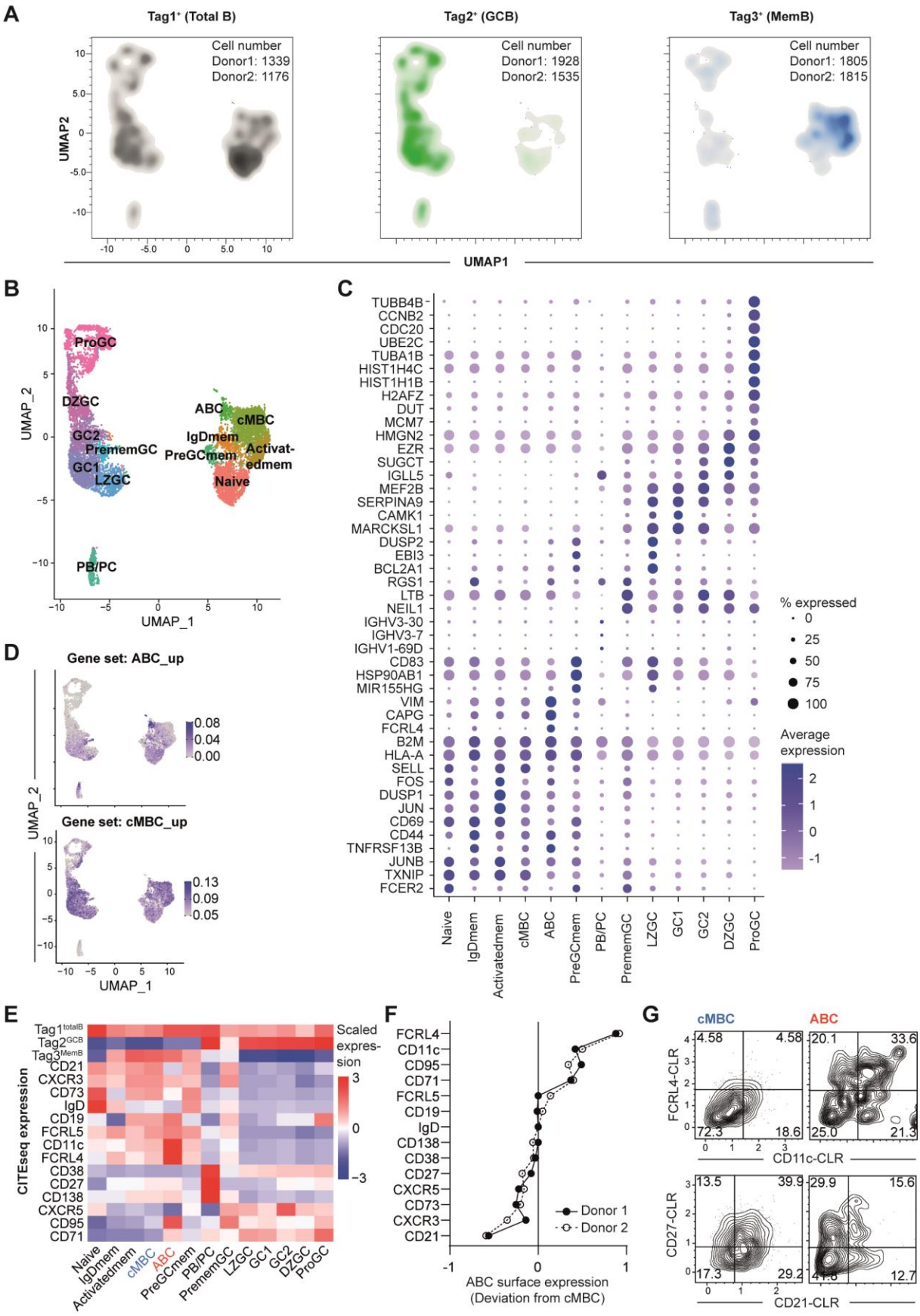

A
Tag1+ (Total B)
Tag2+ (GCB)
Tag3+ (MemB)
Cell number
Donor1: 1339
Donor2: 1176
Donor1: 1928
Donor2: 1535
Donor1: 1805
Donor2: 1815
UMAP2
UMAP1
B
ProGC
DZGC
GC2
PrememGC
GC1
LZGC
ABC
cMBC
IgDmem
PreGCmem
Activat-
edmem
Naive
PB/PC
UMAP_2
UMAP_1
C
TUBB4B
CCNB2
CDC20
UBE2C
TUBA1B
HIST1H4C
HIST1H1B
H2AFZ
DUT
MCM7
HMGN2
EZR
SUGCT
IGLL5
MEF2B
SERPINA9
CAMK1
MARCKSL1
DUSP2
EBI3
BCL2A1
RGS1
LTB
NEIL1
IGHV3-30
IGHV3-7
IGHV1-69D
CD83
HSP90AB1
MIR155HG
VIM
CAPG
FCRL4
B2M
HLA-A
SELL
FOS
DUSP1
JUN
CD69
CD44
TNFRSF13B
JUNB
TXNIP
FCER2
Naive
IgDmem
Activatedmem
cMBC
ABC
PreGCmem
PB/PC
PrememGC
LZGC
GC1
GC2
DZGC
ProGC
% expressed
0
25
50
75
100
Average
expression
2
1
0
-1
D
Gene set: ABC_up
0.08
0.04
0.00
Gene set: cMBC_up
0.13
0.09
0.05
UMAP_2
UMAP_1
E
Tag1totalB
Tag2GCB
Tag3MemB
CD21
CXCR3
CD73
IgD
CD19
FCRL5
CD11c
FCRL4
CD38
CD27
CD138
CXCR5
CD95
CD71
CITEseq expression
Scaled
expres-
sion
3
0
−3
Naive
IgDmem
Activatedmem
cMBC
ABC
PreGCmem
PB/PC
PrememGC
LZGC
GC1
GC2
DZGC
ProGC
F
FCRL4
CD11c
CD95
CD71
FCRL5
CD19
IgD
CD138
CD38
CD27
CXCR5
CD73
CXCR3
CD21
Donor 1
Donor 2
-1
0
1
ABC surface expression
(Deviation from cMBC)
G
cMBC
ABC
4.58
4.58
20.1
33.6
72.3
18.6
25.0
21.3
FCRL4-CLR
CD11c-CLR
13.5
39.9
29.9
15.6
17.3
29.2
41.6
12.7
CD27-CLR
CD21-CLR

Figure 4.2 Unsupervised clustering and the identification of ABCs
Human tonsil total B cells, GC B cells and memory B cells were sorted, labelled by CITE-seq antibodies and mixed at 1:1:1 ratio for 10X scRNA-seq. The data was processed through Cellranger pipeline and analyzed by Seurat package. Experiment design and gating strategy for sorting. (A) UMAP of sorted total B cells, GC B cells and memory B cells labelled by Hashtag 1, 2 and 3 respectively. (B) UMAP of B cells with identified clusters. (C) Dot plot showing the expression of marker genes for each cell cluster identified by FindAllMarkers function. (D) Feature plot showing the average expressions of ABC and cMBC gene sets. (E) Heatmap of surface marker expression of all identified clusters by CITE-seq. The gene express is scaled by z score. (F) Statistics showing the relative expression of indicated surface markers in ABC comparing to cMBC. The deviation values were calculated as ABC CLR median value minus cMBC CLR median value. (G) Representative FACS plot showing the expression of indicated genes by CITE-seq.

VDJ analysis showed that ABCs were not notably different from cMBCs with respect to their antibody isotype usage (**Figure 4.3A**). Moreover, they had only marginally lower frequencies of somatic hypermutation (SHM) compared to cMBCs and GC B cells (**Figure 4.3B**), suggesting the majority of tonsillar ABCs are GC-experienced, though there was a weak inverse correlation between the SHM rate and the strength of expression of ABC genes (**Figure 4.3C**)

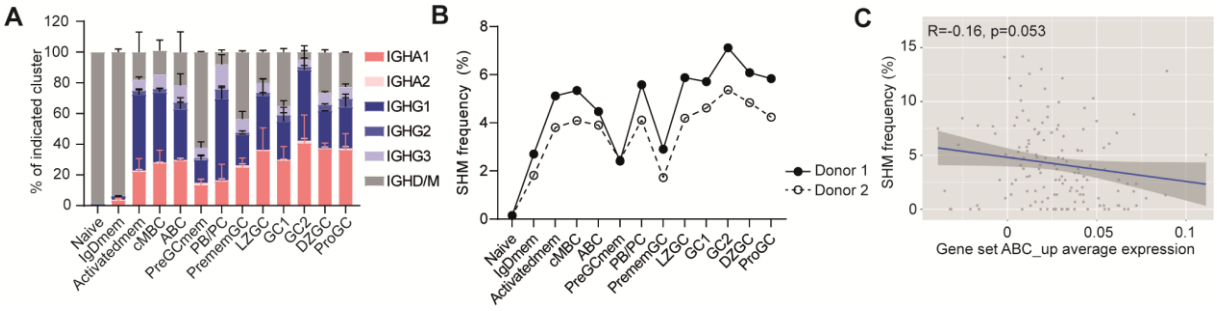


Figure 4.3 VDJ analysis of human tonsil B cells

Human tonsil total B cells, GC B cells and memory B cells were sorted, labelled by CITE-seq antibodies and mixed at 1:1:1 ratio for 10X scRNA-seq. The data was processed through Cellranger pipeline and analyzed by Seurat package. VDJ sequences were mapped to IMGT database through V-QUEST to identify heavy chain usage and mutation frequency for each cell. (S) The distribution of immunoglobulin isotypes for each cell clusters. (B) Statistics of SHM frequency for each cell clusters. (C) Pearson correlation analysis between the SHM frequency and the average expressions of ABC genes in ABC cluster.

To determine what may be the key TFs required for the formation of ABCs we examined which TFs were significantly upregulated in ABCs compared to cMBCs within our dataset. This revealed the top 5 candidate TFs: SOX5; ZBED2; BHLHE40; ZEB2; and ZBTB32 (**Figure 4.4A, B**). BHLHE40 has been showed to restrain the proliferation of GC and Tfh cells (Rauschmeier et al. 2022), and ZBTB32 has been showed to suppress the secondary response of memory B cells (Jash et al. 2016). However, none of these TFs have been studied in terms of ABCs formation. Ideally, the key TFs should promote ABC genes and suppresses cMBC genes, therefore, we performed an analysis to determine which TF correlated most strongly in individual cells with the expression of the ABC geneset, the surface expression of CD11c, and which were most inversely correlated with cMBC gene expression and surface CD27 expression. In this analysis ZEB2 was the top candidate, though expression of BHLHE40 and SOX5 also correlated with the expression of ABC geneset (**Figure 4.4C**).

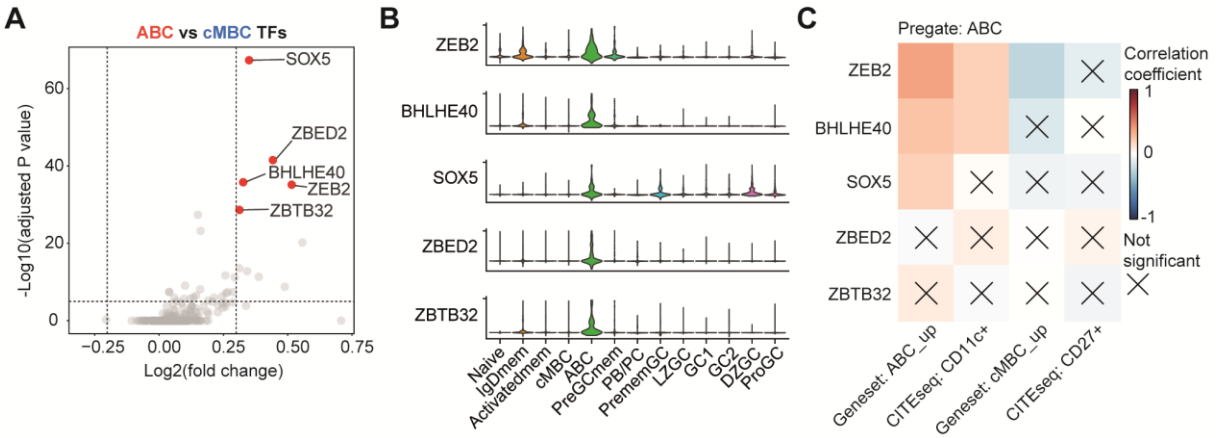


Figure 4.4 Transcription factor analysis of ABC and cMBC

Human tonsil total B cells, GC B cells and memory B cells were sorted, labelled by CITE-seq antibodies and mixed at 1:1:1 ratio for 10X scRNA-seq. The data was processed through Cellranger pipeline and analyzed by Seurat package. (A) Volcano plot comparing the expressions of TFs in ABC and cMBC, the top5 TFs were highlighted. (B) Stacked violin plot showing the expression of the top5 TFs in all clusters. (C) Pearson correlation tests between TFs expression and indicated features in each cell in ABC cluster were performed and the summarized as heatmap.

Murine ABCs induced by infection or immunization share a transcriptomic program with human ABCs

To determine the function of ABCs and to facilitate the identification of key factors driving ABC development we sought to develop experimentally tractable mouse models of ABC formation. In agreement with others (Perez-Mazliah et al. 2018), we identified a population of $CD11c^+$ B cells after infection with rodent malaria *P. chabuadi* (**Figure 4.5A, B**) that had a distinct transcriptional profile from cMBCs (**Figure 4.5C**). This transcriptomic program included many genes homologous to those expressed by human ABCs (**Figure 4.5D**). The similarity of these cells to human ABCs was confirmed by gene set enrichment analysis (GSEA) against the signature genes of our human tonsil ABC cluster (NES 2.37; FDR<0.001).

To study antigen specific ABCs after immunization, we transferred Igh^{g2A10} cells that are specific for the *P. falciparum* circumsporozoite protein (PfCSP) (McNamara et al. 2020) to MD4 recipients and immunized with *P. berghei*-PfCSP sporozoites (PfCSP-SPZ) that express the PfCSP in place of the endogenous *P. berghei* CSP protein (Espinosa et al. 2017). We expected that this would result in the induction of ABCs as our previous analysis showed that immunization of humans with *Plasmodium* sporozoites induces a population of PfCSP-specific ABCs after primary immunization (Sutton et al. 2021), The use of recipient Ig-transgenic mice which express an irrelevant BCR - in this case specific for hen egg lysozyme - is an established approach for the study of rare memory cell populations (Zuccarino-Catania and Shlomchik 2015). Immune responses were assessed 4, 10 and 21 days post-immunization (**Figure 4.5E**). At all time points $CD11c^+$ and $CD11c^-$ switched B cells were identified in the spleen, LN, blood and bone marrow of mice (**Figure 4.5F, G**), with ABCs being most enriched in the spleen where they peaked on day 10 accounting for ~15% of memory Igh^{g2A10} cells (**Figure 4.5G**). The identity of these cells as ABCs was confirmed by bulk RNA-seq analysis of $CD11c^+$ and $CD11c^-$ cells at each timepoint. $CD11c^+$ B cells were clearly separated from $CD11c^-$ B cells by PCA analysis especially at days 10 and 21 (**Figure 4.5H**). $CD11c^+$ Igh^{g2A10} cells also up-regulated many known ABC genes and down-regulated cMBC genes (**Figure 4.5I**), which was further confirmed by GSEA comparing to our human tonsillar ABC (NES 2.14; FDR<0.001).

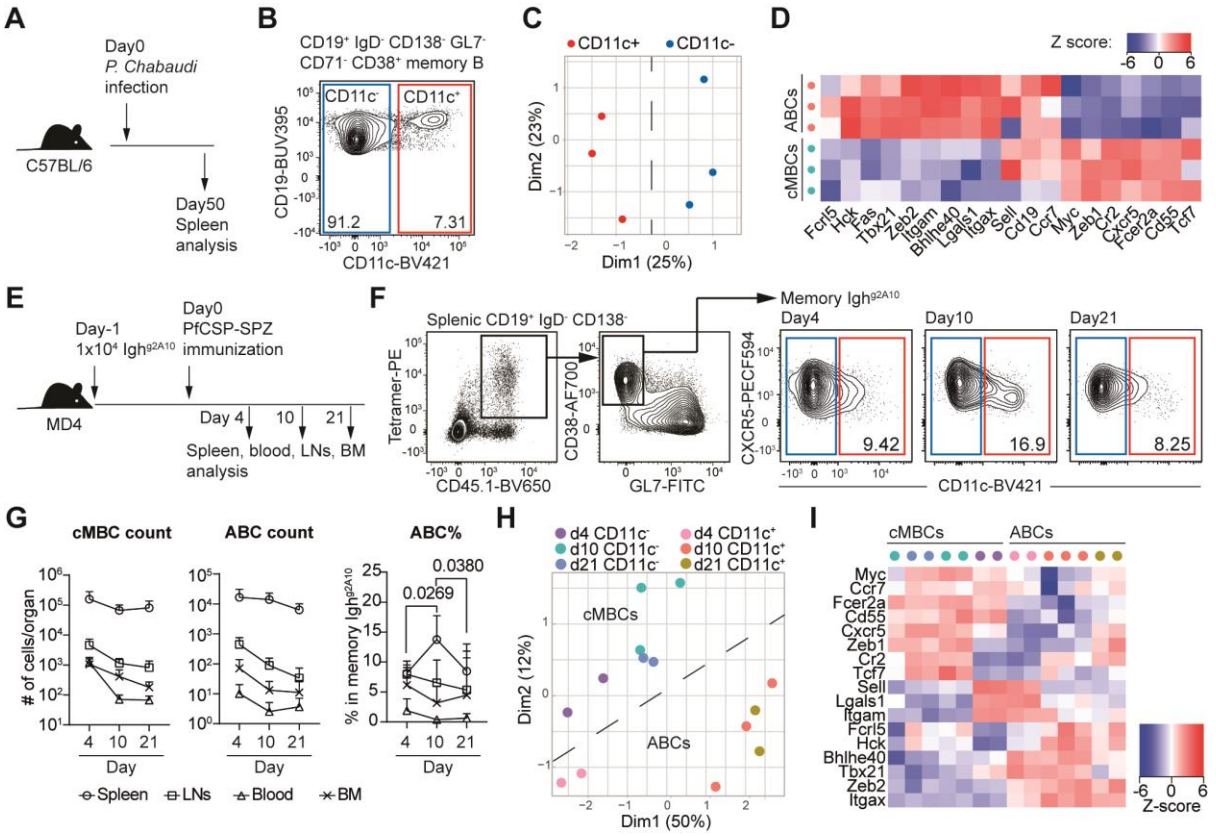


Figure 4.5 Murine ABCs induced by infection or immunization share a transcriptomic program with human ABCs

(A-D) WT mice were infected by IV injecting 10^5 P. *chabuadi* iRBC and spleens were analysed on day50. CD71$^-$ resting CD11c$^+$ and CD11c$^-$ memory B cells were also sorted for bulk RNA-seq analysis. (A) Experiment design. (B) Representative FACS plot for ABC formation. (C) PCA comparing the transcriptomes of CD11c$^+$ and CD11c$^-$ cells. (D) Heatmap showing the expression of ABC and cMBC signature genes in all samples. (E-J) 10^4 Igh^{g2A10} B cells were IV injected into MD4 mice, followed by PfCSP-SPZ immunization. Spleens, blood, LNs and BM were analysed on day4, day10 and day21 after immunization. Splenic memory Igh^{g2A10} CD11c$^-$ and CD11c$^+$ cells were also sorted for bulk RNA-seq analysis. (E) Experiment design. (F) Representative FACS plot for splenic ABCs percentages in memory Igh^{g2A10}. (G) Statistics of cMBC and ABC Igh^{g2A10} total cell count and percentage in memory cells. (H) PCA plot comparing the transcriptomes of CD11c$^-$ and CD11c$^+$ memory Igh^{g2A10} cells on all timepoints. (I) Heatmap showing the expression of ABC and cMBC signature genes in all samples. The results in (G) were pooled from two independent experiments. The P values were calculated by One-way ANOVA.

ABCs are not PB/PC precursors and have poorer survival and proliferation potential than cMBCs

Having identified ABCs in murine models we next set out to determine their longevity and ability to contribute to recall responses. We transferred *P. chabuadi* infection-induced ABCs or cMBCs into naïve recipients to track their survival. We further measured the expression of PB/PC and GC B cell markers in resting settings as it has been suggested that ABCs are PC precursors (**Figure 4.6A**). While transferred cMBCs in the spleen were generally maintained, only about 20% ABCs could be detected 14 days after the transfer (**Figure 4.6B**). Moreover, during the experiment neither ABCs nor cMBCs significantly upregulated the expression of CD138 or GL7 (**Figure 4.6C, D**), suggesting ABCs did not spontaneously become PB/PC or GC B cells. To assess recall responses by ABCs and test whether ABCs were poised to differentiate into PB/PC upon antigen re-stimulation, we transferred PfCSP-SPZ immunization primed *Igh*^{g2A10} ABCs or cMBCs into naïve MD4 recipients followed by PfCSP-SPZ or recombinant CSP in alum immunization (**Figure 4.6E**). Both cMBCs and ABCs could mount recall responses, however the responses by ABCs were smaller in magnitude (**Figure 4.6F, G**). For both memory cell types the majority of cells differentiated into PB/PCs upon restimulation (**Figure 4.6F, G**). Collectively our data suggests that ABCs are exhausted-like cells with poorer survival and proliferation potential than cMBCs.

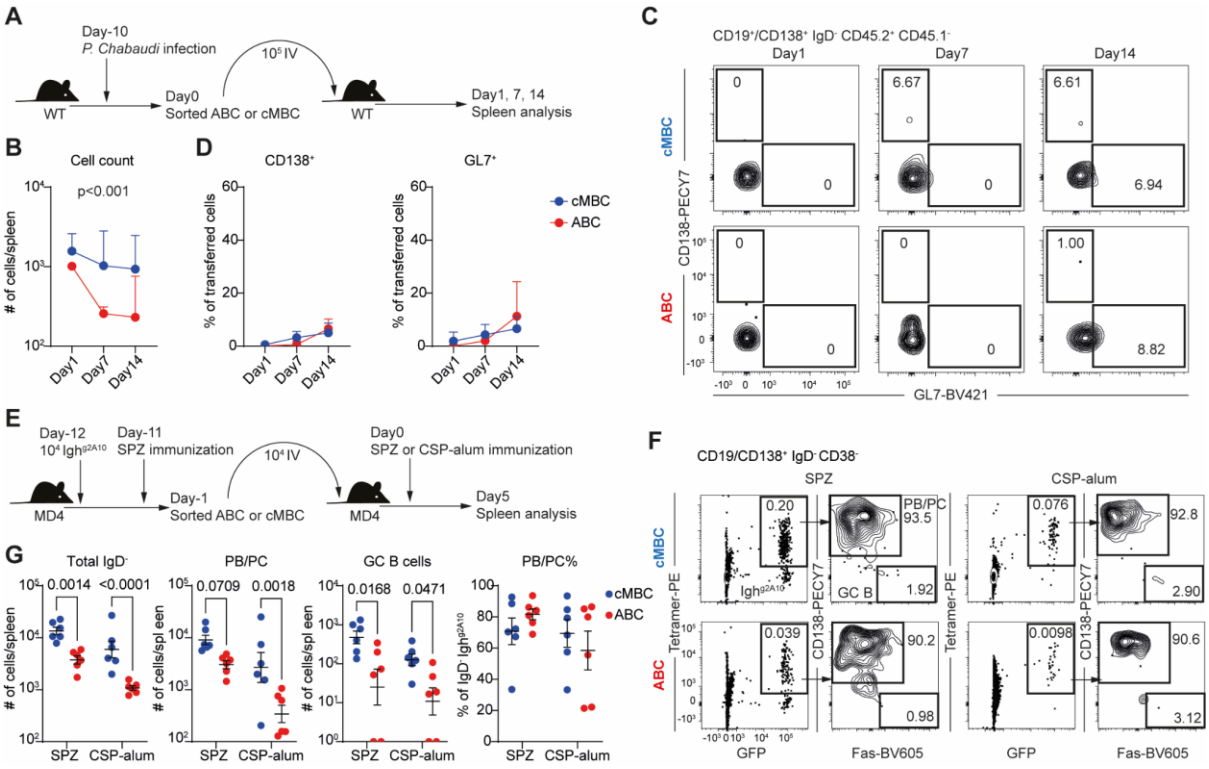


Figure 4.6 ABCs are not PB/PC precursors and have poorer survival and proliferation potential than cMBCs

(A-D) WT mice were infected by 10^5 *P. chabaudi* parasites, after 10 days $CD19^+$ IgD^- $CD138^-$ $GL7^-$ $CD38^+$ $CD11c^+$ and $CD11c^-$ ABCs and cMBCs were sorted and transferred into naïve WT congenic recipients, followed by spleen analysis on day1, 7 and 14 after the cell transfer. (A) Experiment design. (B) Statistic showing the cell count of transferred ABC and cMBC in the spleen. (C) Representative FACS plots and (D) statistics showing the indicated features of transferred ABCs and cMBCs. (E-G) 10^4 Igh^{g2A10} B cells were transferred into MD4 mice followed by PfCSP-SPZ immunization, and the spleens were collected 10 days after the immunization to sort $CD19^+$ IgD^- $CD138^-$ $GL7^-$ $CD38^+$ $CD11c^+$ and $CD11c^-$ Igh^{g2A10} ABCs and cMBCs. 10^4 sorted ABCs or cMBCs were transferred into naïve MD4 recipients, followed by PfCSP-SPZ or CSP-alum immunization and spleen analysis on day5 post immunization. (E) Experiment design. (F) Representative FACS plots and (G) statistics showing the numbers or percentage of indicated Igh^{g2A10} populations. Results were pooled from two independent experiments for (B, D, G). The P values were calculated by two-way ANNOVA for (B, G).

A CRISPR/Cas9 knock-out screen identifies Zeb2 directs the expression of ABC and suppression of cMBC genes

We next set out to identify the key factors that drive the ABC transcriptional program. Firstly, we wanted to know if ABC formation was driven by one or more lineage defining transcription factors. Secondly, we hoped that if the ABCs could be ablated by the loss of a key transcription factor we could gain insights into their normal function. We first defined conditions to induce ABCs in vitro (**Figure 4.7A**). Consistent with previous reports (Naradikian et al. 2016, Manni et al. 2018, Wang et al. 2018, Levack et al. 2020), we found that IL21 was critical for driving the development of a $CD11c^{+}$ $CD21^{-}$ cells in vitro after stimulation with anti-CD40 (**Figure 4.7B, C**). Based on our analysis of tonsil single cell dataset we further found that ABCs had upregulated TGFβ signaling pathway (**Figure 4.7D**). In agreement with this, the addition of TGFβ modestly enhanced ABC formation, while anti-TGFβ almost completely abolished ABC formation in vitro **(Figure 4.7B, C**). Bulk RNA-seq analysis confirmed the identity of our in vitro induced ABCs with upregulation of ABC genes and downregulation of cMBC genes (**Figure 4.7E, F**).

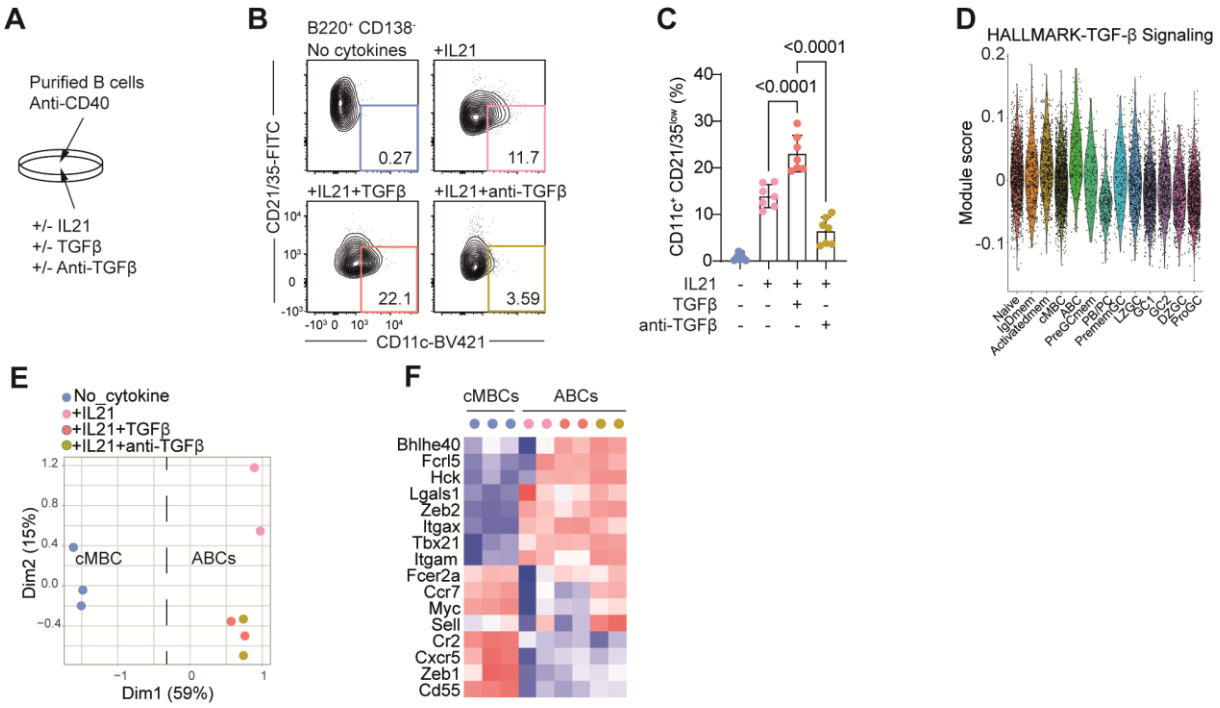


Figure 4.7 The in vitro differentiation of ABC-like cells

MACS enriched mouse B cells were cultured with anti-CD40 w/wo IL-21 and TGFβ for 4 days followed by FACS analysis or sorting ($B220^+$ $CD138^-$) for bulk RNA-seq analysis. (A) Experiment design. (B) Representative FACS plot and (C) statistics of $CD11c^+$ $CD21/35^{lo}$ cells for each culture condition. (D) The violin plot showing the enrichment scores for TGFβ signalling pathway in all clusters in the human tonsil B cell dataset. (E) PCA plot comparing the transcriptomes of B cells cultured under indicated conditions. (F) Heatmap showing the expression of ABC and cMBC signature genes in all samples. The results were pooled from two independent experiments for (C). The P values were calculated by One-way ANOVA.

Because the ABCs are transcriptomically similar in human and mouse, the key TFs for ABC should also be shared between these RNA-seq datasets. Therefore, we calculated the upregulated genes in these datasets and identified 5 highly conserved TFs including *Zeb2*, *Tbx21*, *Bhlhe40*, *Tcf7l2* and *Mafb* for downstream analysis (**Figure 4.8A**). We also included *Bhlhe41* and *Maf* because they belong the same families as Bhlhe40 and Mafb respectively. Runx2 was previously proposed to play a role in human ABC formation (Ehrhardt et al. 2008) and so was also included.

To validate these TFs, we developed an optimized Cas9 ribonuclear protein (RNP) mediated gene knock-out (KO) method to delete individual TFs in cultured B cells. To optimize the gene KO efficiency, we co-transfected our target gene RNP with a sub-optimal concentration of GFP RNP into GFP^+ B cells (**Figure 4.8B**). The co-transfection of GFP RNP will allow us to identify the successfully transfected GFP^- cells with enhanced target gene KO efficiency. We validated this method by KO of Bcl6, the master TF for GC B cells and as predicted, found an enhanced KO efficiency in GFP^- cells (80-90% indel) compared to GFP^+ counterparts (60-75% indel) as determined by Sanger sequencing (**Figure 4.8C**). The robustness of this method was functionally confirmed by the nearly complete abolishment of GC B cells in the transferred GFP^- Igh^{g2A10} cells following transfer to MD4 mice and PfCSP-SPZ immunization (**Figure 4.8D-F**). Because previous studies have been conflicting on whether ABC formation requires GC (Levack et al. 2020, Song et al. 2022), we also analyzed the percentages of ABCs in memory Igh^{g2A10} B cells and found Bcl6 was dispensable for ABC formation (**Figure 4.8E, G**).

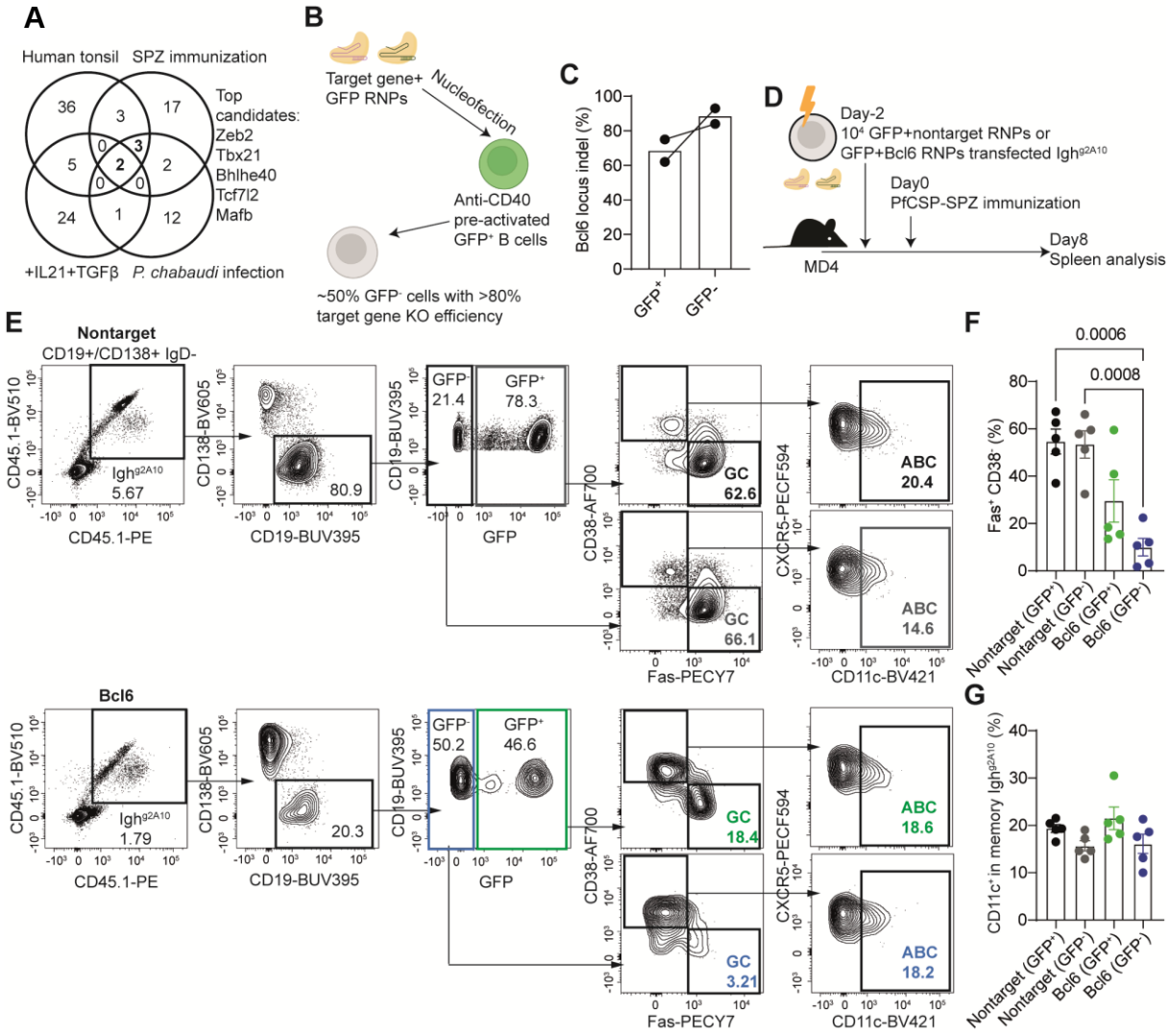


Figure 4.8 An optimized Cas9-RNP based gene knock-out method in mouse B cells
(A) Venn plot showing the numbers of significantly upregulated TFs in indicated RNA-seq datasets by comparing ABC vs cMBC. Threshold were set as follows: Human tonsil (FDR ≤0.01), the other datasets: (P ≤0.05, Log2FC ≥0.5). (B-G) MACS enriched GFP-expressing Igh^{g2A10} B cells were pre-activated by anti-CD40, nucleofected with Bcl6 and GFP Cas9 RNPs and transferred into MD4 mice followed by PfCSP-SPZ immunization. A small aliquot of cells was rested in anti-CD40 supplemented media for 24h, followed by gDNA extraction, PCR and Sanger sequencing to evaluate indel frequency by ICE analysis. (B) Experiment design. (C) Statistics comparing the Bcl6 indel frequency in GFP$^+$ and GFP$^-$ transfected cells. (D) Experiment design, (E) representative FACS plots and statistics showing the percentages of (F) Fas$^+$ CD38$^-$ GC B cells and (G) ABCs in transfected memory Igh^{g2A10} B cells. The results for (C, F, G) were pooled from two independent experiments. The P values were calculated by One-way ANNOVA.

We then used our optimized Cas9 KO method to study the roles of the 8 selected TFs in ABC formation in vitro plus Il21r as the positive control (**Figure 4.9A**). As expected, the KO of Il21r resulted in reduced ABC formation compared to nontarget control (**Figure 4.9B, C**). Notably, we found that in our screen only the KO of *Zeb2* reduced ABC formation (**Figure 4.9B, C**). Similar results were obtained by repeating the experiment using our PfCSP-SPZ immunization in vivo model, in which we also found only the KO of *Zeb2* impaired Igh^{g2A10} ABC formation (**Figure 4.9D-F**).

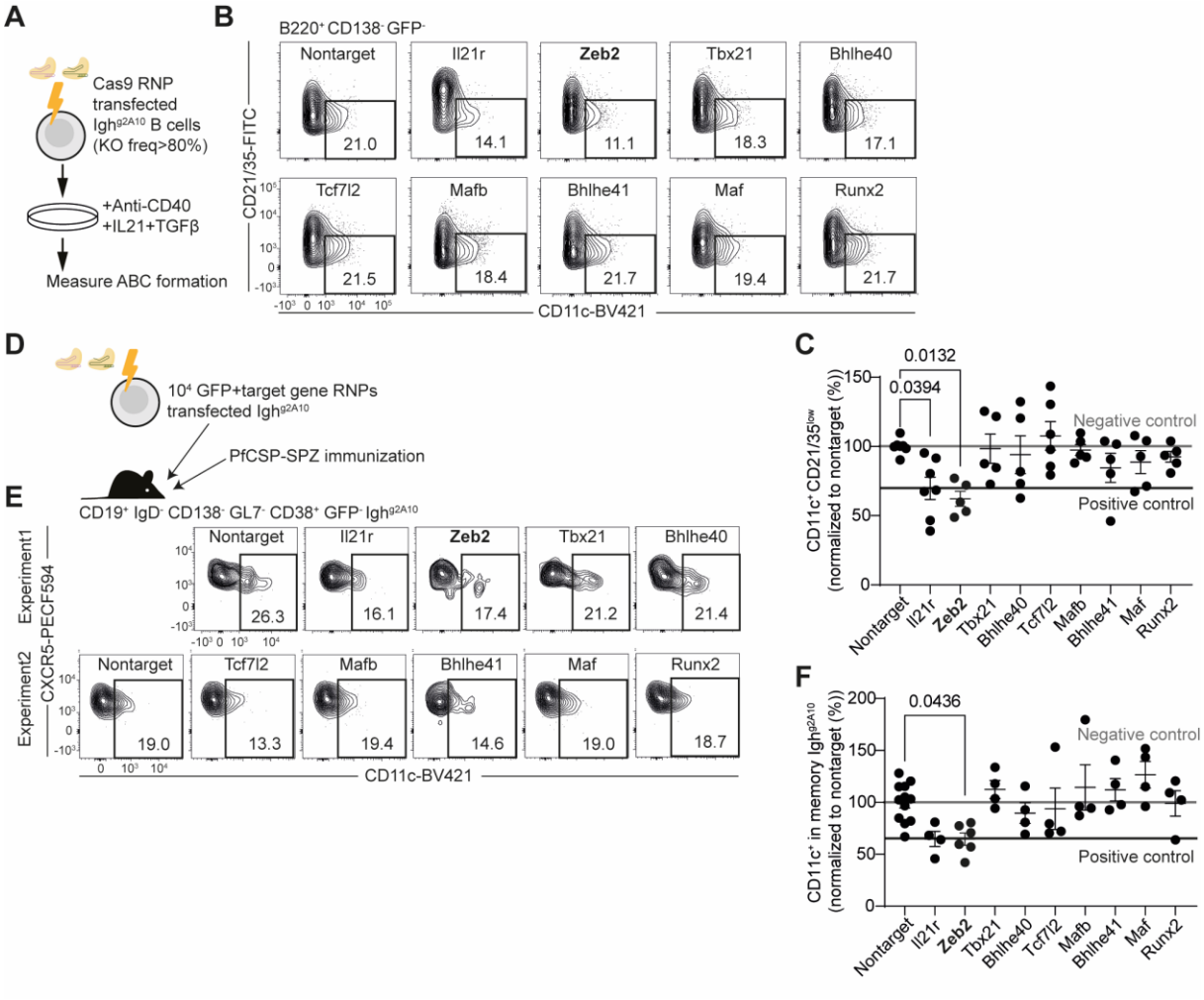


Figure 4.9 CRISPR Cas9 screen identified Zeb2 was required for ABC formation
(A-C) MACS purified B cells were pre-activated by 10ug/ml anti-CD40 for 24h, followed by Cas9-RNP nucleofection and cultured for 4 days with anti-CD40, IL-21 and TGFβ followed by FACS analysis. (A) Experiment design. (B) Representative FACS plots and (C) statistics showing the ratio of $CD11c^{+}$ $CD21/35^{lo}$ cells in $CD138^{-}$ B cells transfected by different Cas9 RNPs. (D-F) MACS enriched GFP-expressing Igh^{g2A10} B cells were pre-activated by anti-CD40, nucleofected with Cas9 RNPs and transferred into MD4 mice followed by PfCSP-SPZ immunization. (D) Experiment design, (E) representative FACS plots and (F) statistics showing the percentages of ABCs in GFP^{-} memory Igh^{g2A10} B cells. The results were pooled from four independent experiments for (C, F). Data normalization was performed in (C, F) by calculating the percentage of nontarget in each individual experiment and then the data was pooled. The P values were calculated by One-way ANOVA.

The importance of Zeb2 for ABC formation in vitro was further confirmed using *Zeb2*$^{fl/fl}$ mice (Higashi et al. 2002) crossed with a *Cd23*$^{Cre/+}$ line (**Figure 4.10A**). Zeb2 is required for pre-pro to pro B cell transition (Li et al. 2017, Scott and Omilusik 2019), therefore we chose *Cd23*$^{Cre/+}$which expresses on mature B cells to avoid B cell developmental defect. Consistent with our CRISPR/Cas9 results, we found a strong reduction in ABC formation in both heterozygous *Zeb2*$^{fl/+}$ and *Zeb2*$^{fl/fl}$ B cells compared to *Cd23*$^{Cre/+}$ control B cells (**Figure 4.10A-C**). In addition to CD11c, we also checked the expression of other ABC and cMBC markers and again found significant impairments of the up-regulation of ABC markers (Hck, CD11b, CD72, CD19) and the down-regulation of cMBC markers (CD21, CD23, CD55, CD24) by flow cytometry (**Figure 4.10D, E**) and by qPCR analysis (**Figure 4.10F**). In summary, among all the highly expressed TFs by ABCs we found only Zeb2 was required for the acquisition of the ABC phenotype and the suppression of cMBC phenotype.

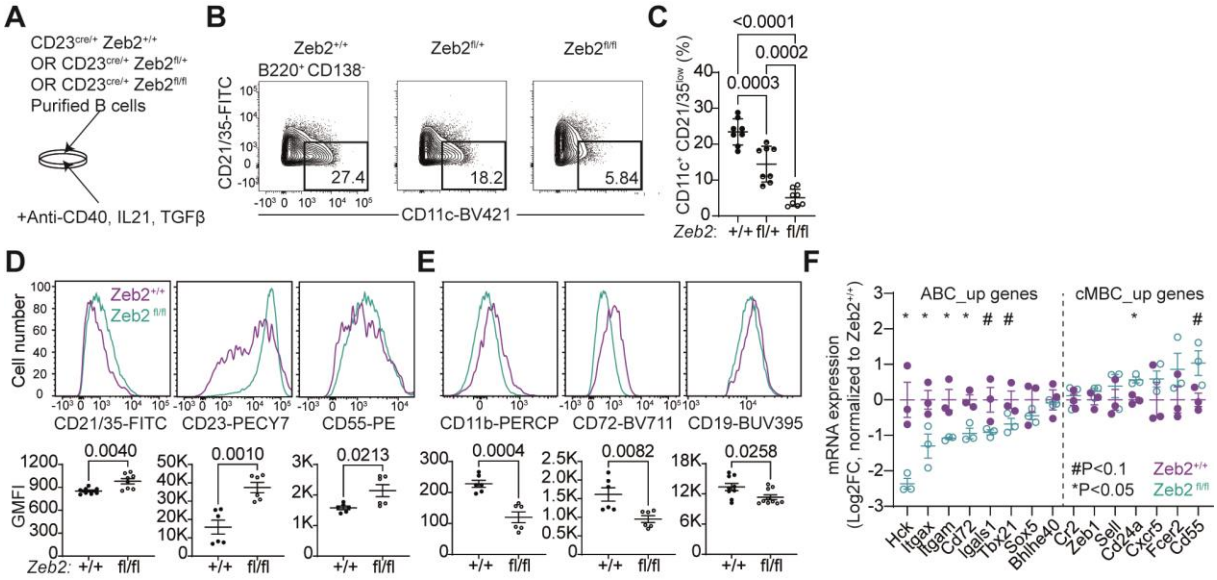


Figure 4.10 Zeb2 was required for the expression of ABC and suppression of cMBC genes
(A-C) MACS purified Cd23$^{cre/+}$ Zeb2$^{+/+,\ fl/+\ or\ fl/fl}$ B cells were cultured with anti-CD40, IL-21 and TGFβ for 4 days followed by FACS analysis. (A) Experiment design. (B) Representative FACS plots and (C) statistics showing the percentages of CD11c^{+} CD21/35^{lo} cells in CD138^{-} B cells. (D-E) MACS purified Cd23$^{cre/+}$ Zeb2$^{+/+\ or\ fl/fl}$ B cells were cultured with anti-CD40, IL-21 and TGFβ for 4 days followed by FACS analysis Representative histograms and statistics showing the GMFI of (D) cMBC and (E) ABC markers. (F) MACS purified Cd23$^{cre/+}$ Zeb2$^{+/+\ or\ fl/fl}$ B cells were cultured with anti-CD40 w/wo IL-21, TGFβ for 3 days followed by qPCR analysis. The statistic of indicated genes was showed. The results for (C, D, E) were pooled from two independent experiments. The P values were calculated by One-way ANOVA for (C) and student's t-test for (D, E, F).

Zeb2 deficiency selectively impairs the development of ABCs in mice and humans

To investigate the functional role of ABCs in vivo we infected *Cd23*$^{Cre/+}$ *Zeb2*$^{fl/fl}$ mice with *P. chabuadi* (**Figure 4.11A**). We observed a near complete abolishment of circulating ABCs in *Zeb2*$^{fl/fl}$ mice as well as a ~50% reduction in *Zeb2*$^{fl/+}$ mice compared to the *Cd23*$^{Cre/+}$ controls (**Figure 4.11B, C**). Surprisingly, no significant changes were detected in the development of other B cell lineages including cMBCs, total IgD^{-}, and PB/PCs (**Figure 4.11C**). *Zeb2* $^{fl/fl}$ mice also had comparable anti-parasite antibody titers, the elimination of parasites, and similar pathogenesis as indicated by the weight loss to the *Cd23*$^{Cre/+}$ control mice (**Figure 4.11D**). Analysis of the splenic B cells also showed a specific defect in ABC formation but no significant changes in cMBCs, total IgD^{-} B cells, PB/PC and GC B cells formation (**Figure 4.11E-H**), confirming a selective reduction of ABCs in the memory B cell population (**Figure 4.11E, F**). Collectively, the results indicate that ABCs are dispensable for the effective B cell responses to primary *P. chabuadi* infection.

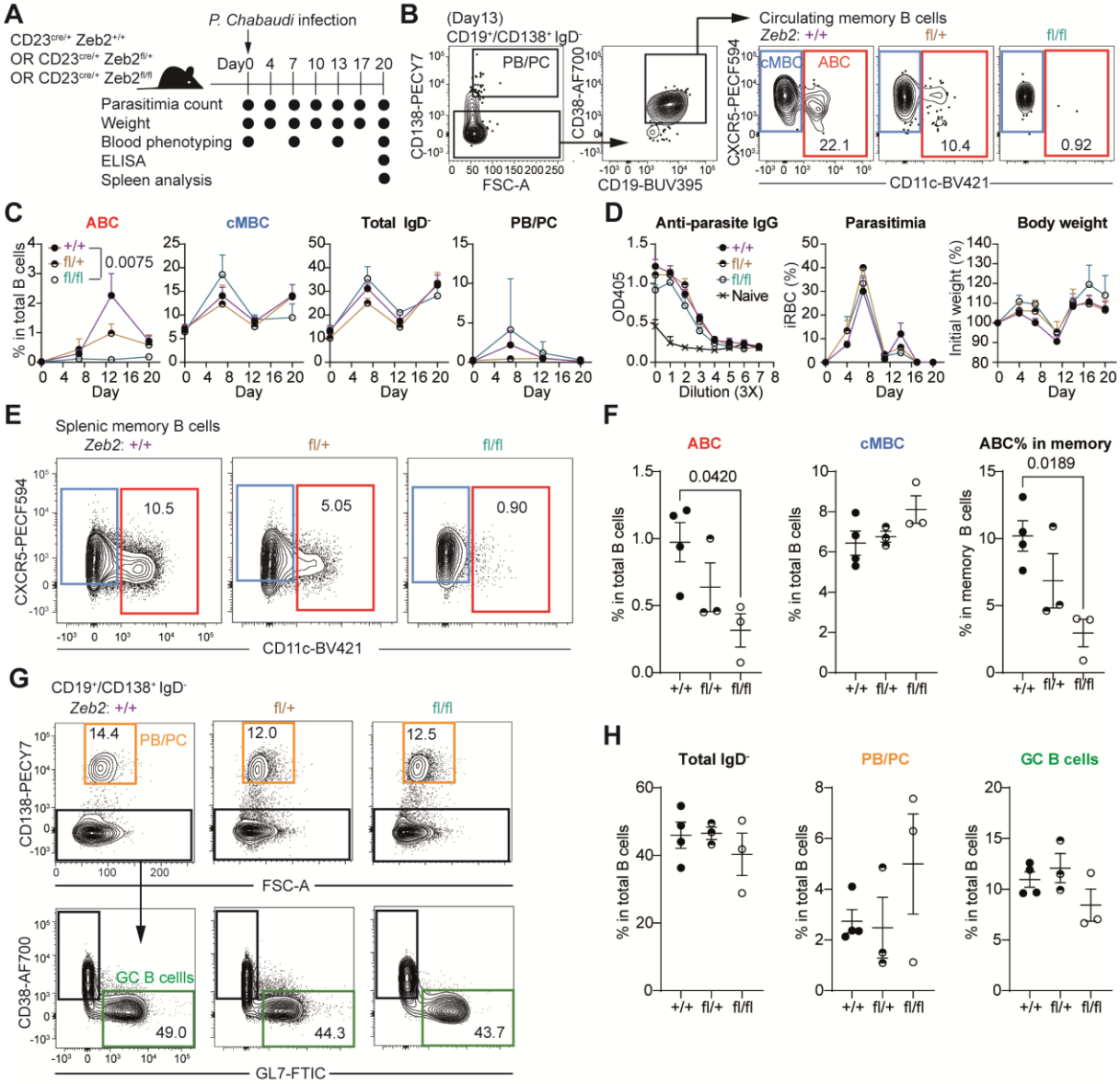


Figure 4.11 Zeb2 deficiency selectively impaired ABC formation during *P. chabuadi* infection

Cd23$^{cre/+}$ *Zeb2*$^{+/+,\ fl/+\ or\ fl/fl}$ mice were infected by *P. chabuadi* followed by a time-course study. (A) Experiment design, (B) representative FACS plots and (C) statistics showing the percentages of indicated B cell subsets in PBMC. (D) Statistics showing anti-parasite total IgG antibody titers (3-times serial diluted from 1:30), the parasitemia and body weight of mice. (E) Representative FACS plots and (F) statistics showing the percentages of splenic ABCs and cMBCs. (G) Representative FACS plots and (H) statistics showing the percentages if splenic IgD$^-$, PB/PC and GC B cells. The P values were calculated by Two-way ANNOVA for (C) and One-way ANNOVA for (F).

Because *P. chabuadi* infection causes a highly inflammatory persistent disease, we asked whether Zeb2 deficiency still selectively abolished ABC formation PfCSP-SPZ immunization. To do this, we further crossed *Cd23*$^{Cre+}$ *Zeb2*$^{fl/fl}$ mice to the Igh^{g2A10} background and transferred splenocytes from the resulting mice into MD4 animals to investigate the outcome of antigen specific B cell with Zeb2 deficiency in response to PfCSP-SPZ immunization (**Figure 4.12A**). We found Zeb2 deficient Igh^{g2A10} cells had significantly reduced ABC responses with a lower percentage of ABCs compared to mice that received *Zeb2*$^{+/+}$ *Cd23*$^{Cre/+}$ Igh^{g2A10} B cells with *Zeb2*$^{fl/+}$*Cd23*$^{Cre/+}$ Igh^{g2A10} heterozygotes having an intermediate phenotype (**Figure 4.12B, C**). However again, the numbers and proportions of other Igh^{g2A10} B cell populations including cMBCs, total IgD^{-} cells, PB/PC and GC B cells were not affected (**Figure 4.12B-E**). In summary, Zeb2 deficiency selectively impairs ABC formation without affecting other B cell populations in response to primary infection and immunization.

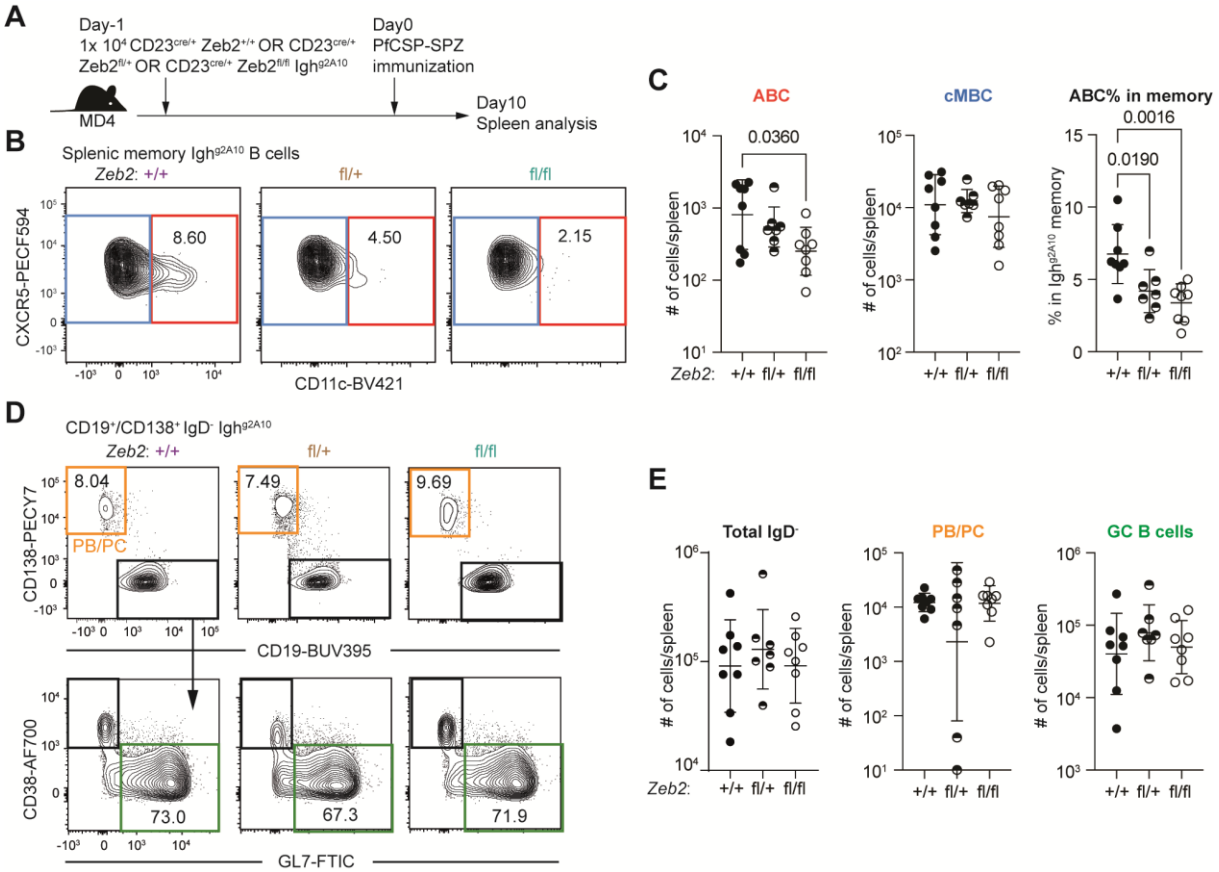


Figure 4.12 Zeb2 deficiency selectively impaired ABC formation upon PfCSP-SPZ immunization

10^4 *Cd23*$^{cre/+}$ *Zeb2*$^{+/+,\ fl/wt\ or\ fl/fl}$ Igh^{g2A10} B cells were transferred into MD4 mice followed by PfCSP-SPZ immunization, the spleens were analyzed on day10. (A) Experiment design. (B) Representative FACS plot and (C) statistics showing the formation of ABC and cMBC Igh^{g2A10} B cell populations. (D) Representative FACS plot and (E) statistics showing the formation of IgD$^-$, PB/PC and GC Igh^{g2A10} B cell populations. Results were pooled from two independent experiments for (C, E). The P values were calculated by One-way ANNOVA.

Finally, to test the role of ZEB2 in ABC formation in human we examined B cells in patients with Mowat-Wilson Syndrome (MWS). MWS is a rare genetic disease in humans (incidence: 1 per 50000-70000 live birth (Evans et al. 2012)) caused by ZEB2 haploinsufficiency resulting in developmental abnormalities. Overt immune system defects have not been reported in MWS patients and they have normal immunoglobulin levels and responses to vaccinations (Frith et al. 2021). However, previous studies also reported that MWS patients have significantly reduced percentages of $CD8^+$ T cell in lymphocytes (Frith et al. 2021) and decreased percentages of B cells in lymphocytes though this was not statistically significant (Omilusik et al. 2015). The fact that murine heterozygous *CD23*$^{Cre/+}$ *Zeb2* $^{fl/+}$ cells had reduced ABC formation both in vitro and in vivo implied that it might be possible to observe a similar phenotype in MWS patients. We therefore analyzed the PBMC samples from 5 MWS patients and 6 age/sex matched healthy controls (HCs) and calculated the percentages of ABCs as well as PB/PC and cMBCs in IgD^- B cells. We calculated the percentage in IgD^- B cells because different people might have large variation on the numbers of IgD^+ naïve B cells due to different antigen exposure history. Notably, we found MWS patients had significantly decreased ratios of ABCs in IgD^- B cells compared to HCs and the percentages of other B cell subsets in PB/PC and cMBCs appeared to be normal (**Figure 4.13A, B**). These results suggest ZEB2 was required for the human ABC formation and its defect did not affect the formation of other B cell lineages.

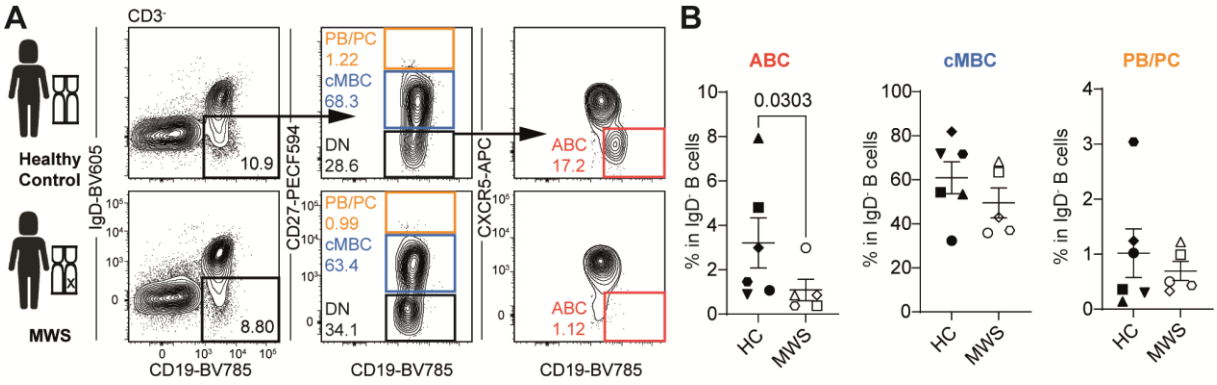


Figure 4.13 ZEB2 haploinsufficient MWS patients had reduced circulating ABCs
PBMC samples from 5 MWS patients and 6 healthy controls were FACS analyzed. (A) Representative FACS plots and (B) statistics showing the percentages of indicated B cell populations. The P values were calculated by Mann Whitney U test.

4.3 Discussion

In this study we identified Zeb2 as the lineage defining TF for ABCs. While Zeb2 clearly is required for the expression of ABC genes and the suppression of cMBC genes, an outstanding question is through what molecular mechanism that Zeb2 controls the expression of these genes. More specifically, whether Zeb2 expression alone is sufficient to drive ABC-like features, or requires cooperation with other TFs like Tbet, needs to be further investigated. Intriguingly, a recent preprint study suggests Zeb2 can directly bind to the promoter region of CD11c and CD11b and promote ABC formation through Jak-Stat pathway (Gu et al. 2021). Consistent with this, Zeb2 KO bone marrow cells also have reduced Stat3 and Stat5 phosphorylation in response to IL-3 and IL-6 treatment (Li et al. 2017). However, whether Jak-Stat acts upstream or downstream of Zeb2 to promote ABC needs to be confirmed. We have found that IL21 treatment in vitro could promote Zeb2 expression and the immediate downstream signaling cascade of IL-21 receptor is through Stat1 and Stat3 (Leonard et al. 2008). In the immune system, Zeb2 has been reported to play roles in the development of lymphocytes including plasmacytoid dendritic cells (DCs) (Scott et al. 2016, Wu et al. 2016), classical DC2 (Scott et al. 2016, Wu et al. 2016), NK cells (van Helden et al. 2015) and terminal differentiated effector CD8 T cells (Dominguez et al. 2015, Omilusik et al. 2015). Interestingly, in $CD8^+$ T cells the Zeb2 expression is dependent on T-bet whereas both our qPCR data and the preprint study suggest that Zeb2 acts independently of Tbet in B cells. Moreover, in contrast to B cells, Zeb2 KO in macrophages promotes the expression of CD11c and CD11b. Therefore, it seems the spectrum of Zeb2-regulated genes is highly context dependent and cannot be inferred from other cell types. Another interesting observation is the reciprocal expression of Zeb2 and Zeb1 in ABCs vs cMBCs. Zeb1 and Zeb2 can suppress the expression of each other in $CD8^+$ T cells (Guan et al. 2018), therefore, it is interesting to further study the role of Zeb1 in ABC formation.

Phenotypically, we showed Zeb2 KO-mediated ABC depletion had no significant impact on overall B cell responses upon immunization, infection in mice and natural antigen exposure in human. Therefore, the original function of ABCs is likely neither of the current hypotheses. If the cells were exhausted-like cell to limit the immune activation we would have expected to observe an enhanced response. If the cells were PB/PC precursors or specialized antigen presenting cells (especially priming Tfh cells), we would have expected to observe a decreased response. Despite our result suggesting ABCs might be an exhausted-like non-functional cell type, we acknowledge the models included in our study might be insufficient because these are both primary immunization and infection models, whereas ABC might play a unique role in

repeated immunization or persistent infection. Intriguingly, in our tonsil scRNA-seq dataset the expression of ABC signature genes was inversely correlated with the SHM frequency, and in HIV infected individuals and LCMV infected mouse ABCs were generated outside of the germinal center where SHM is not favored (Austin et al. 2019, Song et al. 2022), therefore, we reasoned that ABCs might be a quality control mechanism of the host, especially during chronic infection where large numbers of memory B cells are generated, allowing for reduced life-span and functionality of low affinity memory B cells, thereby resulting in improving the average affinity of long-lived memory B cells. In light of this, B cell specific Zeb2 KO mice might be more susceptible to re-infection because their memory B cell pool is crowded with poor affinity BCRs.

Chapter5- The protective B-cell and cognate T-cell epitopes are spatially restricted in response to large parasite invasion

5.1 Introduction

Efficient antibody production requires B cells to uptake and present antigens to the T cells. After engagement with the B-cell receptor (BCR), antigens are taken up by B cells through distinct mechanisms depending on the antigen properties (e.g., size, soluble or membrane bound) including direct endocytosis, phagocytosis and immune synapse formation (Batista and Harwood 2009). However, this understanding is established on the basis of studying antigens such as immunoglobulin-bound immune complexes (Suzuki et al. 2009), membrane-bound antigens (Fleire et al. 2006), antigen-coated beads (Martinez-Riano et al. 2018), and viruses and bacteria (Parra et al. 2012, Hong et al. 2018), whereas no literature has investigated how B cells acquire and present antigens from large parasites, despite this high prevalence of parasite (e.g. *Plasmodium* and helminth) infection humans and the ability of these infections to elicit humoral responses (Harris and Gause 2011, Seder et al. 2013, Weiss et al. 2019, Sartorius et al. 2021). More importantly, because antigens used by these previous studies are predominately nanometer-sized which can be internalized by B cells as a whole, it is of interest to investigate whether B cells use different mechanisms to take up antigens from large parasites (e.g., *Plasmodium* sporozoite, >10 μm in length), a scenario in which the complete uptake of the pathogen by B cells seems unlikely.

A key question is the restriction of T cell help for B cell responses. If B cells are capable of phagocytosing small pathogens they may potentially obtain help from T cells specific for a broad range of parasite antigens. In support of this following small virus infection (e.g. vesicular stomatitis virus (Junt et al. 2007), influenza virus (Scherle et al. 1986)) B cells are capable to take up the whole pathogen. So all antigen specific Tfh cells, regardless of their epitopes, can help B cells. However, whether these findings can be extended to larger pathogens remains unknow. Interestingly, in studies with vaccinia virus (large DNA virus, 200-300 nm in diameter) whose genome encoding >200 genes, $CD4^+$ T cell help to B cells was nontransferable to other virion protein specificities, suggesting that the free individual vaccinia viral proteins, rather than the whole virion, are the physiological B cell antigens (Sette et al. 2008). Although previous studies on virus have significant helped us to understand the collaboration of B and T cells during infection, to date, we have very limited knowledge about the antigen specificity of cognate B and T cells during large parasite infection. Whether antigen specific B cells can be helped by all antigen specific Tfh cells, or only by a subset of T cells targeting a few parasite antigen epitopes remains a major outstanding question. This is important because such knowledge will help us to focus on those functional T cell epitopes to improve malaria vaccines.

In particular it may be that if B cells can be helped by T cells with a range of specificities this may results in more sustained GC reactions and ultimately better B cell responses.

In this research, we aim to investigate how antigen specific B cells uptake antigens from *Plasmodium* sporozoites for presentation to cognate T cells. Here, using a variety of approaches including direct imaging of B cells-parasite interactions, the assessment of Tfh cell repertoires after immunization and the development of transgenic parasites expressing model T cell antigens in different spatial locations, we show that antigen specific B cells could uptake the surface but not the interior antigens from the SPZs in a "surface pinching" manner, resulting in the cognate T cell epitopes are spatially restricted to the surface antigens of SPZs.

5.2 Results

Igh^{g2A10} B cells pinch off CSP from SPZs

To find out how antigen specific B cells acquire antigen(s) from the SPZs, we incubated WT or Igh^{g2A10} B cells with PfCSP-SPZs that express GFP internally and examined the surface expression levels of PfCSP by flow cytometry (**Figure 5.1A**). As expected, a significantly higher proportion of Igh^{g2A10} B cells were GFP^{+} compared to WT cells indicating that they directly bound sporozoites in an antigen specific manner (**Figure 5.1B)**. If the SPZs were completely phagocytosed by the B cells, we would expect the complete abolition of the PfCSP signal by surface staining, however, we found the CSP levels were lower, but not altogether abolished in on SPZ-binding Igh^{g2A10} B cells compared to WT B cells (**Figure 5.1C, D**), suggesting that Igh^{g2A10} B cells were unlikely to phagocytose the whole SPZs but might specifically extract PfCSP or potentially internalize PfCSP along with other SPZs surface antigens. To visualize this process, we performed image flow cytometry analysis on PfCSP-SPZ-interacting B cells (**Figure 5.1A**). In this analysis we included antibodies to Igβ in order to visualize the location of the BCR and the formation of any immunological synapse. We found that on WT B cells, there was no co-localization of the PfCSP or GFP signal with BCR as indicated by Igβ staining (WT example 1, 2, 3) (**Figure 5.1E**). In contrast, the BCR of Igh^{g2A10} cells was colocalized strongly PfCSP (Igh^{g2A10} example 1, 2), and in one case we found the B220 and Igβ appeared to have spread onto the PfCSP-SPZ (Igh^{g2A10} example 3) (**Figure 5.1E**). A more formal analysis showed significantly increased co-localization scores of Igβ with CSP in Igh^{g2A10} compared to WT B cells (**Figure 5.1F**). However, there was no localization of Igβ with GFP, suggesting that Igh^{g2A10} B cells gather and extract the surface PfCSP from PfCSP-SPZs, however we did not see evidence of uptake of internally expressed GFP (**Figure 5.1F**).

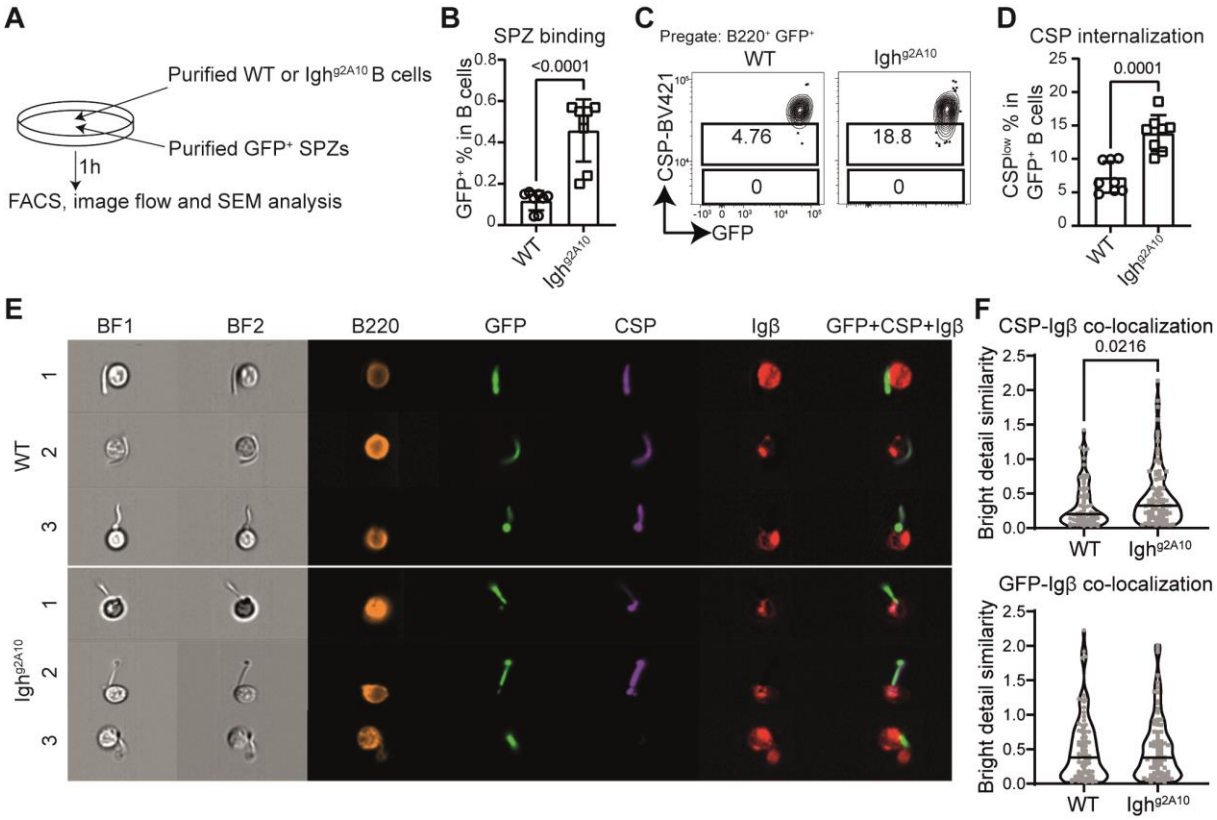


Figure 5.1 Igh^{g2A10} B cells preferentially internalise CSP from SPZs
10^4 Igh^{g2A10} B cells were incubated with sorted GFP$^+$ PfCSP-SPZs in complete RPMI for 1h, followed by flow cytometry and image flow cytometry analysis. (A) Experiment design. (B) Statistics of the frequencies of SPZ-bound B cells. (C) Representative FACS plot and (D) statistics showing the surface staining of PfCSP on SPZ-bound B cells. (E) Representative images showing the bright field (BF1/2) and the expression of indicated antigens. (F) Statistics comparing the co-localization scores of CSP-Igβ and GFP-Igβ. The scores were calculated by the build-in co-localization function of IDEAS software. The results of (B, C, D) were pooled from three independent experiments, of (F) were two independent experiments. The P values in (B, D, F) were calculated by Student's t test.

To further understand this antigen uptake process, we performed similar experiments but used scanning electronic microscopy (SEM) to visualize the B cell-SPZ interaction. For WT B cells, we found all the PfCSP-SPZs appeared loosely attached to them (**Figure 5.2A**). In contrast, we found (6/18 SPZ binding) Igh^{g2A10} B cells spread their cell membrane onto PfCSP-SPZs to form a close interaction (**Figure 5.2B**). In Example 1 & 2, around 30%-50% of the surface area of the SPZs were completely covered by the Igh^{g2A10} B cell membrane, while in Example 3 only <10% surface area of the SPZs was covered by B cell membrane. It is unclear whether BCR affinity contributes to such differences. Nevertheless, these results suggest Igh^{g2A10} B cells cannot phagocytose the PfCSP-SPZs but may pinch off surface antigen PfCSP from SPZs.

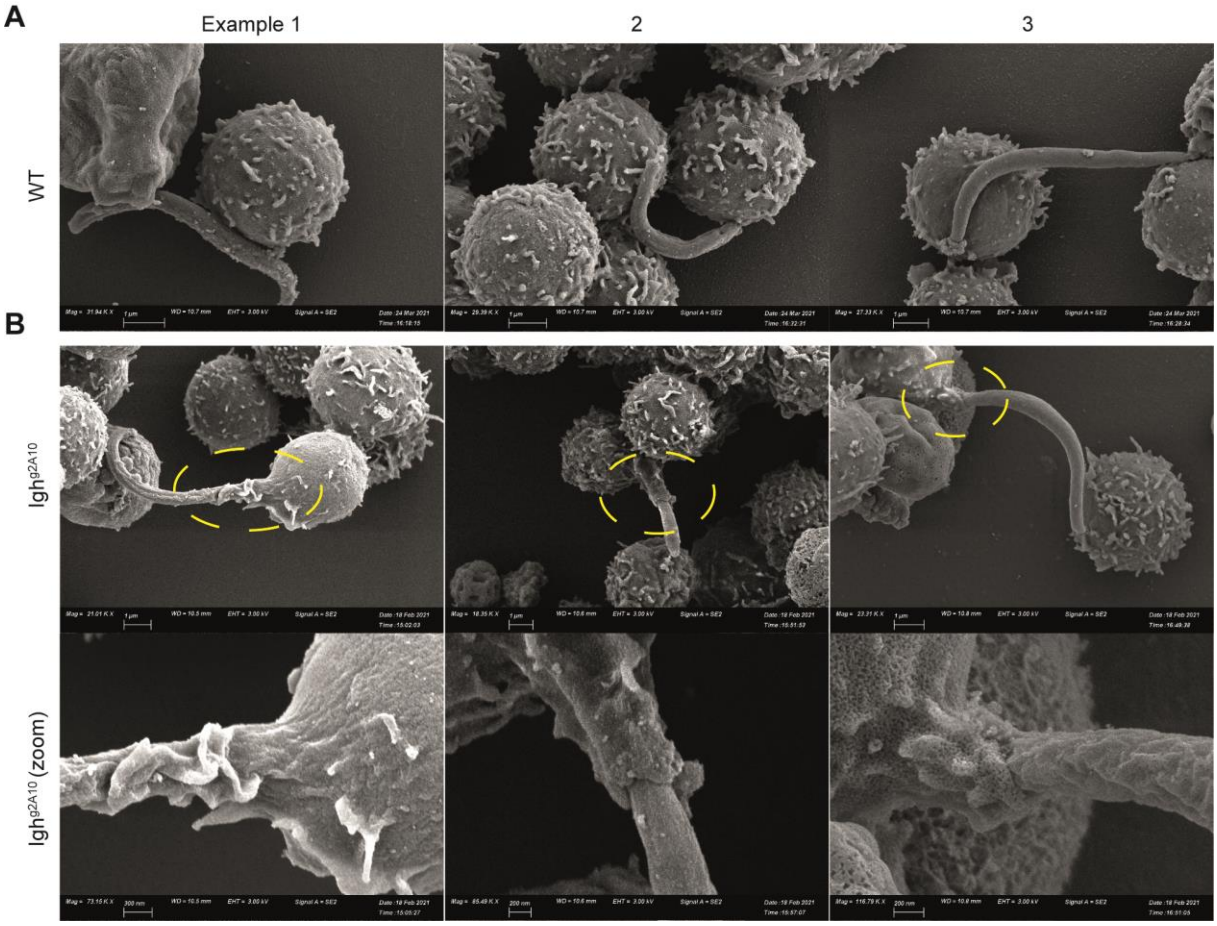


Figure 5.2 Scanning electron microscopy to visualize the B cell-SPZ interaction
10^4 Igh^{g2A10} B cells cells were incubated with sorted GFP$^+$ PfCSP-SPZs in complete RPMI for 1h and fixed by 1% glutaraldehyde for SEM analysis. Representative SEM pictures showing the interaction between PfCSP-SPZs and (A) WT or (B) Igh^{g2A10} B cells. The results were representative of two independent experiments.

Igh^{g2A10} B cells were predominately helped by PfCSP-specific Tfh cells upon SPZ immunization

The "surface pinching" manner might result in Igh^{g2A10} B cells only taking up PfCSP in a highly specific process, or it may permit the uptake of other SPZ surface antigens. In the first scenario we would expect all helping T cells to be PfCSP specific, however in the second scenario a more diverse range of helper T cells would be observed. To test these hypotheses we therefore measured the TCR repertoire of Tfh cells that are helping Igh^{g2A10} B cells upon PfCSP-SPZs immunization and compared this to the range of TCRs responding to immunization with recombinant PfCSP.

Specifically, we transferred Igh^{g2A10} B cells into MD4 mice followed by recombinant PfCSP (protein; Group 1 CSP) or PfCSP-SPZ (whole parasite; Group 2 SPZ) immunization and sorted the bulk Tfh cells on day10 post immunization for 10X scRNA-seq and TCR-seq (**Figure 5.3A**). Because the recipient MD4 mice only express irrelevant BCRs that were specific to hen egg lysozyme, all the Tfh cells should be maintained by the transferred Igh^{g2A10} B cells. We included 2 additional control groups: 1) mice that were immunized by PfCSP-SPZ but had no Igh^{g2A10} B cells transferred (Group3 NoB), and 2) untreated mice (Group4 NT), which could be used to identify the background TCR clones (**Figure 5.3A**).

Unsupervised clustering of all Tfh cells revealed 4 major clusters, conventional Tfh, memory Tfh, regulatory Tfh and proliferating Tfh (**Figure 5.3B**) with the expression of signature genes Bcl6, Sell, Foxp3 and Mki67 respectively (**Figure 5.3C**). Reclustering on conventional Tfh cells revealed more subsets including PreTfh1/2, GCTfh1/2 (marked by GC Tfh markers Ascl2, IL21 and Hif1a) (Liu et al. 2014, Cho et al. 2019), activated Tfh1/2 (Nfkbid) and an unassigned "SPZ Tfh" population that were induced by SPZ immunization (**Figure 5.3D, E**). As expected, CSP and SPZ groups had more GC and activated Tfh cells than NT and NoB groups (**Figure 5.3F**).

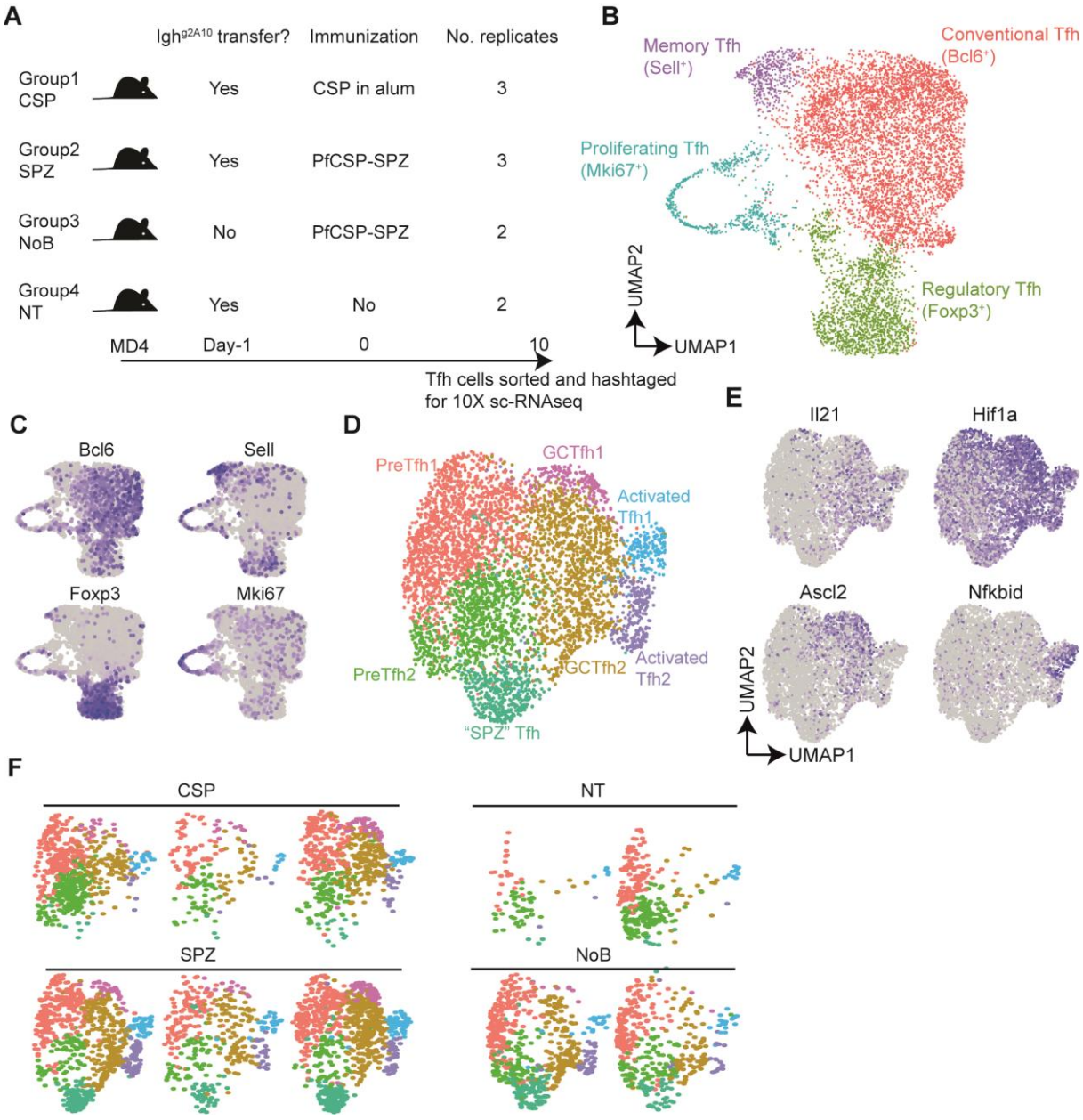


Figure 5.3 Unsupervised clustering of Tfh cells

MD4 mice were transferred w/wo 10^4 Igh^{g2A10} B cells, left untreated or immunized with recombinant PfCSP-alum or PfCSP-SPZs. CD4$^+$ CXCR5$^+$ PD-1$^+$ Tfh cells were sorted on day10 post treatment, hashtaged and pooled for 10X scRNA-seq analysis. The raw sequencing data was processed by Cellranger pipeline and the output was analysed by Seurat and scReptoire R packages. (A) Experiment design. (B) UMAP of all Tfh cells. (C) Feature plots showing the expression of indicated signature genes. (D) The conventional Tfh cluster was reclustered and analyzed. The UMAP of reclustered conventional Tfh cells. (E) The feature plots showing the expression of markers genes of conventional Tfh cell sub-clusters. (F) The UMAP of reclustered conventional Tfh cells of each mouse.

To investigate TCR repertoire composition in response to different vaccinations, we analyzed the clonal diversity of each group of mice based on different algorisms. As expected, NT mice had no clonal selection stress thus had the highest clonal diversity. The diversity of NoB cell group sit between CSP/SPZ and NT in general. There was no clear difference between CSP and SPZ groups (**Figure 5.4A**), despite the fact SPZ are estimated to express ~3000 genes and are supposed to induce a much more diverse TCR repertoire than CSP-alum. Even assuming the possibility that B cells may take up only sporozoite surface antigens, mass-spectrometry studies suggest that at least 42 proteins are highly likely to be present on the sporozoite surface (Swearingen et al. 2016). These results suggest Igh^{g2A10} B cells only captured a few antigens, or perhaps only PfCSP, during SPZs immunization.

To understand the antigen specificity of the Tfh cells, we ranked the TCR alpha chain based on their abundance and studied the 10 most abundant alpha chain usage. We focused on alpha chain because structurally it has more contacts with MHC-II loaded peptide than beta chain, therefore the peptide recognition predominately depends on alpha chain (Yokosuka et al. 2002). We found the top1 alpha chain, TRAV10J31, was highly enriched in CSP and SPZ groups compared with NoB and NT groups (**Figure 5.4B**), and there was no significant difference in the relative abundance of this clone between CSP and SPZ groups (**Figure 5.4C**), suggesting that Igh^{g2A10} helping Tfh cells were predominately PfCSP specific.

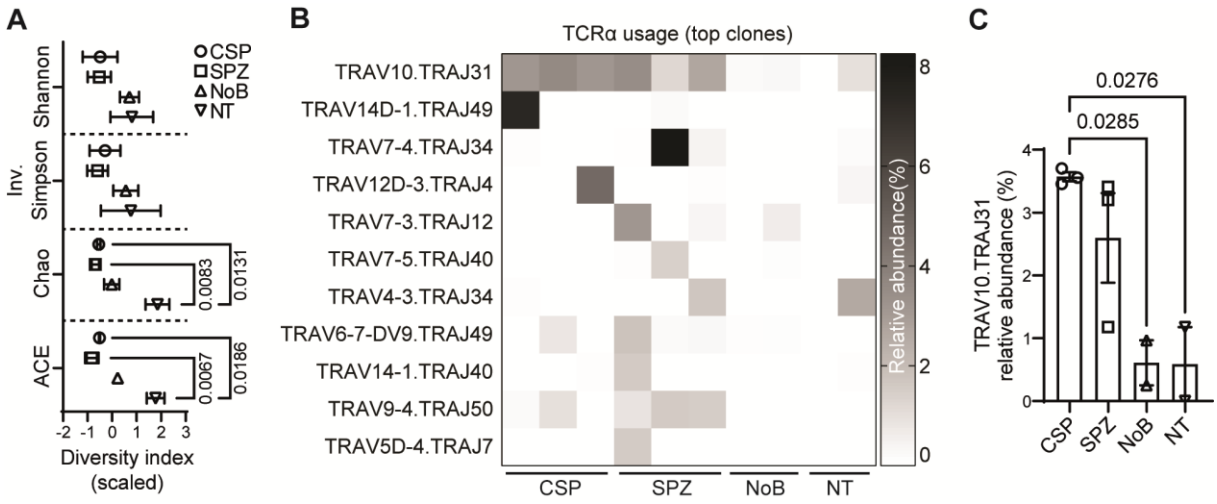


Figure 5.4 The TCR analysis of Tfh cells

MD4 mice were transferred w/wo 10^4 Igh^{g2A10} B cells, left untreated or immunized with recombinant PfCSP-alum or PfCSP-SPZs. $CD4^+$ $CXCR5^+$ $PD\text{-}1^+$ Tfh cells were sorted on day10 post treatment, hashtaged and pooled for 10X scRNA-seq analysis. The raw sequencing data was processed by Cellranger pipeline and the output was analysed by Seurat and scReptoire R packages. (A) The clonal diversity of all Tfh cells in each group of mice based on 4 different algorisms. (B) The heatmap showing the relative abundance (number of cells with this clone/ total cell number *100%) of the top10 most abundant alpha chain usage in all conventional Tfh cells. (C) The statistics showing the relative abundance of the TRAV10J31 clones in each mouse. The P values in (C) were calculated by Two-way ANOVA.

To confirm the antigen specificity of the TRAV10J31 clone was PfCSP, we further analyzed the associated beta chains and found a highly restricted beta chain usage with the majority using TRBV26 paired with different J genes (**Figure 5.5A**). We then cloned the TRAV10J31 alpha chain and TRBV26DJ2-1 beta chain into pES4 and p3A9cbTCR vectors respectively (Barnden et al. 1998), and co-microinjected the linearized DNA fragments into fertilized eggs to generate TCR transgenic mice (**Figure 5.5B**). After screening, we acquired one founder strain named CST-II. Flow cytometry analysis of $CD4^+$ T cells from CST-II mice showed the successful expression of TRBV26 (Vβ3) (**Figure 5.5C**). In the absence of a TRAV10 antibody, we analyzed the expression of a major V gene Vα2 and found its expression was reduced in CST-II mice, suggesting successful expression of the TRAV10 transgene (**Figure 5.5C**). In vitro culture of the CST-II splenocytes with recombinant PfCSP protein induced significant activation of CST-II $CD4^+$ T cells (**Figure 5.5D**). The PfCSP molecule can be divided into 4 regions: N-terminus, minor repeat (NANPNVDP)n, major repeat (NANP)n and C-terminus. To determine the domain level specificity of CST-II cells, we incubated CST-II splenocytes with peptides corresponding to each domain and found CST-II cells were specific to major NANP central repeat region with at least 6 repeats (**Figure 5.5D**). In contrast the cells were unable to respond to the minor repeat region which has itself been identified as HLA-DQ restricted $CD4^+$ T cell epitope in humans (Wahl et al. 2022). In sum, these results suggest Igh^{g2A10} B cells mainly take up PfCSP upon PfCSP-SPZ immunization, although at this stage we could not exclude the possibility that Igh^{g2A10} B cells also uptake other bystander SPZs antigens especially those present on the SPZs surface.

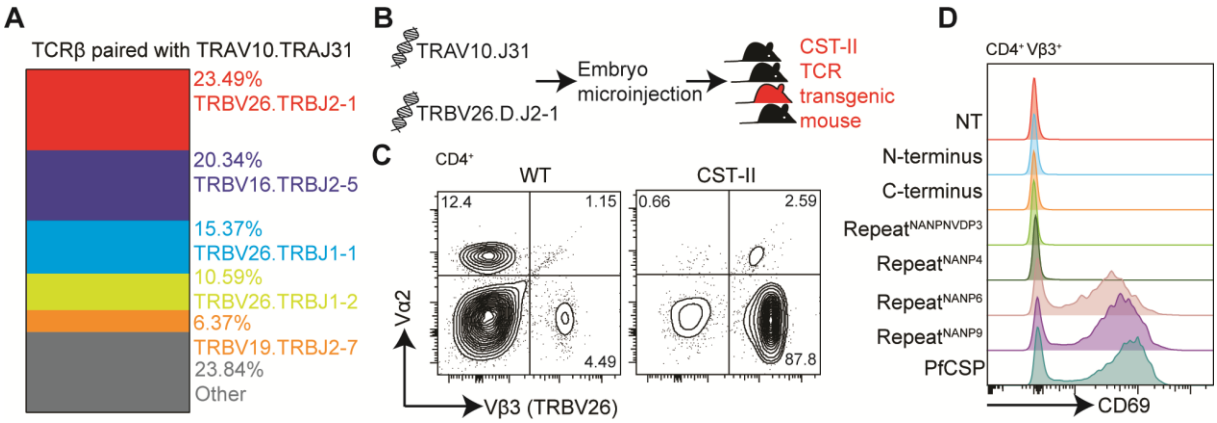


Figure 5.5 The generation of PfCSP-specific TCR transgenic mice (CST-II)
(A) Statistics showing the percentages of beta chains that were paired with TRAV10J31. (B) The original TRAV10J31 and TRBV26DJ2-1 sequences from scRNA-seq were cloned into the pES4 and p3A9cbetaTCR plasmids respectively. The plasmids were digested, purified and microinjected into fertilized eggs. One transgenic founder (CST-II) was identified by PCR and the transgenes were confirmed that could be passed to the offspring. (C) The PBMCs were isolated from the CST-II mice and stained by indicated antibodies. Representative FACS plots were showed. (D) The splenocytes of CST-II mice were cultured for 24h in complete RPMI in the presence w/wo 1ug/ml indicated antigen, and the expression of CD69 was measured. Indicated FACS plots were showed. Results of (C, D) were representative of >5 experiments.

Igh^{g2A10} B cells took up surface but not interior antigens from PfCSP-SPZs

Our previous imaging and scRNA-seq analysis suggested Igh^{g2A10} B cells predominately extracted PfCSP from the PfCSP-SPZ, however, it is unclear whether they could also acquire other SPZ antigens, given that B cells are known to capture the bystander molecules that are co-presented on the membrane with cognate antigen (Suzuki et al. 2009). Therefore, we aimed to test whether Igh^{g2A10} B cells could also be helped by Tfh cells that were specific to spatially different SPZ antigens.

To find out whether Igh^{g2A10} B cells could take up the SPZ interior antigens, we used PbT-II cells which were specific to the heat shock protein 90 (Hsp90) expressed internally by SPZs (**Figure 5.6A**) (Enders et al. 2021). We transferred Igh^{g2A10} B cells w/wo irrelevant T cells, (OT-II), PbT-II or CST-II cells into CD28KO mice followed by PfCSP-SPZ immunization to examine the formation of Igh^{g2A10} GC B cells (**Figure 5.6B**). We used CD28KO mice as recipients because they have no endogenous CD4^{+} T cell responses, and all T cell help to Igh^{g2A10} B cells will only be provided by the transferred transgenic T cells. Consistent with our imaging and scRNA-seq data suggesting Igh^{g2A10} B cells cannot take up the interior SPZ antigens, PbT-II cells were incapable to support the formation of Igh^{g2A10} GC B cells upon PfCSP-SPZ immunization (**Figure 5.6C, D**), despite of the efficient priming and proliferation of PbT-II cells (**Figure 5.6E, F**).

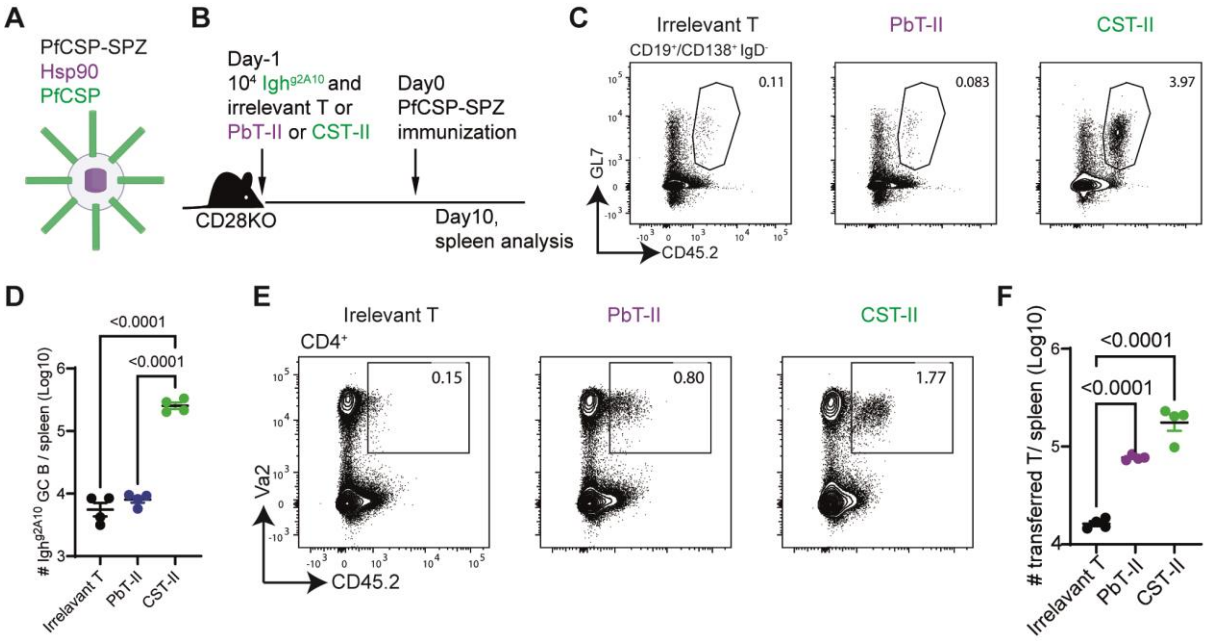


Figure 5.6 PbT-II cells were incapable to help Igh^{g2A10} B cell responses upon PfCSP-SPZ immunization

10^4 Igh^{g2A10} B cells and either 10^5 OT-II (irrelevant T), PbT-II or CST-II cells were transferred into CD28KO mice followed by PfCSP-SPZs immunization. (A) Schematics of PfCSP-SPZ. (B) Experiment design. (C) Representative FACS plots and (D) statistics showing the formation of Igh^{g2A10} GC B cells under indicated conditions. (E) Representative FACS plots and (F) statistics showing the numbers of indicated transferred T cells. The results were representative of two independent experiments. The P values were calculated by One-way ANOVA in (E).

One limitation of the previous experiment is that the PbT-II epitope is expressed in the interior of the sporozoite. However it is also possible that $CD4^+$ T cells that are specific to the bystander surface antigens may provide help to B cells. To test this hypothesis, we took advantage of the availability of parasites that express two copies of CSP on the surface of the cells: PfCSP and the endogenous PbCSP. We transferred CST-II cells into CD28KO mice and immunized them by such SPZs co-expressing PfCSP and PbCSP to examine the anti-PbCSP antibody response (**Figure 5.7A, B**). In this scenario the transferred CST-II cells, which only target PfCSP but not PbCSP, are specific for the bystander surface antigen. Surprisingly, we found CST-II cells efficiently supported humoral response to PbCSP (**Figure 5.7C, D**), conversely mice that received control (OT-II cells) were not able to generate PbCSP specific antibody responses suggesting that PbCSP-specific B cells could take up the bystander PfCSP during SPZ immunization and obtain help from CST-II cells.

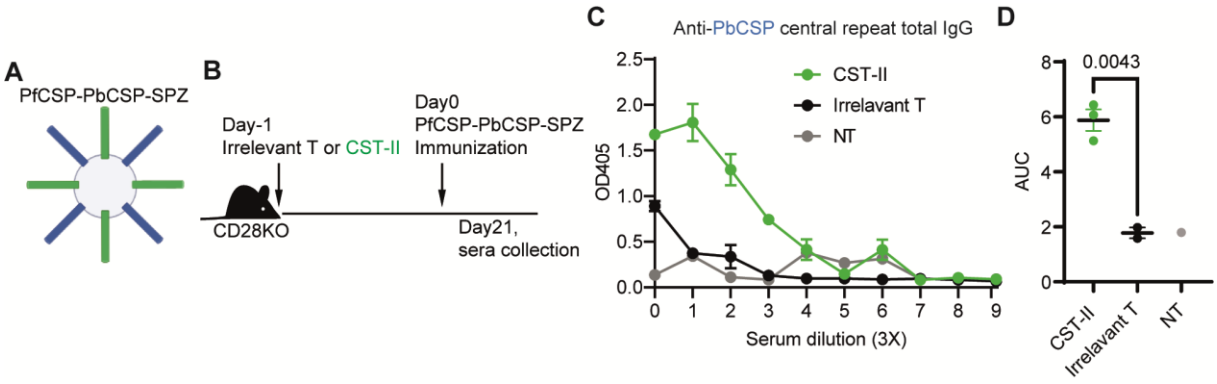


Figure 5.7 Anti-PbCSP humoral responses were supported by CST-II cells that were specific to the bystander surface antigen PfCSP

CD28KO mice were transferred by either 10^5 OT-II (irrelevant T) or CST-II cells, followed by PfCSP-PbCSP-SPZs immunization. The sera were collected on day21 and the anti-PbCSP central repeat region total IgG was measured by ELISA. (A) Schematics of PfCSP-PbCSP-SPZ. (B) Experiment design and (C-D) statistics showing the antibody titers. The P value was calculated by One-way ANOVA.

Our previous experiments demonstrated that surface antigen specific B cells could also take up bystander surface antigens but not internal antigens. However, such conclusion might be limited. Due to the use of different transgenic $CD4^+$ T cells, our results can also be interpreted as the CST-II cells have a stronger preference to form Tfh cells over PbT-II cells due to their intrinsically different TCR signalling strength. Therefore, we wanted to generate parasites in which the same T cell epitope was expressed in spatially different locations of the SPZs. A particular challenge in this system is that there is no established method for targeting antigens to the sporozoite surface. However we took advantage of the fact that the sporozoite can tolerate having two copies of CSP on the surface (Singer and Frischknecht 2021), and we successfully expressed OVA on the surface with SPZs by fusion it with CSP, using the well-established transgenic technique for *Plasmodium berghei* (Janse et al. 2006) which has been previously used to generate GFP expressing SPZs. In this experiment in addition to using our standard PfCSP-SPZ (**Figure 5.8A**), we generated three transgenic SPZ strains that expressed OVA on spatially different locations in the SPZs: 1) PfCSP-$OVA^{mcherry}$-SPZ, which expressed $OVA^{mcherry}$ internally (**Figure 5.8B**) (Lin et al. 2014). 2) PfCSP-$PbCSPOVA^{flag}$ -SPZ, which expressed OVA^{flag}-PbCSP fusion protein on the surface (**Figure 5.8C)**. 3) PbCSP-$PfCSPOVA^{flag}$-SPZ, which expressed OVA^{flag} inside the PfCSP molecule (**Figure 5.8D**). The identities of these parasites were validated by flow cytometry using PfCSP-SPZs as the control (**Figure 5.8A-D**).

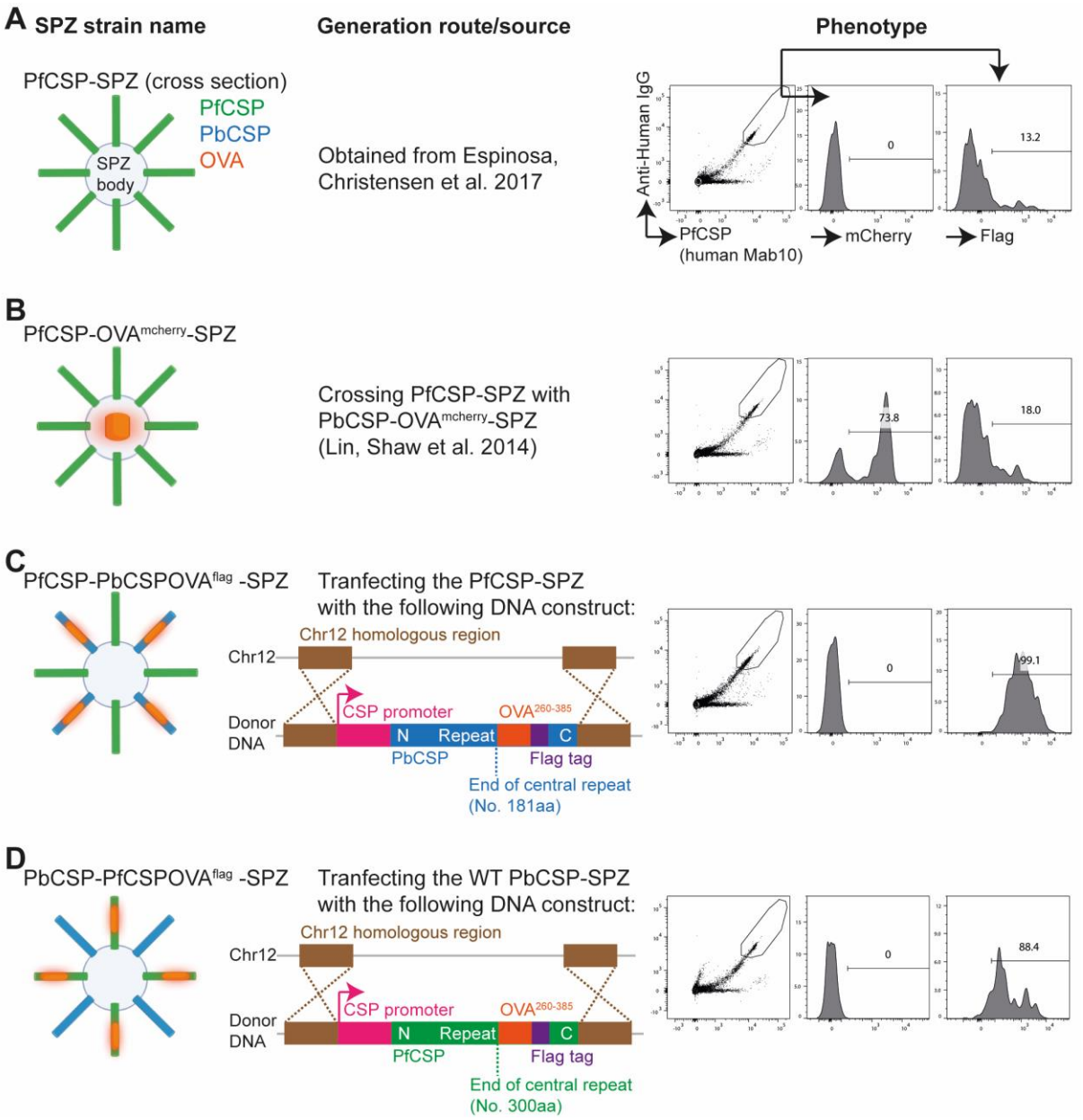


Figure 5.8 The generation and phenotyping of the transgenic SPZs

Transgenic SPZs were generated and dissected from mosquito salivary glands, stained by indicated antibodies and analyzed by flow cytometry. Schematics of the antigens expression, generation routes and the phenotypes of the SPZs were showed for: (A) PfCSP-SPZ, (B) PfCSP-OVA^{mcherry}-SPZ, (C) PfCSP-PbCSPOVA^{flag}-SPZ and (D) PbCSP-PfCSPOVA^{flag}-SPZ. The results were representative of three independent experiments.

We transferred Igh^{g2A10} B cells and OT-II cells into CD28KO mice followed by immunization with each type of the parasites to look at the formation of Igh^{g2A10} GC B cells (**Figure 5.9A**). We also included no T cell transferred and CST-II cell transferred groups for each parasite strains as endogenous control, given that the yield of SPZs had great variation between batches and it was unlikely to keep the total dose of PfCSP consistent between all parasite strains in one experiment. As expected, the transfer of CST-II significantly enhanced GC Igh^{g2A10} B cells responses after immunization with all parasites (**Figure 5.9B-C)**. However, OT-II cells efficiently helped Igh^{g2A10} B cells only when the OVA was located on the surface (PfCSP-PbCSPOVA^{flag} and PbCSP-PfCSPOVA^{flag}) of the SPZ (**Figure 5.9B-C)**. Moreover, OT-II cells were inferior than CST-II when OVA was expressed in PbCSP while comparable with CST-II when expressed inside PfCSP (**Figure 5.9C)**. Altogether, our result suggests the helper function of OT-II cells in relate to the OVA location in SPZs is: same surface antigen> bystander surface antigen >> interior antigen. Of note, the lack of Igh^{g2A10} GC B cells formation in response to PfCSP-OVA^{mcherry}-SPZ was not caused by inefficient OT-II cells activation and expansion (**Figure 5.9D-E**). These results demonstrated that Igh^{g2A10} B cells could uptake PfCSP and bystander surface antigens but not interior antigens of SPZs. In sum, by different models we demonstrated that antigen specific B cells can uptake surface but not interior antigens from the SPZs and present to cognate T cells to form GC reaction.

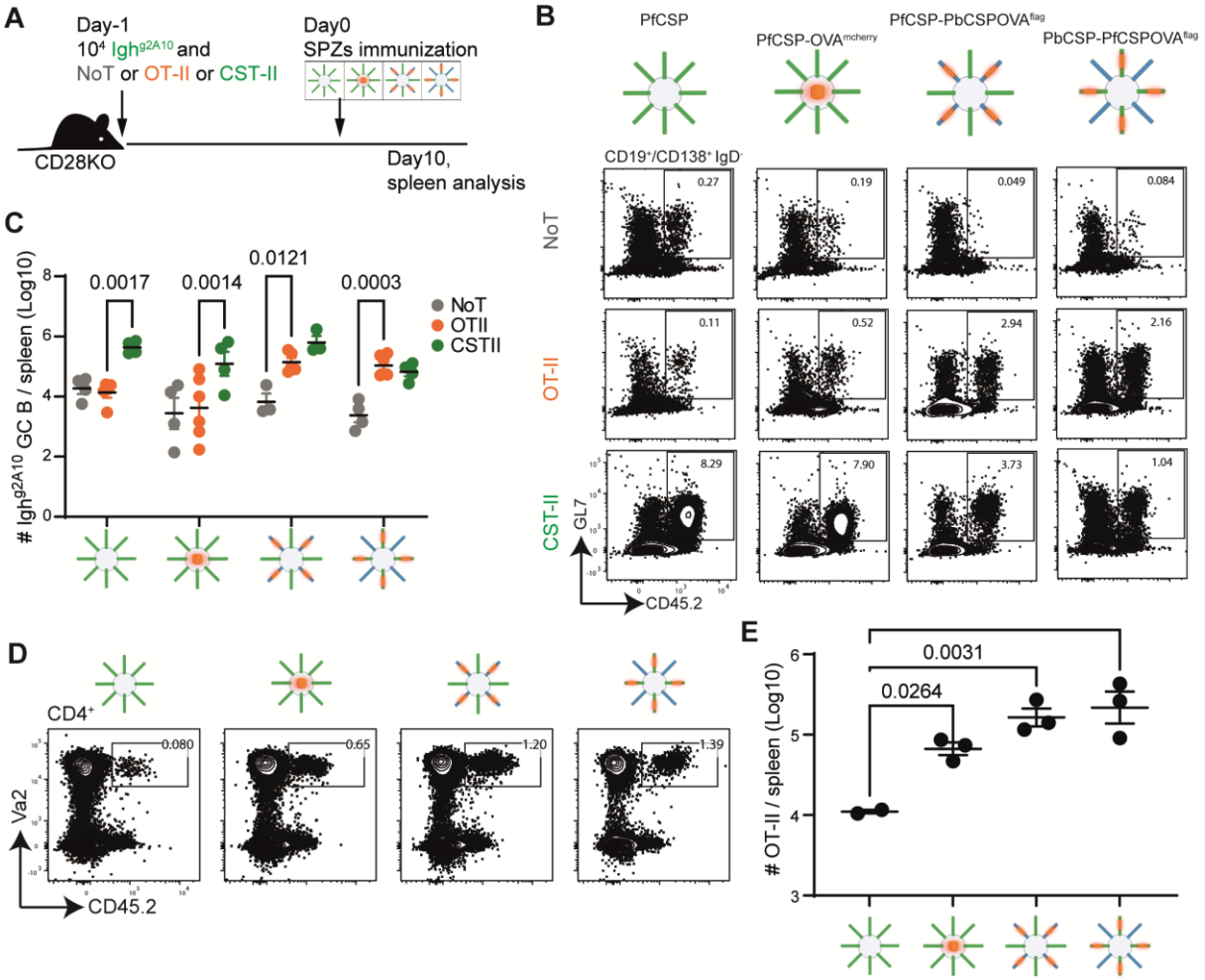


Figure 5.9 OT-II cells were only capable to help Igh^{g2A10} B cells when OVA was located on the SPZ surface

10^4 Igh^{g2A10} B cells w/wo 10^5 OT-II or CST-II cells were transferred into CD28KO mice followed by indicated SPZs immunization. (A) Experiment design. (B) Representative FACS plots and (C) statistics showing the formation of Igh^{g2A10} GC B cells under indicated conditions. (D) Representative FACS plots and (E) statistics showing the numbers of OT-II cells. The results were pooled from two independent experiments for (C) and representative for two experiments in (E). The P values were calculated by Two-way ANOVA and in (C) and by One-way ANOVA in (E).

5.3 Discussion

Here we demonstrated that PfCSP-specific Igh^{g2A10} B cells deployed a "surface pinching" mechanism to uptake PfCSP and bystander surface antigens from the SPZs, resulting in the specificity of cognate Tfh cells were restricted to the SPZ surface antigens. One outstanding question is when Igh^{g2A10} B cells bind to the SPZs, by which molecular mechanism that they pinch off and internalize these antigens. Upon BCR binding, B cells are known to use myosin IIa–mediated contraction force to pinch off the membrane-bound antigens (Natkanski et al. 2013). Additionally, lysosomes are secreted by B cells around the immune synapse to further facilitate the antigen extraction (Yuseff et al. 2011). Therefore, it is interesting to find out whether these mechanisms are also applied during B cell and SPZ interaction. Once 'pinching off' the antigens, B cells are known to internalize the antigen by clathrin mediated endocytosis, caveolin mediated endocytosis, endophilin mediated endocytosis and phagocytosis (McShane and Malinova 2022). Given that there are established small molecule drugs that can specifically inhibit these pathways (Roper et al. 2019), it is of interest to use these drugs to study the molecular mechanism of B cell antigen uptake.

Our analysis of the Igh^{g2A10} helping Tfh cell epitope suggests that B and T cells can successfully form cognate interaction when their corresponding antigen epitopes are co-localized on the surface of the SPZs, and do not necessarily have to be in the same molecule. Our result seems to contradict to a previous study using vaccinia virus infection model and showed that the efficient B and T cell interaction required both B and T cells were specific to the same virion antigen (Sette et al. 2008). However, such difference may be caused by the intrinsic features of the pathogens. While the irradiated SPZs are structurally intact, during vaccinia virus infection, the lymph nodes might be enriched with separated viral antigens due to the killing of $CD8^+$ T cells and the release of unassembled viral antigens from infected cells. Further work should use a variety of different pathogens to unveil the general rule of the B-T cell epitope relationship in response to large pathogen infection. Perhaps the best theory is related to the spatial proximity, the closer of the B and T cell epitopes, the more likely of the corresponding cells to establish an efficient cognate interaction.

We showed that Igh^{g2A10} B cells could directly extract surface antigens from SPZs, however, it remains to be further studied whether native sporozoites directly interact with B cells in vivo, or whether SPZ antigens are presented to B cells through dendritic cells or macrophages, given that these cells can phagocytose SPZs (Radtke et al. 2015) and can also present naïve antigens

to B cells (Batista and Harwood 2009). Upon immunization with particulate antigens, subcapsular sinus macrophages are proposed to 'filter' these antigens and release them into the B cell follicle. The released antigens are transferred by bystander B cells onto the follicular dendritic cells (FDC) for long-term deposition (Carrasco and Batista 2007). Therefore, we can speculate that Igh^{g2A10} B cells might also acquire SPZ antigens from subcapsular sinus macrophages and FDCs. However, it seems unlikely that a bystander B cell could 'drag' an intact SPZ and deposit it onto the FDC, thus it will be interesting to look at the forms of the SPZ antigens in vivo (e.g. partially digested by the macrophage and deposited onto the FDCs) by immunofluorescence staining.

Additionally, SPZ immunization can induce GC responses >60 days (McNamara et al. 2020). Whether the 'surface pinching' mechanism still apply to the B cells (predominately GC B cells) at this extended stage is of interest to find out. Notably, GC B cells have different immune synapse structure with naïve B cells (Nowosad et al. 2016, Kwak et al. 2018). While naïve B cells gather antigens towards the center of immune synapse before endocytosis, GC B cells form multiple small antigen clusters before endocytosis. Therefore, where (e.g. directly from SPZs or FDCs) and how (e.g. surface pinching like naïve B cells) GC B cells acquire SPZ antigens is an interesting follow up question. Additionally, GC is a locus where B cells undergo SHM to increase antibody affinity, whether B cells with higher affinity are more capable to uptake the bystander surface antigens and obtain the help from Tfh cells is an interesting future direction.

At last, whole attenuate SPZs has been used in human trials and demonstrated promising protection against control malaria infection in malaria-naïve adults (Seder et al. 2013). However, recent phase I clinical trials conducted in African adults showed efficacy only around 48% to prevent malaria infection throughout the wet season (Sissoko et al. 2017). Moreover, recent phase II trials in infants failed to confer protection against malaria infection within the 6 months follow-up period (Oneko et al. 2021). These results raise the question whether our 'default' humoral response against whole SPZ is suboptimal. Intriguingly, SPZ immunization induces strong early plasmablast responses which may dampen efficient humoral responses against malaria (Vijay et al. 2020, McNamara et al. 2022). Of note, the formation of these early plasmablast is mainly T cell independent, suggesting the lack of cognate T cell help at this stage (McNamara et al. 2022). Based on the our 'surface pinching' observation, there might be an imbalance between the numbers of surface antigen specific T cells and B cells, given that SPZs are estimated to express over 3000 genes and most them should be internally, while only a few

surface antigen specific T cells are capable to help the B cells. Provided that enhancing the Tfh response, for example by knocking out inhibitory receptors CTLA4, promotes the expansion of GC B cells (Abbott et al. 2017). Thus a priming immunization with peptide pool containing the multiple T cell epitopes of the SPZ surface antigens might increase the numbers of Tfh cells, reduce the early plasmablast formation, and prolong the following GC reaction and affinity maturation, result in memory B cells and long-lived plasma cell generation with higher affinity to the sporozoite antigens. Therefore diversification of the T cell epitopes located on the surface of SPZs might be helpful to improve the efficacy of the whole parasite vaccines.

Chapter6- Discussion

The general topic of my thesis is to understand the collaboration of T and B cell in response to vaccination, and this topic is targeted from three aspects: the Tfh cells, the ABCs and the antigen presentation of B cells. Based on these findings, outstanding mechanistic questions and translational implications will be discussed.

6.1 Understand the mechanism of T cell persistence and to improve the longevity of vaccine-induced protection

The results in Chapter3 demonstrated that Tfh17 cells had better immunological memory persistence over Tfh1 and Tfh2 cells. However, it is unclear why Tfh17 cells have a central-memory like phenotype and have survival and proliferative advantages. Of note, Th17 cells also have such central-memory features, or 'stem-like' features as suggested by other papers (Kryczek et al. 2011, Muranski et al. 2011, Schnell et al. 2021), therefore, the mechanism is very likely to be associated with shared features with Th17 cells. Th17 cells are known to express higher levels of Il7ra, which potentially triggers stronger IL-7-Bcl2 signaling to promote their survival (Arbelaez et al. 2015). Recent studies in the EAE model also identified a population of $Tcf7^+$ stem-like Th17 cells that can replenish proinflammatory Th17 cells under stimulation. Therefore, to understand the factors that drive the central-memory features of Th17 cells is crucial to optimize vaccine regimens to induce long-lived memory $CD4^+$ T cells.

It has been well-documented that the choice of adjuvants can modulate Th1/2/17 and Tfh1/2/17 responses (He et al. 2013), an extended question is whether a Th17-polarized vaccination strategy should be favored in future vaccine design. In the light of the Covid-19 pandemic, several novel vaccine formulations (mRNA and adenovirus based) have been introduced into humans which provide protection in the initial 3-6 months observation period (Polack et al. 2020). However, it is now clear that these vaccines are poorly protective against breakthrough infection with the emerging SARS-CoV2 variant strains (e.g. BA.5) though they are still effective against severe symptoms (Aggarwal et al. 2022, Kirsebom et al. 2022). Intriguingly, although the binding affinity of vaccine induced antibodies against SARS-CoV2 variants are reduced 3.9, 8.8 and at least 10-fold for the Delta (B.1.167.2), Beta (B.1.351) and Omicron (B.1.1.529) strains respectively, T cell responses to the variants are largely preserved (Kent et al. 2022), which might explain why these Covid-19 vaccines still prevent severe symptoms (Kirsebom et al. 2022). Nevertheless, these findings highlight the crucial role of vaccine induced antigen specific T cells.

However, current licensed Covid-19 mRNA vaccines induce strong Th1-polarized responses with very few Th17 cells (Bettini and Locci 2021), which raises concerns regarding whether the longevity of these vaccines is sub-optimal. However, unlike antibodies whose protective role can be easily evaluated, the protection induced by memory T cells is hard to assess in humans (Kent et al. 2022), and no approved vaccines are reported to induce Th17-biased responses, therefore future studies comparing the longevity and protective role of Th1, Th2 or Th17-adjuvanted vaccines should be conducted to elucidate the most proper type of T cell responses. Another limiting factor we should consider is the availability of adjuvants, because to date the majority of adjuvants that are developed and studied trigger Th1 or Th2 responses. Only CFA and CAF01 are reported to induce Th1/Th17 responses in mice but these adjuvants are highly inflammatory and are unlikely to be used in humans (Coffman et al. 2010).

Another implication of this study is the potential pathogenic role of autoreactive Th17 and Tfh17 cells in the chronic inflammatory diseases. Because these cells are long-lived and retain proliferation potential, they might contribute to the disease relapses. In support of this, Th17 cells have a high degree of plasticity and can trans-differentiate into Th1, Th2, and Tfh cells (McGeachy 2013, Wu et al. 2018). Moreover, in autoimmune patients, Th17 cells are often correlated with disease activity (McGeachy 2013). Thus, while Th17-polarization is preferred in vaccination to induce long-lived memory cells, it should be used with caution in patients with autoimmune disorders.

6.2 The development and functions of ABCs, and its possible role in "gate-keeping" the antibody affinity

ABCs are found not only enriched in patients with chronic infection and autoimmunity but also present in healthy individuals, suggesting that ABCs are not "atypical" but a normal and conserved B cell population. With the emerging of scRNA-seq technique, this has been further confirmed that a cluster of ABCs are always identified in unsupervised clustering of peripheral memory B cells (Gao and Cockburn 2022). A key aim of this study was to determine whether these ubiquitous cells have a functional role in healthy immune responses. However, while we were able to show that ZEB2 deficient mice lack ABCs, this absence had no significant impact on the general humoral response, making it difficult to interpret the normal function of this population.

Understanding the origin of ABCs may be helpful to understand their role in immune system. During LCMV infection, ABCs were found to be localized around the T-B border and marginal zone areas, and they had low SHM rate, suggesting they were predominately generated from the interaction with extrafollicular T cells (Song et al. 2022). Consistent with this, both our scRNA-seq data and human lymph node analysis suggests ABCs have lower levels of somatic hypermutation than cMBCs (Austin et al. 2019). The association of ABCs with low SHM rate and likely the poor BCR affinity might explain ABCs are enriched in autoreactive BCR clones (Isnardi et al. 2010). However, it is unclear why the extrafollicular region favors the generation of ABCs. According to our in vitro result, anti-CD40 and IL21 are sufficient to drive ABCs formation, however both stimuli are common T cell derived signals which can also promote the generation of PC and GC B cells. One hypothesis is when B cells are stimulated by cognate T cells but their BCR signaling is weak, they will likely form ABCs. On the contrary, if they receive stronger BCR signaling, they will become PC or GC B cells. This theory can explain the association of ABC with lower SHM rate. Co-transferring transgenic B cells with low and high antigen affinity and determining their ability to form ABC populations upon immunization or infection can address this hypothesis.

Given that ABCs also have exhausted like features with poor survival and poor proliferation potential upon antigen recall responses, another hypothesis is that the natural function of ABC might be a method of the host to turn low affinity B cells into ABCs, to allow for their eliminated from the immune system. If this is true, the antibody affinity but not the titer will be altered in Zeb2 deficient mice during immunization or infection. A systemic analysis of the SHM frequency of B cells in Zeb2 KO mice can test this hypothesis. Moreover, it has been argued that ABCs are pathogenic in the context of autoimmune disease, mainly because they are enriched for autoantibody clones and correlate with disease progression (Isnardi et al. 2010, Jenks et al. 2018, Wang et al. 2018). Although these findings seem to contradict the hypothesis that ABCs are exhausted cells, to date, there is insufficient evidence to support the view that "ABCs are drivers of autoimmunity". For example, there is no viable method that can specifically deplete ABCs, thus how can we draw the conclusion of "ABCs can drive autoimmunity" till we can carry out the proper experiment to deplete ABCs in autoimmune animal models? Additionally, the link of ABCs to PCs-precursors are generally based on in vitro culture experiments in the presence of TLR7/9 agonist which is non-physiological. Therefore, although many papers are claiming that ABCs are pathogenic for autoimmunity, I am not convinced of it and believe ABCs might instead have a protective role. Based on my

result, ABCs might play a protective role in autoimmunity because otherwise they will become either PC or GC and thus exacerbate disease progression. In conclusion, testing the outcome of Zeb2 deficiency in autoimmune animal models is a critical future experiment.

ZEB2 haploinsufficiency causing MWS is mostly manifest via developmental defects, while the immune system of these patients appears to be largely normal, which is consistent with our result in Zeb2 deficient mouse upon immunization and infection. However, there are case reports showing that MWS patients are susceptible to bacterial infection like Streptococcus pneumoniae (Nevarez Flores et al. 2019), and some patients require the injection of immunoglobulin due to recurrent infection (Adam et al. 2019). Therefore, it is likely that the ABCs are required for optimal humoral response in human. However, ZEB2 plays key roles in other immune cell subsets including effector $CD8^+$ T cells and type 2 conventional dendritic cells (Scott et al. 2016, Wu et al. 2016, Dominguez et al. 2015, Omilusik et al. 2015), so any immune deficiencies may be related to these other defects. Due to the extremely rare occurrence of MWS, it might be difficult to conduct clinical trials but a systemic review of these patients' history, especially the report of any types of infection, might be helpful to understand the role of ABCs in otherwise healthy humans.

6.3 Insights in improving the malaria vaccine

In the past two decades several vaccines candidates were designed against malaria. These include vaccines targeting the pre-erythrocytic stage of *Plasmodium falciparum* and *vivax*, the erythrocytic stage and sexual stage of *Plasmodium falciparum* (Duffy and Patrick Gorres 2020). Because the species *falciparum* is the major cause of severe malaria symptoms, and there is lack of evidence suggesting vaccine against the erythrocytic stage is effective in human trials (Duffy and Patrick Gorres 2020), I will focus on the vaccine targeting the pre-erythrocytic stage of *Plasmodium falciparum*.

It has been showed that the anti-CSP antibody titer that is required to confer 50% protection against *Plasmodium falciparum* infection in human is 121 EU/ml (White et al. 2015). Of note, this is a very high antibody titer compared to that of most infectious diseases such as measles and tetanus. As a result, it is difficult to maintain the protective level of anti-CSP antibody for more than a year without further booster vaccination (Cockburn and Seder 2018). Therefore, vaccine that only targets the CSP molecule seems unlikely to offer long-lived protection. The

attempts should be made to increase the spectrum of antigens that are targeted by a given malaria vaccine.

In theory, whole parasite vaccine should offer superior protection than subunit vaccine because it contains all the antigen epitopes that are expressed on the sporozoite (SPZ). Moreover, it elicits liver tissue resident $CD4^+$ and $CD8^+$ T cells which should offer additional layer of protection (Lefebvre and Harty 2020). However, clinical trial results in African adults found that the efficacy of this vaccine was at best comparable to subunit vaccine candidates RTS, S/AS01 and R21 Matrix-M (Sissoko et al. 2017), challenging the idea that broader antigen epitope offers better protection. Intriguingly, plasmablasts induced by attenuated SPZ in humans were 30-50% CSP specific (McNamara et al. 2020), suggesting that SPZ induced humoral response is dominated by anti-CSP antibody. This scenario is probably caused by the membrane pinching mechanism described in Chapter5 which results in the immunodominant CSP specific B cells "using up" all the opportunities to interact with cognate T cells and repelling other minor B cell clones out of the GC reaction. Therefore, reducing the quantities of CSP on the surface might be helpful to avoid CSP domination and to induce a more diverse humoral response (Chatterjee et al, 2021).

Another attempt should be made is to characterize antigens that are presented on the surface of SPZ during invasion into hepatocytes, because targeting these antigens might be more effective to block the infection of the parasite than targeting CSP. Interestingly, recently a scRNA-seq study demonstrated that SPZ started to express a different transcriptomic profile when exposed to the mammal host environment within as few as 1h (Real et al. 2021). Perhaps the transcriptomes of SPZs in the liver is very different from the SPZs localize in the lymphoid organs where B cells recognize the SPZs, causing the antibodies that B cells produce miss the important epitopes that are crucial for hepatocyte invasion. Therefore, future experiment should compare the proteomics of SPZs isolated from the salivary glands with SPZs that are invading the hepatocytes to search for potential vaccine targets.

6.4 Concluding remarks

Understanding the T and B cell collaboration in response to infection and vaccination provides important knowledge in the development of vaccines. In my thesis, Tfh17 cells were reported to favor the longevity of immunological memory, while $CD11c^+$ atypical B cells had a redundant role in vaccination, suggesting modulating the production of these cell populations

could be a strategy to optimize our current vaccines. The analysis of the TCR repertoire upon large parasite immunization further shows the choice of T cell epitope may affect the antibody production during whole pathogen immunization. In sum, this project has revealed several key information regarding the T and B cell collaboration in response to immunization and infection.

Chapter7- References

Abbott, R. K., M. Silva, J. Labuda, M. Thayer, D. W. Cain, P. Philbrook, S. Sethumadhavan, S. Hatfield, A. Ohta and M. Sitkovsky (2017). "The GS Protein-coupled A2a Adenosine Receptor Controls T Cell Help in the Germinal Center." J Biol Chem **292**(4): 1211-1217.

Adam, M. P., J. Conta and L. J. Bean (2019). "Mowat-wilson syndrome."

Aggarwal, A., A. Akerman, V. Milogiannakis, M. R. Silva, G. Walker, A. O. Stella, A. Kindinger, T. Angelovich, E. Waring, S. Amatayakul-Chantler, N. Roth, S. Manni, T. Hauser, T. Barnes, A. Condylios, M. Yeang, M. Wong, T. Jean, C. S. P. Foster, D. Christ, A. C. Hoppe, M. L. Munier, D. Darley, M. Churchill, D. J. Stark, G. Matthews, W. D. Rawlinson, A. D. Kelleher and S. G. Turville (2022). "SARS-CoV-2 Omicron BA.5: Evolving tropism and evasion of potent humoral responses and resistance to clinical immunotherapeutics relative to viral variants of concern." EBioMedicine **84**: 104270.

Akiyama, M., K. Suzuki, K. Yamaoka, H. Yasuoka, M. Takeshita, Y. Kaneko, H. Kondo, Y. Kassai, T. Miyazaki, R. Morita, A. Yoshimura and T. Takeuchi (2015). "Brief Report: Number of Circulating Follicular Helper 2 T Cells Correlates With IgG4 and Interleukin-4 Levels and Plasmablast Numbers in IgG4-Related Disease." Arthritis & Rheumatology **67**(9): 2476-2481.

Albershardt, T. C., B. M. Iritani and A. Ruddell (2012). "Evaluation of reference genes for quantitative PCR analysis of mouse lymphocytes." J Immunol Methods **384**(1-2): 196-199.

Ambegaonkar, A. A., K. Kwak, H. Sohn, J. Manzella-Lapeira, J. Brzostowski and S. K. Pierce (2020). "Expression of inhibitory receptors by B cells in chronic human infectious diseases restricts responses to membrane-associated antigens." Science advances **6**(30): eaba6493.

Ambegaonkar, A. A., S. Nagata, S. K. Pierce and H. Sohn (2019). "The Differentiation in vitro of Human Tonsil B Cells With the Phenotypic and Functional Characteristics of T-bet+ Atypical Memory B Cells in Malaria." Front Immunol **10**: 852.

Aranburu, A., N. Hook, N. Gerasimcik, B. Corleis, W. Ren, A. Camponeschi, H. Carlsten, O. Grimsholm and I. L. Martensson (2018). "Age-associated B cells expanded in autoimmune mice are memory cells sharing H-CDR3-selected repertoires." Eur J Immunol **48**(3): 509-521.

Arazi, A., D. A. Rao, C. C. Berthier, A. Davidson, Y. Liu, P. J. Hoover, A. Chicoine, T. M. Eisenhaure, A. H. Jonsson, S. Li, D. J. Lieb, F. Zhang, K. Slowikowski, E. P. Browne, A. Noma, D. Sutherby, S. Steelman, D. E. Smilek, P. Tosta, W. Apruzzese, E. Massarotti, M. Dall'Era, M. Park, D. L. Kamen, R. A. Furie, F. Payan-Schober, W. F. Pendergraft, 3rd, E. A. McInnis, J. P. Buyon, M. A. Petri, C. Putterman, K. C. Kalunian, E. S. Woodle, J. A. Lederer, D. A. Hildeman, C. Nusbaum, S. Raychaudhuri, M. Kretzler, J. H. Anolik, M. B. Brenner, D. Wofsy, N. Hacohen, B. Diamond and S. L. E. n. Accelerating Medicines Partnership in (2019). "The

immune cell landscape in kidneys of patients with lupus nephritis." Nat Immunol **20**(7): 902-914.

Arbelaez, C. A., S. Glatigny, R. Duhen, G. Eberl, M. Oukka and E. Bettelli (2015). "IL-7/IL-7 Receptor Signaling Differentially Affects Effector CD4+ T Cell Subsets Involved in Experimental Autoimmune Encephalomyelitis." J Immunol **195**(5): 1974-1983.

Arstila, T. P., A. Casrouge, V. Baron, J. Even, J. Kanellopoulos and P. Kourilsky (1999). "A direct estimate of the human αβ T cell receptor diversity." Science **286**(5441): 958-961.

Attaf, M., E. Huseby and A. K. Sewell (2015). "alphabeta T cell receptors as predictors of health and disease." Cell Mol Immunol **12**(4): 391-399.

Austin, J. W., C. M. Buckner, L. Kardava, W. Wang, X. Zhang, V. A. Melson, R. G. Swanson, A. J. Martins, J. Q. Zhou and K. B. Hoehn (2019). "Overexpression of T-bet in HIV infection is associated with accumulation of B cells outside germinal centers and poor affinity maturation." Science translational medicine **11**(520): eaax0904.

Aye, R., H. J. Sutton, E. W. Nduati, O. Kai, J. Mwacharo, J. Musyoki, E. Otieno, J. Wambua, P. Bejon, I. A. Cockburn and F. M. Ndungu (2020). "Malaria exposure drives both cognate and bystander human B cells to adopt an atypical phenotype." Eur J Immunol **50**(8): 1187-1194.

Baiyegunhi, O., B. Ndlovu, F. Ogunshola, N. Ismail, B. D. Walker, T. Ndung'u and Z. M. Ndhlovu (2018). "Frequencies of Circulating Th1-Biased T Follicular Helper Cells in Acute HIV-1 Infection Correlate with the Development of HIV-Specific Antibody Responses and Lower Set Point Viral Load." J Virol **92**(15).

Barnden, M. J., J. Allison, W. R. Heath and F. R. Carbone (1998). "Defective TCR expression in transgenic mice constructed using cDNA-based α-and β-chain genes under the control of heterologous regulatory elements." Immunology and cell biology **76**(1): 34-40.

Batista, F. D. and N. E. Harwood (2009). "The who, how and where of antigen presentation to B cells." Nat Rev Immunol **9**(1): 15-27.

Baumjohann, D., S. Preite, A. Reboldi, F. Ronchi, K. M. Ansel, A. Lanzavecchia and F. Sallusto (2013). "Persistent Antigen and Germinal Center B Cells Sustain T Follicular Helper Cell Responses and Phenotype." Immunity **38**(3): 596-605.

Bentebibel, S. E., S. Lopez, G. Obermoser, N. Schmitt, C. Mueller, C. Harrod, E. Flano, A. Mejias, R. A. Albrecht, D. Blankenship, H. Xu, V. Pascual, J. Banchereau, A. Garcia-Sastre, A. K. Palucka, O. Ramilo and H. Ueno (2013). "Induction of ICOS+CXCR3+CXCR5+ TH cells correlates with antibody responses to influenza vaccination." Sci Transl Med **5**(176): 176ra132.

Berland, R. and H. H. Wortis (2002). "Origins and functions of B-1 cells with notes on the role of CD5." Annu Rev Immunol **20**: 253-300.

Bettini, E. and M. Locci (2021). "SARS-CoV-2 mRNA Vaccines: Immunological Mechanism and Beyond." Vaccines (Basel) **9**(2).

Bouneaud, C., Z. Garcia, P. Kourilsky and C. Pannetier (2005). "Lineage relationships, homeostasis, and recall capacities of central- and effector-memory CD8 T cells in vivo." J Exp Med **201**(4): 579-590.

Breitfeld, D., L. Ohl, E. Kremmer, J. Ellwart, F. Sallusto, M. Lipp and R. Förster (2000). "Follicular B helper T cells express CXC chemokine receptor 5, localize to B cell follicles, and support immunoglobulin production." The Journal of experimental medicine **192**(11): 1545-1552.

Bretscher, P. A. (1999). "A two-step, two-signal model for the primary activation of precursor helper T cells." Proceedings of the National Academy of Sciences **96**(1): 185-190.

Briquet, S., C. Marinach, O. Silvie and C. Vaquero (2020). "Preparing for Transmission: Gene Regulation in Plasmodium Sporozoites." Front Cell Infect Microbiol **10**: 618430.

Bruce, M. G., D. Bruden, D. Hurlburt, C. Zanis, G. Thompson, L. Rea, M. Toomey, L. Townshend-Bulson, K. Rudolph, L. Bulkow, P. R. Spradling, R. Baum, T. Hennessy and B. J. McMahon (2016). "Antibody Levels and Protection After Hepatitis B Vaccine: Results of a 30-Year Follow-up Study and Response to a Booster Dose." J Infect Dis **214**(1): 16-22.

Butler, N. S., N. W. Schmidt, A. M. Vaughan, A. S. Aly, S. H. Kappe and J. T. Harty (2011). "Superior antimalarial immunity after vaccination with late liver stage-arresting genetically attenuated parasites." Cell Host Microbe **9**(6): 451-462.

Cannons, J. L., K. T. Lu and P. L. Schwartzberg (2013). "T follicular helper cell diversity and plasticity." Trends Immunol **34**(5): 200-207.

Carrasco, Y. R. and F. D. Batista (2007). "B cells acquire particulate antigen in a macrophage-rich area at the boundary between the follicle and the subcapsular sinus of the lymph node." Immunity **27**(1): 160-171.

Chan, J.-A., J. R. Loughland, F. de Labastida Rivera, A. SheelaNair, D. W. Andrew, N. L. Dooley, B. D. Wines, F. H. Amante, L. Webb, P. M. Hogarth, J. S. McCarthy, J. G. Beeson, C. R. Engwerda and M. J. Boyle (2020). "Th2-like T Follicular Helper Cells Promote Functional Antibody Production during Plasmodium falciparum Infection." Cell Reports Medicine **1**(9).

Charles, E. D., C. Brunetti, S. Marukian, K. D. Ritola, A. H. Talal, K. Marks, I. M. Jacobson, C. M. Rice and L. B. Dustin (2011). "Clonal B cells in patients with hepatitis C virus-associated

mixed cryoglobulinemia contain an expanded anergic CD21low B-cell subset." Blood **117**(20): 5425-5437.

Chaudhury, S., R. S. MacGill, A. M. Early, J. S. Bolton, C. R. King, E. Locke, T. Pierson, D. F. Wirth, D. E. Neafsey and E. S. Bergmann-Leitner (2021). "Breadth of humoral immune responses to the C-terminus of the circumsporozoite protein is associated with protective efficacy induced by the RTS,S malaria vaccine." Vaccine **39**(6): 968-975.

Chen, G., B. Ning and T. Shi (2019). "Single-Cell RNA-Seq Technologies and Related Computational Data Analysis." Front Genet **10**: 317.

Chen, Z., N. Wang, Y. Yao and D. Yu (2022). "Context-dependent regulation of follicular helper T cell survival." Trends Immunol **43**(4): 309-321.

Chevalier, N., D. Jarrossay, E. Ho, D. T. Avery, C. S. Ma, D. Yu, F. Sallusto, S. G. Tangye and C. R. Mackay (2011). "CXCR5 expressing human central memory CD4 T cells and their relevance for humoral immune responses." The Journal of Immunology **186**(10): 5556-5568.

Cho, S. H., A. L. Raybuck, J. Blagih, E. Kemboi, V. H. Haase, R. G. Jones and M. R. Boothby (2019). "Hypoxia-inducible factors in CD4(+) T cells promote metabolism, switch cytokine secretion, and T cell help in humoral immunity." Proc Natl Acad Sci U S A **116**(18): 8975-8984.

Cho, Y. B., I. G. Lee, Y. H. Joo, S. H. Hong and Y. J. Seo (2020). "TCR Transgenic Mice: A Valuable Tool for Studying Viral Immunopathogenesis Mechanisms." Int J Mol Sci **21**(24).

Clyde, D. F., H. Most, V. McCarthy and J. Vanderberg (1973). "Immunization of man against sporozoite-induced falciparum malaria." American journal of medical sciences **266**(3): 169-177.

Cockburn, I. A. and R. A. Seder (2018). "Malaria prevention: from immunological concepts to effective vaccines and protective antibodies." Nat Immunol **19**(11): 1199-1211.

Coffman, R. L., A. Sher and R. A. Seder (2010). "Vaccine adjuvants: putting innate immunity to work." Immunity **33**(4): 492-503.

Coppi, A., C. Pinzon-Ortiz, C. Hutter and P. Sinnis (2005). "The Plasmodium circumsporozoite protein is proteolytically processed during cell invasion." J Exp Med **201**(1): 27-33.

Crawford, A., J. M. Angelosanto, C. Kao, T. A. Doering, P. M. Odorizzi, B. E. Barnett and E. J. Wherry (2014). "Molecular and transcriptional basis of CD4(+) T cell dysfunction during chronic infection." Immunity **40**(2): 289-302.

Crotty, S. (2011). "Follicular helper CD4 T cells (TFH)." Annu Rev Immunol **29**: 621-663.

Crotty, S. (2014). "T follicular helper cell differentiation, function, and roles in disease." Immunity **41**(4): 529-542.

Crotty, S. (2019). "T Follicular Helper Cell Biology: A Decade of Discovery and Diseases." Immunity **50**(5): 1132-1148.

Curtidor, H., C. Reyes, A. Bermúdez, M. Vanegas, Y. Varela and M. E. Patarroyo (2017). "Conserved Binding Regions Provide the Clue for Peptide-Based Vaccine Development: A Chemical Perspective." Molecules **22**(12): 2199.

Cyster, J. G. and C. D. C. Allen (2019). "B Cell Responses: Cell Interaction Dynamics and Decisions." Cell **177**(3): 524-540.

Dan, J. M., C. S. Lindestam Arlehamn, D. Weiskopf, R. da Silva Antunes, C. Havenar-Daughton, S. M. Reiss, M. Brigger, M. Bothwell, A. Sette and S. Crotty (2016). "A Cytokine-Independent Approach To Identify Antigen-Specific Human Germinal Center T Follicular Helper Cells and Rare Antigen-Specific CD4+ T Cells in Blood." J Immunol **197**(3): 983-993.

Dan, J. M., J. Mateus, Y. Kato, K. M. Hastie, E. D. Yu, C. E. Faliti, A. Grifoni, S. I. Ramirez, S. Haupt, A. Frazier, C. Nakao, V. Rayaprolu, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette and S. Crotty (2021). "Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection." Science **371**(6529).

Dan, J. M., J. Mateus, Y. Kato, K. M. Hastie, E. D. Yu, C. E. Faliti, A. Grifoni, S. I. Ramirez, S. Haupt, A. Frazier, C. Nakao, V. Rayaprolu, S. A. Rawlings, B. Peters, F. Krammer, V. Simon, E. O. Saphire, D. M. Smith, D. Weiskopf, A. Sette and S. Crotty (2021). "Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection." Science: eabf4063.

Datoo, M. S., M. H. Natama, A. Somé, O. Traoré, T. Rouamba, D. Bellamy, P. Yameogo, D. Valia, M. Tegneri, F. Ouedraogo, R. Soma, S. Sawadogo, F. Sorgho, K. Derra, E. Rouamba, B. Orindi, F. Ramos Lopez, A. Flaxman, F. Cappuccini, R. Kailath, S. Elias, E. Mukhopadhyay, A. Noe, M. Cairns, A. Lawrie, R. Roberts, I. Valéa, H. Sorgho, N. Williams, G. Glenn, L. Fries, J. Reimer, K. J. Ewer, U. Shaligram, A. V. S. Hill and H. Tinto (2021). "Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised controlled trial." The Lancet **397**(10287): 1809-1818.

De Silva, N. S. and U. Klein (2015). "Dynamics of B cells in germinal centres." Nat Rev Immunol **15**(3): 137-148.

Deng, J., Y. Wei, V. R. Fonseca, L. Graca and D. Yu (2019). "T follicular helper cells and T follicular regulatory cells in rheumatic diseases." Nat Rev Rheumatol **15**(8): 475-490.

Didierlaurent, A. M., B. Laupèze, A. Di Pasquale, N. Hergli, C. Collignon and N. Garçon (2017). "Adjuvant system AS01: helping to overcome the challenges of modern vaccines." Expert review of vaccines **16**(1): 55-63.

Dominguez, C. X., R. A. Amezquita, T. Guan, H. D. Marshall, N. S. Joshi, S. H. Kleinstein and S. M. Kaech (2015). "The transcription factors ZEB2 and T-bet cooperate to program cytotoxic T cell terminal differentiation in response to LCMV viral infection." J Exp Med **212**(12): 2041-2056.

Dondorp, A. M. and L. Von Seidlein (2017). 117 - Malaria. Infectious Diseases (Fourth Edition). J. Cohen, W. G. Powderly and S. M. Opal, Elsevier**:** 1014-1025.e1011.

Du, S. W., T. Arkatkar, F. Al Qureshah, H. M. Jacobs, C. D. Thouvenel, K. Chiang, A. D. Largent, Q. Z. Li, B. Hou, D. J. Rawlings and S. W. Jackson (2019). "Functional Characterization of CD11c(+) Age-Associated B Cells as Memory B Cells." J Immunol **203**(11): 2817-2826.

Du, S. W., T. Arkatkar, H. M. Jacobs, D. J. Rawlings and S. W. Jackson (2019). "Generation of functional murine CD11c(+) age-associated B cells in the absence of B cell T-bet expression." Eur J Immunol **49**(1): 170-178.

Duffy, P. E. and J. Patrick Gorres (2020). "Malaria vaccines since 2000: progress, priorities, products." NPJ Vaccines **5**(1): 48.

Ehrhardt, G. R., A. Hijikata, H. Kitamura, O. Ohara, J. Y. Wang and M. D. Cooper (2008). "Discriminating gene expression profiles of memory B cell subpopulations." J Exp Med **205**(8): 1807-1817.

Ehrhardt, G. R., J. T. Hsu, L. Gartland, C. M. Leu, S. Zhang, R. S. Davis and M. D. Cooper (2005). "Expression of the immunoregulatory molecule FcRH4 defines a distinctive tissue-based population of memory B cells." J Exp Med **202**(6): 783-791.

Enders, M. H., G. Bayarsaikhan, S. Ghilas, Y. C. Chua, R. May, M. N. de Menezes, Z. Ge, P. S. Tan, A. Cozijnsen and V. Mollard (2021). "Plasmodium berghei Hsp90 contains a natural immunogenic I-Ab-restricted antigen common to rodent and human Plasmodium species." Current research in immunology **2**: 79-92.

Enders, M. H., G. Bayarsaikhan, S. Ghilas, Y. C. Chua, R. May, M. N. de Menezes, Z. Ge, P. S. Tan, A. Cozijnsen, V. Mollard, K. Yui, G. I. McFadden, M. H. Lahoud, I. Caminschi, A. W. Purcell, R. B. Schittenhelm, L. Beattie, W. R. Heath and D. Fernandez-Ruiz (2021). "Plasmodium berghei Hsp90 contains a natural immunogenic I-Ab-restricted antigen common to rodent and human Plasmodium species." Current Research in Immunology **2**: 79-92.

Espinosa, D. A., D. Christensen, C. Munoz, S. Singh, E. Locke, P. Andersen and F. Zavala (2017). "Robust antibody and CD8(+) T-cell responses induced by P. falciparum CSP adsorbed to cationic liposomal adjuvant CAF09 confer sterilizing immunity against experimental rodent malaria infection." NPJ Vaccines **2**.

Eto, D., C. Lao, D. DiToro, B. Barnett, T. C. Escobar, R. Kageyama, I. Yusuf and S. Crotty (2011). "IL-21 and IL-6 are critical for different aspects of B cell immunity and redundantly induce optimal follicular helper CD4 T cell (Tfh) differentiation." PLoS One **6**(3): e17739.

Evans, E., S. Einfeld, D. Mowat, J. Taffe, B. Tonge and M. Wilson (2012). "The behavioral phenotype of Mowat-Wilson syndrome." Am J Med Genet A **158A**(2): 358-366.

Ferguson, D. J., A. E. Balaban, E. M. Patzewitz, R. J. Wall, C. S. Hopp, B. Poulin, A. Mohmmed, P. Malhotra, A. Coppi, P. Sinnis and R. Tewari (2014). "The repeat region of the circumsporozoite protein is critical for sporozoite formation and maturation in Plasmodium." PLoS One **9**(12): e113923.

Fernandez-Ruiz, D., W. Y. Ng, L. E. Holz, J. Z. Ma, A. Zaid, Y. C. Wong, L. S. Lau, V. Mollard, A. Cozijnsen, N. Collins, J. Li, G. M. Davey, Y. Kato, S. Devi, R. Skandari, M. Pauley, J. H. Manton, D. I. Godfrey, A. Braun, S. S. Tay, P. S. Tan, D. G. Bowen, F. Koch-Nolte, B. Rissiek, F. R. Carbone, B. S. Crabb, M. Lahoud, I. A. Cockburn, S. N. Mueller, P. Bertolino, G. I. McFadden, I. Caminschi and W. R. Heath (2016). "Liver-Resident Memory CD8(+) T Cells Form a Front-Line Defense against Malaria Liver-Stage Infection." Immunity **45**(4): 889-902.

Flannigan, K. L., V. L. Ngo, D. Geem, A. Harusato, S. A. Hirota, C. A. Parkos, N. W. Lukacs, A. Nusrat, V. Gaboriau-Routhiau, N. Cerf-Bensussan, A. T. Gewirtz and T. L. Denning (2017). "IL-17A-mediated neutrophil recruitment limits expansion of segmented filamentous bacteria." Mucosal Immunol **10**(3): 673-684.

Fleire, S. J., J. P. Goldman, Y. R. Carrasco, M. Weber, D. Bray and F. D. Batista (2006). "B cell ligand discrimination through a spreading and contraction response." Science **312**(5774): 738-741.

Frischknecht, F. and K. Matuschewski (2017). "Plasmodium Sporozoite Biology." Cold Spring Harb Perspect Med **7**(5).

Frith, K., C. M. L. Munier, L. Hastings, D. Mowat, M. Wilson, N. Seddiki, R. Macintosh, A. D. Kelleher, P. Gray and J. J. Zaunders (2021). "The Role of ZEB2 in Human CD8 T Lymphocytes: Clinical and Cellular Immune Profiling in Mowat-Wilson Syndrome." Int J Mol Sci **22**(10).

Gao, X. and I. A. Cockburn (2022). "The development and function of CD11c(+) atypical B cells - insights from single cell analysis." Front Immunol **13**: 979060.

Gao, X. and I. A. Cockburn (2022). "The development and function of CD11c+ atypical B cells-insights from single cell analysis." Frontiers in Immunology **13**.

Gao, X., H. Wang, Z. Chen, P. Zhou and D. Yu (2020). "An optimized method to differentiate mouse follicular helper T cells in vitro." Cell Mol Immunol **17**(7): 779-781.

Gaya, M., P. Barral, M. Burbage, S. Aggarwal, B. Montaner, A. Warren Navia, M. Aid, C. Tsui, P. Maldonado, U. Nair, K. Ghneim, P. G. Fallon, R. P. Sekaly, D. H. Barouch, A. K. Shalek, A. Bruckbauer, J. Strid and F. D. Batista (2018). "Initiation of Antiviral B Cell Immunity Relies on Innate Signals from Spatially Positioned NKT Cells." Cell **172**(3): 517-533 e520.

Germain, R. N. (2002). "T-cell development and the CD4-CD8 lineage decision." Nat Rev Immunol **2**(5): 309-322.

Giles, J. R., M. Kashgarian, P. A. Koni and M. J. Shlomchik (2015). "B Cell-Specific MHC Class II Deletion Reveals Multiple Nonredundant Roles for B Cell Antigen Presentation in Murine Lupus." J Immunol **195**(6): 2571-2579.

Glass, D. R., A. G. Tsai, J. P. Oliveria, F. J. Hartmann, S. C. Kimmey, A. A. Calderon, L. Borges, M. C. Glass, L. E. Wagar, M. M. Davis and S. C. Bendall (2020). "An Integrated Multi-omic Single-Cell Atlas of Human B Cell Identity." Immunity **53**(1): 217-232 e215.

Goel, R. R., M. M. Painter, S. A. Apostolidis, D. Mathew, W. Meng, A. M. Rosenfeld, K. A. Lundgreen, A. Reynaldi, D. S. Khoury and A. Pattekar (2021). "mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern." Science **374**(6572): abm0829.

Goodnow, C. C., J. Crosbie, S. Adelstein, T. B. Lavoie, S. J. Smith-Gill, R. A. Brink, H. Pritchard-Briscoe, J. S. Wotherspoon, R. H. Loblay and K. Raphael (1988). "Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice." Nature **334**(6184): 676-682.

Gowthaman, U., J. S. Chen, B. Zhang, W. F. Flynn, Y. Lu, W. Song, J. Joseph, J. A. Gertie, L. Xu, M. A. Collet, J. D. S. Grassmann, T. Simoneau, D. Chiang, M. C. Berin, J. E. Craft, J. S. Weinstein, A. Williams and S. C. Eisenbarth (2019). "Identification of a T follicular helper cell subset that drives anaphylactic IgE." Science **365**(6456).

Gu, S., X. Han, C. Yao, H. Ding, J. Zhang, G. Hou, B. Qu, H. Zhou, Z. Ying, Z. Ye, J. Qian, Q. Guo, S. Chen, C. G. Vinuesa, D. Dai and N. Shen (2021). "The transcription factor Zeb2 drives differentiation of age-associated B cells." bioRxiv: 2021.2007.2024.453633.

Hale, J. S., B. Youngblood, D. R. Latner, A. U. Mohammed, L. Ye, R. S. Akondy, T. Wu, S. S. Iyer and R. Ahmed (2013). "Distinct memory CD4+ T cells with commitment to T follicular helper- and T helper 1-cell lineages are generated after acute viral infection." Immunity **38**(4): 805-817.

Hao, Y., S. Hao, E. Andersen-Nissen, W. M. Mauck, 3rd, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R.

Gottardo, P. Smibert and R. Satija (2021). "Integrated analysis of multimodal single-cell data." Cell **184**(13): 3573-3587 e3529.

Guan, T., C. X. Dominguez, R. A. Amezquita, B. J. Laidlaw, J. Cheng, J. Henao-Mejia, A. Williams, R. A. Flavell, J. Lu and S. M. Kaech (2018). "ZEB1, ZEB2, and the miR-200 family form a counterregulatory network to regulate CD8(+) T cell fates." J Exp Med **215**(4): 1153-1168.

Hao, Y., P. O'Neill, M. S. Naradikian, J. L. Scholz and M. P. Cancro (2011). "A B-cell subset uniquely responsive to innate stimuli accumulates in aged mice." Blood **118**(5): 1294-1304.

Hardy, R. R., C. E. Carmack, S. A. Shinton, J. D. Kemp and K. Hayakawa (1991). "Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow." The Journal of experimental medicine **173**(5): 1213-1225.

Harms Pritchard, G. and M. Pepper (2018). "Memory B cell heterogeneity: Remembrance of things past." J Leukoc Biol **103**(2): 269-274.

Harris, N. and W. C. Gause (2011). "To B or not to B: B cells and the Th2-type immune response to helminths." Trends Immunol **32**(2): 80-88.

He, J., L. M. Tsai, Y. A. Leong, X. Hu, C. S. Ma, N. Chevalier, X. Sun, K. Vandenberg, S. Rockman, Y. Ding, L. Zhu, W. Wei, C. Wang, A. Karnowski, G. T. Belz, J. R. Ghali, M. C. Cook, D. S. Riminton, A. Veillette, P. L. Schwartzberg, F. Mackay, R. Brink, S. G. Tangye, C. G. Vinuesa, C. R. Mackay, Z. Li and D. Yu (2013). "Circulating precursor CCR7(lo)PD-1(hi) CXCR5(+) CD4(+) T cells indicate Tfh cell activity and promote antibody responses upon antigen reexposure." Immunity **39**(4): 770-781.

Heesters, B. A., P. Chatterjee, Y. A. Kim, S. F. Gonzalez, M. P. Kuligowski, T. Kirchhausen and M. C. Carroll (2013). "Endocytosis and recycling of immune complexes by follicular dendritic cells enhances B cell antigen binding and activation." Immunity **38**(6): 1164-1175.

Heit, A., F. Schmitz, S. Gerdts, B. Flach, M. S. Moore, J. A. Perkins, H. S. Robins, A. Aderem, P. Spearman, G. D. Tomaras, S. C. De Rosa and M. J. McElrath (2017). "Vaccination establishes clonal relatives of germinal center T cells in the blood of humans." J Exp Med **214**(7): 2139-2152.

Herati, R. S., A. Muselman, L. Vella, B. Bengsch, K. Parkhouse, D. Del Alcazar, J. Kotzin, S. A. Doyle, P. Tebas, S. E. Hensley, L. F. Su, K. E. Schmader and E. J. Wherry (2017). "Successive annual influenza vaccination induces a recurrent oligoclonotypic memory response in circulating T follicular helper cells." Sci Immunol **2**(8).

Higashi, Y., M. Maruhashi, L. Nelles, T. Van de Putte, K. Verschueren, T. Miyoshi, A. Yoshimoto, H. Kondoh and D. Huylebroeck (2002). "Generation of the floxed allele of the SIP1

(Smad-interacting protein 1) gene for Cre-mediated conditional knockout in the mouse." Genesis **32**(2): 82-84.

Hinrichs, C. S., R. Spolski, C. M. Paulos, L. Gattinoni, K. W. Kerstann, D. C. Palmer, C. A. Klebanoff, S. A. Rosenberg, W. J. Leonard and N. P. Restifo (2008). "IL-2 and IL-21 confer opposing differentiation programs to CD8+ T cells for adoptive immunotherapy." Blood **111**(11): 5326-5333.

Hirota, K., J. E. Turner, M. Villa, J. H. Duarte, J. Demengeot, O. M. Steinmetz and B. Stockinger (2013). "Plasticity of Th17 cells in Peyer's patches is responsible for the induction of T cell-dependent IgA responses." Nat Immunol **14**(4): 372-379.

Hoffman, S. L., L. M. Goh, T. C. Luke, I. Schneider, T. P. Le, D. L. Doolan, J. Sacci, P. De la Vega, M. Dowler and C. Paul (2002). "Protection of humans against malaria by immunization with radiation-attenuated Plasmodium falciparum sporozoites." The Journal of infectious diseases **185**(8): 1155-1164.

Holla, P., B. Dizon, A. A. Ambegaonkar, N. Rogel, E. Goldschmidt, A. K. Boddapati, H. Sohn, D. Sturdevant, J. W. Austin and L. Kardava (2021). "Shared transcriptional profiles of atypical B cells suggest common drivers of expansion and function in malaria, HIV, and autoimmunity." Science Advances **7**(22): eabg8384.

Hong, S., Z. Zhang, H. Liu, M. Tian, X. Zhu, Z. Zhang, W. Wang, X. Zhou, F. Zhang, Q. Ge, B. Zhu, H. Tang, Z. Hua and B. Hou (2018). "B Cells Are the Dominant Antigen-Presenting Cells that Activate Naive CD4(+) T Cells upon Immunization with a Virus-Derived Nanoparticle Antigen." Immunity **49**(4): 695-708 e694.

Hoogeboom, R. and P. Tolar (2016). "Molecular Mechanisms of B Cell Antigen Gathering and Endocytosis." Curr Top Microbiol Immunol **393**: 45-63.

Horns, F., C. L. Dekker and S. R. Quake (2020). "Memory B Cell Activation, Broad Anti-influenza Antibodies, and Bystander Activation Revealed by Single-Cell Transcriptomics." Cell Rep **30**(3): 905-913 e906.

Huang, J., S. Ruan, X. Wu and X. Zhou (2018). "Seasonal transmission dynamics of measles in China." Theory Biosci **137**(2): 185-195.

Isnardi, I., Y. S. Ng, L. Menard, G. Meyers, D. Saadoun, I. Srdanovic, J. Samuels, J. Berman, J. H. Buckner, C. Cunningham-Rundles and E. Meffre (2010). "Complement receptor 2/CD21- human naive B cells contain mostly autoreactive unresponsive clones." Blood **115**(24): 5026-5036.

Ivashkiv, L. B. (2018). "IFNgamma: signalling, epigenetics and roles in immunity, metabolism, disease and cancer immunotherapy." Nat Rev Immunol **18**(9): 545-558.

Jackson, S. W., N. E. Scharping, N. S. Kolhatkar, S. Khim, M. A. Schwartz, Q. Z. Li, K. L. Hudkins, C. E. Alpers, D. Liggitt and D. J. Rawlings (2014). "Opposing impact of B cell-intrinsic TLR7 and TLR9 signals on autoantibody repertoire and systemic inflammation." J Immunol **192**(10): 4525-4532.

Jacobsen, J. T., W. Hu, T. B. R. Castro, S. Solem, A. Galante, Z. Lin, S. J. Allon, L. Mesin, A. M. Bilate and A. Schiepers (2021). "Expression of Foxp3 by T follicular helper cells in end-stage germinal centers." Science **373**(6552): eabe5146.

Jalili, V., E. Afgan, Q. Gu, D. Clements, D. Blankenberg, J. Goecks, J. Taylor and A. Nekrutenko (2020). "The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2020 update." Nucleic acids research **48**(W1): W395-W402.

Janse, C. J., J. Ramesar and A. P. Waters (2006). "High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei." Nat Protoc **1**(1): 346-356.

Jash, A., Y. Wang, F. J. Weisel, C. D. Scharer, J. M. Boss, M. J. Shlomchik and D. Bhattacharya (2016). "ZBTB32 Restricts the Duration of Memory B Cell Recall Responses." J Immunol **197**(4): 1159-1168.

Jenks, S. A., K. S. Cashman, E. Zumaquero, U. M. Marigorta, A. V. Patel, X. Wang, D. Tomar, M. C. Woodruff, Z. Simon, R. Bugrovsky, E. L. Blalock, C. D. Scharer, C. M. Tipton, C. Wei, S. S. Lim, M. Petri, T. B. Niewold, J. H. Anolik, G. Gibson, F. E. Lee, J. M. Boss, F. E. Lund and I. Sanz (2018). "Distinct Effector B Cells Induced by Unregulated Toll-like Receptor 7 Contribute to Pathogenic Responses in Systemic Lupus Erythematosus." Immunity **49**(4): 725-739 e726.

Jia, Q., H. Chu, Z. Jin, H. Long and B. Zhu (2022). "High-throughput single-small es, Cyrillicell sequencing in cancer research." Signal Transduct Target Ther **7**(1): 145.

Johnson, J. L., R. L. Rosenthal, J. J. Knox, A. Myles, M. S. Naradikian, J. Madej, M. Kostiv, A. M. Rosenfeld, W. Meng, S. R. Christensen, S. E. Hensley, J. Yewdell, D. H. Canaday, J. Zhu, A. B. McDermott, Y. Dori, M. Itkin, E. J. Wherry, N. Pardi, D. Weissman, A. Naji, E. T. L. Prak, M. R. Betts and M. P. Cancro (2020). "The Transcription Factor T-bet Resolves Memory B Cell Subsets with Distinct Tissue Distributions and Antibody Specificities in Mice and Humans." Immunity **52**(5): 842-855 e846.

Johnston, R. J., Y. S. Choi, J. A. Diamond, J. A. Yang and S. Crotty (2012). "STAT5 is a potent negative regulator of TFH cell differentiation." J Exp Med **209**(2): 243-250.

Johnston, R. J., A. C. Poholek, D. DiToro, I. Yusuf, D. Eto, B. Barnett, A. L. Dent, J. Craft and S. Crotty (2009). "Bcl6 and Blimp-1 are reciprocal and antagonistic regulators of T follicular helper cell differentiation." Science **325**(5943): 1006-1010.

Junt, T., E. A. Moseman, M. Iannacone, S. Massberg, P. A. Lang, M. Boes, K. Fink, S. E. Henrickson, D. M. Shayakhmetov, N. C. Di Paolo, N. van Rooijen, T. R. Mempel, S. P. Whelan and U. H. von Andrian (2007). "Subcapsular sinus macrophages in lymph nodes clear lymph-borne viruses and present them to antiviral B cells." Nature **450**(7166): 110-114.

Kappe, S. H., C. A. Buscaglia and V. Nussenzweig (2004). "Plasmodium sporozoite molecular cell biology." Annu Rev Cell Dev Biol **20**: 29-59.

Keller, B., V. Strohmeier, I. Harder, S. Unger, K. J. Payne, G. Andrieux, M. Boerries, P. T. Felixberger, J. J. Landry and A. Nieters (2021). "The expansion of human T-bethighCD21low B cells is T cell dependent." Science Immunology **6**(64): eabh0891.

Kent, S. J., D. S. Khoury, A. Reynaldi, J. A. Juno, A. K. Wheatley, E. Stadler, E. John Wherry, J. Triccas, S. C. Sasson, D. Cromer and M. P. Davenport (2022). "Disentangling the relative importance of T cell responses in COVID-19: leading actors or supporting cast?" Nat Rev Immunol **22**(6): 387-397.

Kim, C. C., A. M. Baccarella, A. Bayat, M. Pepper and M. F. Fontana (2019). "FCRL5(+) Memory B Cells Exhibit Robust Recall Responses." Cell Rep **27**(5): 1446-1460 e1444.

King, H. W., N. Orban, J. C. Riches, A. J. Clear, G. Warnes, S. A. Teichmann and L. K. James (2021). "Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics." Science Immunology **6**(56): eabe6291.

Kirsebom, F. C. M., N. Andrews, J. Stowe, S. Toffa, R. Sachdeva, E. Gallagher, N. Groves, A.-M. O'Connell, M. Chand, M. Ramsay and J. L. Bernal (2022). "COVID-19 vaccine effectiveness against the omicron (BA.2) variant in England." The Lancet Infectious Diseases **22**(7): 931-933.

Kisalu, N. K., A. H. Idris, C. Weidle, Y. Flores-Garcia, B. J. Flynn, B. K. Sack, S. Murphy, A. Schon, E. Freire, J. R. Francica, A. B. Miller, J. Gregory, S. March, H. X. Liao, B. F. Haynes, K. Wiehe, A. M. Trama, K. O. Saunders, M. A. Gladden, A. Monroe, M. Bonsignori, M. Kanekiyo, A. K. Wheatley, A. B. McDermott, S. K. Farney, G. Y. Chuang, B. Zhang, N. Kc, S. Chakravarty, P. D. Kwong, P. Sinnis, S. N. Bhatia, S. H. I. Kappe, B. K. L. Sim, S. L. Hoffman, F. Zavala, M. Pancera and R. A. Seder (2018). "A human monoclonal antibody prevents malaria infection by targeting a new site of vulnerability on the parasite." Nat Med **24**(4): 408-416.

Klein, U. and R. Dalla-Favera (2008). "Germinal centres: role in B-cell physiology and malignancy." Nat Rev Immunol **8**(1): 22-33.

Knochelmann, H. M., C. J. Dwyer, S. R. Bailey, S. M. Amaya, D. M. Elston, J. M. Mazza-McCrann and C. M. Paulos (2018). "When worlds collide: Th17 and Treg cells in cancer and autoimmunity." Cell Mol Immunol **15**(5): 458-469.

Knox, J. J., M. Buggert, L. Kardava, K. E. Seaton, M. A. Eller, D. H. Canaday, M. L. Robb, M. A. Ostrowski, S. G. Deeks, M. K. Slifka, G. D. Tomaras, S. Moir, M. A. Moody and M. R. Betts (2017). "T-bet+ B cells are induced by human viral infections and dominate the HIV gp140 response." JCI Insight **2**(8).

Koyasu, S. and K. Moro (2011). "Type 2 innate immune responses and the natural helper cell." Immunology **132**(4): 475-481.

Krautler, N. J., D. Suan, D. Butt, K. Bourne, J. R. Hermes, T. D. Chan, C. Sundling, W. Kaplan, P. Schofield, J. Jackson, A. Basten, D. Christ and R. Brink (2017). "Differentiation of germinal center B cells into plasma cells is initiated by high-affinity antigen and completed by Tfh cells." J Exp Med **214**(5): 1259-1267.

Kryczek, I., E. Zhao, Y. Liu, Y. Wang, L. Vatan, W. Szeliga, J. Moyer, A. Klimczak, A. Lange and W. Zou (2011). "Human TH17 cells are long-lived effector memory cells." Sci Transl Med **3**(104): 104ra100.

Kurucz, R., P. H. Seeberger and D. Varón Silva (2013). Glycosylphosphatidylinositols in Malaria: GPI Biosynthesis and GPI-Derived Proteins. Encyclopedia of Malaria**:** 1-13.

Kwak, K., N. Quizon, H. Sohn, A. Saniee, J. Manzella-Lapeira, P. Holla, J. Brzostowski, J. Lu, H. Xie and C. Xu (2018). "Intrinsic properties of human germinal center B cells set antigen affinity thresholds." Science immunology **3**(29): eaau6598.

Lau, D., L. Y.-L. Lan, S. F. Andrews, C. Henry, K. T. Rojas, K. E. Neu, M. Huang, Y. Huang, B. DeKosky and A.-K. E. Palm (2017). "Low CD21 expression defines a population of recent germinal center graduates primed for plasma cell differentiation." Science immunology **2**(7): eaai8153.

Law, C. W., M. Alhamdoosh, S. Su, X. Dong, L. Tian, G. K. Smyth and M. E. Ritchie (2016). "RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR." F1000Res **5**.

LeBien, T. W. and T. F. Tedder (2008). "B lymphocytes: how they develop and function." Blood **112**(5): 1570-1580.

Lederer, K., D. Castano, D. Gomez Atria, T. H. Oguin, 3rd, S. Wang, T. B. Manzoni, H. Muramatsu, M. J. Hogan, F. Amanat, P. Cherubin, K. A. Lundgreen, Y. K. Tam, S. H. Y. Fan, L. C. Eisenlohr, I. Maillard, D. Weissman, P. Bates, F. Krammer, G. D. Sempowski, N. Pardi

and M. Locci (2020). "SARS-CoV-2 mRNA Vaccines Foster Potent Antigen-Specific Germinal Center Responses Associated with Neutralizing Antibody Generation." Immunity **53**(6): 1281-1295 e1285.

Lee, J. H., J. K. Hu, E. Georgeson, C. Nakao, B. Groschel, T. Dileepan, M. K. Jenkins, G. Seumois, P. Vijayanand, W. R. Schief and S. Crotty (2021). "Modulating the quantity of HIV Env-specific CD4 T cell help promotes rare B cell responses in germinal centers." J Exp Med **218**(2).

Lee, R. D., S. A. Munro, T. P. Knutson, R. S. LaRue, L. M. Heltemes-Harris and M. A. Farrar (2021). "Single-cell analysis identifies dynamic gene expression networks that govern B cell development and transformation." Nat Commun **12**(1): 6843.

Lefebvre, M. N. and J. T. Harty (2020). "You Shall Not Pass: Memory CD8 T Cells in Liver-Stage Malaria." Trends Parasitol **36**(2): 147-157.

Leonard, W. J., R. Zeng and R. Spolski (2008). "Interleukin 21: a cytokine/cytokine receptor system that has come of age." J Leukoc Biol **84**(2): 348-356.

Levack, R. C., K. L. Newell, M. Popescu, B. Cabrera-Martinez and G. M. Winslow (2020). "CD11c(+) T-bet(+) B Cells Require IL-21 and IFN-gamma from Type 1 T Follicular Helper Cells and Intrinsic Bcl-6 Expression but Develop Normally in the Absence of T-bet." J Immunol **205**(4): 1050-1058.

Li, H., F. Borrego, S. Nagata and M. Tolnay (2016). "Fc Receptor-like 5 Expression Distinguishes Two Distinct Subsets of Human Circulating Tissue-like Memory B Cells." J Immunol **196**(10): 4064-4074.

Li, J., T. Riedt, S. Goossens, C. Carrillo Garcia, S. Szczepanski, M. Brandes, T. Pieters, L. Dobrosch, I. Gutgemann, N. Farla, E. Radaelli, P. Hulpiau, N. Mallela, H. Frohlich, R. La Starza, C. Matteucci, T. Chen, P. Brossart, C. Mecucci, D. Huylebroeck, J. J. Haigh and V. Janzen (2017). "The EMT transcription factor Zeb2 controls adult murine hematopoietic differentiation by regulating cytokine signaling." Blood **129**(4): 460-472.

Li, Y., Z. Li and F. Hu (2021). "Double-negative (DN) B cells: an under-recognized effector memory B cell subset in autoimmunity." Clin Exp Immunol **205**(2): 119-127.

Lin, J. W., T. N. Shaw, T. Annoura, A. Fougere, P. Bouchier, S. Chevalley-Maurel, H. Kroeze, B. Franke-Fayard, C. J. Janse, K. N. Couper and S. M. Khan (2014). "The subcellular location of ovalbumin in Plasmodium berghei blood stages influences the magnitude of T-cell responses." Infect Immun **82**(11): 4654-4665.

Lindenstrom, T., J. Woodworth, J. Dietrich, C. Aagaard, P. Andersen and E. M. Agger (2012). "Vaccine-induced th17 cells are maintained long-term postvaccination as a distinct and phenotypically stable memory subset." Infect Immun **80**(10): 3533-3544.

Lindner, S. E., K. E. Swearingen, M. J. Shears, M. P. Walker, E. N. Vrana, K. J. Hart, A. M. Minns, P. Sinnis, R. L. Moritz and S. H. I. Kappe (2019). "Transcriptomics and proteomics reveal two waves of translational repression during the maturation of malaria parasite sporozoites." Nat Commun **10**(1): 4964.

Lino, C. A., J. C. Harper, J. P. Carney and J. A. Timlin (2018). "Delivering CRISPR: a review of the challenges and approaches." Drug Deliv **25**(1): 1234-1257.

Liu, X., X. Chen, B. Zhong, A. Wang, X. Wang, F. Chu, R. I. Nurieva, X. Yan, P. Chen, L. G. van der Flier, H. Nakatsukasa, S. S. Neelapu, W. Chen, H. Clevers, Q. Tian, H. Qi, L. Wei and C. Dong (2014). "Transcription factor achaete-scute homologue 2 initiates follicular T-helper-cell development." Nature **507**(7493): 513-518.

Locci, M., C. Havenar-Daughton, E. Landais, J. Wu, M. A. Kroenke, C. L. Arlehamn, L. F. Su, R. Cubas, M. M. Davis, A. Sette, E. K. Haddad, A. V. I. P. C. P. I. International, P. Poignard and S. Crotty (2013). "Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses." Immunity **39**(4): 758-769.

Longley, R. J., A. M. Salman, M. G. Cottingham, K. Ewer, C. J. Janse, S. M. Khan, A. J. Spencer and A. V. Hill (2015). "Comparative assessment of vaccine vectors encoding ten malaria antigens identifies two protective liver-stage candidates." Scientific reports **5**(1): 1-13.

Lu, K. T., Y. Kanno, J. L. Cannons, R. Handon, P. Bible, A. G. Elkahloun, S. M. Anderson, L. Wei, H. Sun, J. J. O'Shea and P. L. Schwartzberg (2011). "Functional and epigenetic studies reveal multistep differentiation and plasticity of in vitro-generated and in vivo-derived follicular T helper cells." Immunity **35**(4): 622-632.

Luckheeram, R. V., R. Zhou, A. D. Verma and B. Xia (2012). "CD4(+)T cells: differentiation and functions." Clin Dev Immunol **2012**: 925135.

MacLeod, M. K., A. David, A. S. McKee, F. Crawford, J. W. Kappler and P. Marrack (2011). "Memory CD4 T cells that express CXCR5 provide accelerated help to B cells." J Immunol **186**(5): 2889-2896.

MacLeod, M. K., J. W. Kappler and P. Marrack (2010). "Memory CD4 T cells: generation, reactivation and re-assignment." Immunology **130**(1): 10-15.

Malinova, D., L. Wasim, R. Newman, A. Martinez-Riano, N. Engels and P. Tolar (2021). "Endophilin A2 regulates B-cell endocytosis and is required for germinal center and humoral responses." EMBO Rep **22**(9): e51328.

Mandl, J. N., J. P. Monteiro, N. Vrisekoop and R. N. Germain (2013). "T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens." Immunity **38**(2): 263-274.

Manni, M., S. Gupta, E. Ricker, Y. Chinenov, S. H. Park, M. Shi, T. Pannellini, R. Jessberger, L. B. Ivashkiv and A. B. Pernis (2018). "Regulation of age-associated B cells by IRF5 in systemic autoimmunity." Nat Immunol **19**(4): 407-419.

Mariuzza, R. A., P. Agnihotri and J. Orban (2020). "The structural basis of T-cell receptor (TCR) activation: An enduring enigma." J Biol Chem **295**(4): 914-925.

Marshall, H. D., J. P. Ray, B. J. Laidlaw, N. Zhang, D. Gawande, M. M. Staron, J. Craft and S. M. Kaech (2015). "The transforming growth factor beta signaling pathway is critical for the formation of CD4 T follicular helper cells and isotype-switched antibody responses in the lung mucosa." Elife **4**: e04851.

Martinez-Riano, A., E. R. Bovolenta, P. Mendoza, C. L. Oeste, M. J. Martin-Bermejo, P. Bovolenta, M. Turner, N. Martinez-Martin and B. Alarcon (2018). "Antigen phagocytosis by B cells is required for a potent humoral response." EMBO Rep **19**(9).

Maul, R. W. and P. J. Gearhart (2010). "AID and somatic hypermutation." Adv Immunol **105**: 159-191.

McCarron, M. J. and J. C. Marie (2014). "TGF-beta prevents T follicular helper cell accumulation and B cell autoreactivity." J Clin Invest **124**(10): 4375-4386.

McGeachy, M. J. (2013). "Th17 memory cells: live long and proliferate." J Leukoc Biol **94**(5): 921-926.

McNamara, H. A., A. H. Idris, H. J. Sutton, R. Vistein, B. J. Flynn, Y. Cai, K. Wiehe, K. E. Lyke, D. Chatterjee, N. Kc, S. Chakravarty, B. K. Lee Sim, S. L. Hoffman, M. Bonsignori, R. A. Seder and I. A. Cockburn (2020). "Antibody Feedback Limits the Expansion of B Cell Responses to Malaria Vaccination but Drives Diversification of the Humoral Response." Cell Host Microbe **28**(4): 572-585 e577.

McNamara, H. A., M. H. Lahoud, Y. Cai, J. Durrant-Whyte, J. H. O'Connor, I. Caminschi and I. A. Cockburn (2022). "Splenic Dendritic Cells and Macrophages Drive B Cells to Adopt a Plasmablast Cell Fate." Front Immunol **13**: 825207.

McShane, A. N. and D. Malinova (2022). "The Ins and Outs of Antigen Uptake in B cells." Front Immunol **13**: 892169.

Meckiff, B. J., C. Ramirez-Suastegui, V. Fajardo, S. J. Chee, A. Kusnadi, H. Simon, S. Eschweiler, A. Grifoni, E. Pelosi, D. Weiskopf, A. Sette, F. Ay, G. Seumois, C. H. Ottensmeier and P. Vijayanand (2020). "Imbalance of Regulatory and Cytotoxic SARS-CoV-2-Reactive CD4(+) T Cells in COVID-19." Cell **183**(5): 1340-1353 e1316.

Modrzejewski, D., F. Hartung, H. Lehnert, T. Sprink, C. Kohl, J. Keilwagen and R. Wilhelm (2020). "Which factors affect the occurrence of off-target effects caused by the use of CRISPR/Cas: a systematic review in plants." Frontiers in plant science **11**: 574959.

Moir, S., J. Ho, A. Malaspina, W. Wang, A. C. DiPoto, M. A. O'Shea, G. Roby, S. Kottilil, J. Arthos, M. A. Proschan, T. W. Chun and A. S. Fauci (2008). "Evidence for HIV-associated B cell exhaustion in a dysfunctional memory B cell compartment in HIV-infected viremic individuals." J Exp Med **205**(8): 1797-1805.

Moir, S., A. Malaspina, J. Ho, W. Wang, A. C. Dipoto, M. A. O'Shea, G. Roby, J. M. Mican, S. Kottilil, T. W. Chun, M. A. Proschan and A. S. Fauci (2008). "Normalization of B cell counts and subpopulations after antiretroviral therapy in chronic HIV disease." J Infect Dis **197**(4): 572-579.

Morita, R., N. Schmitt, S. E. Bentebibel, R. Ranganathan, L. Bourdery, G. Zurawski, E. Foucat, M. Dullaers, S. Oh, N. Sabzghabaei, E. M. Lavecchio, M. Punaro, V. Pascual, J. Banchereau and H. Ueno (2011). "Human blood CXCR5(+)CD4(+) T cells are counterparts of T follicular cells and contain specific subsets that differentially support antibody secretion." Immunity **34**(1): 108-121.

Mourad, A. (2020). Immunogenetic Aspect of B-Cell Antigen Receptor Diversity Generation. Normal and Malignant B-Cell. A. Mourad. Rijeka, IntechOpen**:** Ch. 2.

Mourad, A. (2020). Introductory Chapter: B-Cells. Normal and Malignant B-Cell. A. Mourad. Rijeka, IntechOpen**:** Ch. 1.

Muellenbeck, M. F., B. Ueberheide, B. Amulic, A. Epp, D. Fenyo, C. E. Busse, M. Esen, M. Theisen, B. Mordmuller and H. Wardemann (2013). "Atypical and classical memory B cells produce Plasmodium falciparum neutralizing antibodies." J Exp Med **210**(2): 389-399.

Muranski, P., Z. A. Borman, S. P. Kerkar, C. A. Klebanoff, Y. Ji, L. Sanchez-Perez, M. Sukumar, R. N. Reger, Z. Yu, S. J. Kern, R. Roychoudhuri, G. A. Ferreyra, W. Shen, S. K. Durum, L. Feigenbaum, D. C. Palmer, P. A. Antony, C. C. Chan, A. Laurence, R. L. Danner, L. Gattinoni and N. P. Restifo (2011). "Th17 cells are long lived and retain a stem cell-like molecular signature." Immunity **35**(6): 972-985.

Mwakingwe-Omari, A., S. A. Healy, J. Lane, D. M. Cook, S. Kalhori, C. Wyatt, A. Kolluri, O. Marte-Salcedo, A. Imeru, M. Nason, L. K. Ding, H. Decederfelt, J. Duan, J. Neal, J. Raiten, G.

Lee, J. C. C. Hume, J. E. Jeon, I. Ikpeama, N. Kc, S. Chakravarty, T. Murshedkar, L. W. P. Church, A. Manoj, A. Gunasekera, C. Anderson, S. C. Murphy, S. March, S. N. Bhatia, E. R. James, P. F. Billingsley, B. K. L. Sim, T. L. Richie, I. Zaidi, S. L. Hoffman and P. E. Duffy (2021). "Two chemoattenuated PfSPZ malaria vaccines induce sterile hepatic immunity." Nature **595**(7866): 289-294.

Nanan, R., A. Rauch, E. Kämpgen, S. Niewiesk and H. W. Kreth (2000). "A novel sensitive approach for frequency analysis of measles virus-specific memory T-lymphocytes in healthy adults with a childhood history of natural measles." Journal of General Virology **81**(5): 1313-1319.

Naradikian, M. S., A. Myles, D. P. Beiting, K. J. Roberts, L. Dawson, R. S. Herati, B. Bengsch, S. L. Linderman, E. Stelekati, R. Spolski, E. J. Wherry, C. Hunter, S. E. Hensley, W. J. Leonard and M. P. Cancro (2016). "Cutting Edge: IL-4, IL-21, and IFN-gamma Interact To Govern T-bet and CD11c Expression in TLR-Activated B Cells." J Immunol **197**(4): 1023-1028.

Natarajan, K. N., Z. Miao, M. Jiang, X. Huang, H. Zhou, J. Xie, C. Wang, S. Qin, Z. Zhao, L. Wu, N. Yang, B. Li, Y. Hou, S. Liu and S. A. Teichmann (2019). "Comparative analysis of sequencing technologies for single-cell transcriptomics." Genome Biol **20**(1): 70.

Natkanski, E., W.-Y. Lee, B. Mistry, A. Casal, J. E. Molloy and P. Tolar (2013). "B cells use mechanical energy to discriminate antigen affinities." Science **340**(6140): 1587-1590.

Nehar-Belaid, D., S. Hong, R. Marches, G. Chen, M. Bolisetty, J. Baisch, L. Walters, M. Punaro, R. J. Rossi, C. H. Chung, R. P. Huynh, P. Singh, W. F. Flynn, J. A. Tabanor-Gayle, N. Kuchipudi, A. Mejias, M. A. Collet, A. L. Lucido, K. Palucka, P. Robson, S. Lakshminarayanan, O. Ramilo, T. Wright, V. Pascual and J. F. Banchereau (2020). "Mapping systemic lupus erythematosus heterogeneity at the single-cell level." Nat Immunol **21**(9): 1094-1106.

Netea, M. G., F. Balkwill, M. Chonchol, F. Cominelli, M. Y. Donath, E. J. Giamarellos-Bourboulis, D. Golenbock, M. S. Gresnigt, M. T. Heneka, H. M. Hoffman, R. Hotchkiss, L. A. B. Joosten, D. L. Kastner, M. Korte, E. Latz, P. Libby, T. Mandrup-Poulsen, A. Mantovani, K. H. G. Mills, K. L. Nowak, L. A. O'Neill, P. Pickkers, T. van der Poll, P. M. Ridker, J. Schalkwijk, D. A. Schwartz, B. Siegmund, C. J. Steer, H. Tilg, J. W. M. van der Meer, F. L. van de Veerdonk and C. A. Dinarello (2017). "A guiding map for inflammation." Nat Immunol **18**(8): 826-831.

Netea, M. G., A. Schlitzer, K. Placek, L. A. B. Joosten and J. L. Schultze (2019). "Innate and Adaptive Immune Memory: an Evolutionary Continuum in the Host's Response to Pathogens." Cell Host Microbe **25**(1): 13-26.

Nevarez Flores, C. G., A. P. Sun and H. Hast (2019). "Mowat-Wilson syndrome presenting with purpura fulminans." Pediatrics **143**(1).

Niessl, J., A. E. Baxter, A. Morou, E. Brunet-Ratnasingham, G. Sannier, G. Gendron-Lepage, J. Richard, G. G. Delgado, N. Brassard, I. Turcotte, R. Fromentin, N. F. Bernard, N. Chomont, J. P. Routy, M. Dube, A. Finzi and D. E. Kaufmann (2020). "Persistent expansion and Th1-like skewing of HIV-specific circulating T follicular helper cells during antiretroviral therapy." EBioMedicine **54**: 102727.

Nishizaki, Y., T. Takagi, F. Matsui and Y. Higashi (2014). "SIP1 expression patterns in brain investigated by generating a SIP1-EGFP reporter knock-in mouse." Genesis **52**(1): 56-67.

Nowosad, C. R., K. M. Spillane and P. Tolar (2016). "Germinal center B cells recognize antigen through a specialized immune synapse architecture." Nat Immunol **17**(7): 870-877.

Nurieva, R. I., Y. Chung, D. Hwang, X. O. Yang, H. S. Kang, L. Ma, Y. H. Wang, S. S. Watowich, A. M. Jetten, Q. Tian and C. Dong (2008). "Generation of T follicular helper cells is mediated by interleukin-21 but independent of T helper 1, 2, or 17 cell lineages." Immunity **29**(1): 138-149.

Nurieva, R. I., Y. Chung, G. J. Martinez, X. O. Yang, S. Tanaka, T. D. Matskevitch, Y.-H. Wang and C. Dong (2009). "Bcl6 mediates the development of T follicular helper cells." Science **325**(5943): 1001-1005.

Nussenzweig, R., J. Vanderberg, H. Most and C. Orton (1967). "Protective immunity produced by the injection of x-irradiated sporozoites of Plasmodium berghei." Nature **216**(5111): 160-162.

Nutt, S. L., P. D. Hodgkin, D. M. Tarlinton and L. M. Corcoran (2015). "The generation of antibody-secreting plasma cells." Nat Rev Immunol **15**(3): 160-171.

Obeng-Adjei, N., S. Portugal, P. Holla, S. Li, H. Sohn, A. Ambegaonkar, J. Skinner, G. Bowyer, O. K. Doumbo, B. Traore, S. K. Pierce and P. D. Crompton (2017). "Malaria-induced interferon-gamma drives the expansion of Tbethi atypical memory B cells." PLoS Pathog **13**(9): e1006576.

Obeng-Adjei, N., S. Portugal, T. M. Tran, T. B. Yazew, J. Skinner, S. Li, A. Jain, P. L. Felgner, O. K. Doumbo, K. Kayentao, A. Ongoiba, B. Traore and P. D. Crompton (2015). "Circulating Th1-Cell-type Tfh Cells that Exhibit Impaired B Cell Help Are Preferentially Activated during Acute Malaria in Children." Cell Rep **13**(2): 425-439.

Omilusik, K. D., J. A. Best, B. Yu, S. Goossens, A. Weidemann, J. V. Nguyen, E. Seuntjens, A. Stryjewska, C. Zweier, R. Roychoudhuri, L. Gattinoni, L. M. Bird, Y. Higashi, H. Kondoh, D. Huylebroeck, J. Haigh and A. W. Goldrath (2015). "Transcriptional repressor ZEB2

promotes terminal differentiation of CD8+ effector and memory T cell populations during infection." J Exp Med **212**(12): 2027-2039.

Oneko, M., L. C. Steinhardt, R. Yego, R. E. Wiegand, P. A. Swanson, N. Kc, D. Akach, T. Sang, J. R. Gutman, E. L. Nzuu, A. Dungani, B. Kim Lee Sim, P. N. Oloo, K. Otieno, D. K. Bii, P. F. Billingsley, E. R. James, S. Kariuki, A. M. Samuels, S. Jongo, W. Chebore, S. Abdulla, C. Daubenberger, M. Mpina, D. Styers, G. E. Potter, G. Abarbanell, T. L. Richie, S. L. Hoffman and R. A. Seder (2021). "Safety, immunogenicity and efficacy of PfSPZ Vaccine against malaria in infants in western Kenya: a double-blind, randomized, placebo-controlled phase 2 trial." Nat Med **27**(9): 1636-1645.

Oxenius, A., M. F. Bachmann, R. M. Zinkernagel and H. Hengartner (1998). "Virus-specific major MHC class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection." European journal of immunology **28**(1): 390-400.

Parra, D., A. M. Rieger, J. Li, Y. A. Zhang, L. M. Randall, C. A. Hunter, D. R. Barreda and J. O. Sunyer (2012). "Pivotal advance: peritoneal cavity B-1 B cells have phagocytic and microbicidal capacities and present phagocytosed antigen to CD4+ T cells." J Leukoc Biol **91**(4): 525-536.

Patel, R. S., J. E. Tomlinson, T. J. Divers, G. R. Van de Walle and B. R. Rosenberg (2021). "Single-cell resolution landscape of equine peripheral blood mononuclear cells reveals diverse cell types including T-bet(+) B cells." BMC Biol **19**(1): 13.

Perez-Mazliah, D., P. J. Gardner, E. Schweighoffer, S. McLaughlin, C. Hosking, I. Tumwine, R. S. Davis, A. J. Potocnik, V. L. Tybulewicz and J. Langhorne (2018). "Plasmodium-specific atypical memory B cells are short-lived activated B cells." Elife **7**.

Pieper, K., B. Grimbacher and H. Eibel (2013). "B-cell biology and development." J Allergy Clin Immunol **131**(4): 959-971.

Pilzecker, B. and H. Jacobs (2019). "Mutating for Good: DNA Damage Responses During Somatic Hypermutation." Front Immunol **10**: 438.

Polack, F. P., S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Perez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, Jr., L. L. Hammitt, O. Tureci, H. Nell, A. Schaefer, S. Unal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Sahin, K. U. Jansen, W. C. Gruber and C. C. T. Group (2020). "Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine." N Engl J Med **383**(27): 2603-2615.

Portugal, S., C. M. Tipton, H. Sohn, Y. Kone, J. Wang, S. Li, J. Skinner, K. Virtaneva, D. E. Sturdevant, S. F. Porcella, O. K. Doumbo, S. Doumbo, K. Kayentao, A. Ongoiba, B. Traore, I.

Sanz, S. K. Pierce and P. D. Crompton (2015). "Malaria-associated atypical memory B cells exhibit markedly reduced B cell receptor signaling and effector function." Elife **4**.

Radtke, A. J., W. Kastenmuller, D. A. Espinosa, M. Y. Gerner, S. W. Tse, P. Sinnis, R. N. Germain, F. P. Zavala and I. A. Cockburn (2015). "Lymph-node resident CD8alpha+ dendritic cells capture antigens from migratory malaria sporozoites and induce CD8+ T cell responses." PLoS Pathog **11**(2): e1004637.

Rakhmanov, M., B. Keller, S. Gutenberger, C. Foerster, M. Hoenig, G. Driessen, M. van der Burg, J. J. van Dongen, E. Wiech and M. Visentini (2009). "Circulating CD21low B cells in common variable immunodeficiency resemble tissue homing, innate-like B cells." Proceedings of the National Academy of Sciences **106**(32): 13451-13456.

Ramesh, A., R. D. Schubert, A. L. Greenfield, R. Dandekar, R. Loudermilk, J. J. Sabatino, Jr., M. T. Koelzer, E. B. Tran, K. Koshal, K. Kim, A. K. Probstel, D. Banerji, S. F. M. S. E. T. University of California, C. Y. Guo, A. J. Green, R. M. Bove, J. L. DeRisi, J. M. Gelfand, B. A. C. Cree, S. S. Zamvil, S. E. Baranzini, S. L. Hauser and M. R. Wilson (2020). "A pathogenic and clonally expanded B cell transcriptome in active multiple sclerosis." Proc Natl Acad Sci U S A **117**(37): 22932-22943.

Rauschmeier, R., A. Reinhardt, C. Gustafsson, V. Glaros, A. V. Artemov, J. Dunst, R. Taneja, I. Adameyko, R. Mansson, M. Busslinger and T. Kreslavsky (2022). "Bhlhe40 function in activated B and TFH cells restrains the GC reaction and prevents lymphomagenesis." J Exp Med **219**(2).

Read, K. A., M. D. Powell, B. K. Sreekumar and K. J. Oestreich (2019). "In Vitro Differentiation of Effector CD4(+) T Helper Cell Subsets." Methods Mol Biol **1960**: 75-84.

Real, E., V. M. Howick, F. A. Dahalan, K. Witmer, J. Cudini, C. Andradi-Brown, J. Blight, M. S. Davidson, S. K. Dogga, A. J. Reid, J. Baum and M. K. N. Lawniczak (2021). "A single-cell atlas of Plasmodium falciparum transmission through the mosquito." Nat Commun **12**(1): 3196.

Reiss, S., A. E. Baxter, K. M. Cirelli, J. M. Dan, A. Morou, A. Daigneault, N. Brassard, G. Silvestri, J. P. Routy, C. Havenar-Daughton, S. Crotty and D. E. Kaufmann (2017). "Comparative analysis of activation induced marker (AIM) assays for sensitive identification of antigen-specific CD4 T cells." PLoS One **12**(10): e0186998.

Ricker, E., M. Manni, D. Flores-Castro, D. Jenkins, S. Gupta, J. Rivera-Correa, W. Meng, A. M. Rosenfeld, T. Pannellini, M. Bachu, Y. Chinenov, P. K. Sculco, R. Jessberger, E. T. L. Prak and A. B. Pernis (2021). "Altered function and differentiation of age-associated B cells contribute to the female bias in lupus mice." Nat Commun **12**(1): 4813.

Rickert, R. C. (2013). "New insights into pre-BCR and BCR signalling with relevance to B cell malignancies." Nat Rev Immunol **13**(8): 578-591.

Rivera-Correa, J., J. J. Guthmiller, R. Vijay, C. Fernandez-Arias, M. A. Pardo-Ruge, S. Gonzalez, N. S. Butler and A. Rodriguez (2017). "Plasmodium DNA-mediated TLR9 activation of T-bet(+) B cells contributes to autoimmune anaemia during malaria." Nat Commun **8**(1): 1282.

Rivera-Correa, J., M. S. Mackroth, T. Jacobs, J. Schulze Zur Wiesch, T. Rolling and A. Rodriguez (2019). "Atypical memory B-cells are associated with Plasmodium falciparum anemia through anti-phosphatidylserine antibodies." Elife **8**.

Roper, S. I., L. Wasim, D. Malinova, M. Way, S. Cox and P. Tolar (2019). "B cells extract antigens at Arp2/3-generated actin foci interspersed with linear filaments." Elife **8**.

RTS,S Clinical Trials Partnership (2015). "Efficacy and safety of RTS, S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial." The Lancet **386**(9988): 31-45.

Rubtsov, A. V., K. Rubtsova, A. Fischer, R. T. Meehan, J. Z. Gillis, J. W. Kappler and P. Marrack (2011). "Toll-like receptor 7 (TLR7)-driven accumulation of a novel CD11c(+) B-cell population is important for the development of autoimmunity." Blood **118**(5): 1305-1315.

Rubtsov, A. V., K. Rubtsova, J. W. Kappler, J. Jacobelli, R. S. Friedman and P. Marrack (2015). "CD11c-Expressing B Cells Are Located at the T Cell/B Cell Border in Spleen and Are Potent APCs." J Immunol **195**(1): 71-79.

Rubtsova, K., A. V. Rubtsov, L. F. van Dyk, J. W. Kappler and P. Marrack (2013). "T-box transcription factor T-bet, a key player in a unique type of B-cell activation essential for effective viral clearance." Proc Natl Acad Sci U S A **110**(34): E3216-3224.

Russell Knode, L. M., M. S. Naradikian, A. Myles, J. L. Scholz, Y. Hao, D. Liu, M. L. Ford, J. W. Tobias, M. P. Cancro and P. J. Gearhart (2017). "Age-Associated B Cells Express a Diverse Repertoire of VH and Vkappa Genes with Somatic Hypermutation." J Immunol **198**(5): 1921-1927.

Rydyznski Moderbacher, C., S. I. Ramirez, J. M. Dan, A. Grifoni, K. M. Hastie, D. Weiskopf, S. Belanger, R. K. Abbott, C. Kim, J. Choi, Y. Kato, E. G. Crotty, C. Kim, S. A. Rawlings, J. Mateus, L. P. V. Tse, A. Frazier, R. Baric, B. Peters, J. Greenbaum, E. Ollmann Saphire, D. M. Smith, A. Sette and S. Crotty (2020). "Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity." Cell **183**(4): 996-1012 e1019.

Sage, P. T., D. Alvarez, J. Godec, U. H. von Andrian and A. H. Sharpe (2014). "Circulating T follicular regulatory and helper cells have memory-like properties." J Clin Invest **124**(12): 5191-5204.

Sahin, U., A. Muik, E. Derhovanessian, I. Vogler, L. M. Kranz, M. Vormehr, A. Baum, K. Pascal, J. Quandt, D. Maurus, S. Brachtendorf, V. Lorks, J. Sikorski, R. Hilker, D. Becker, A. K. Eller, J. Grutzner, C. Boesler, C. Rosenbaum, M. C. Kuhnle, U. Luxemburger, A. Kemmer-Bruck, D. Langer, M. Bexon, S. Bolte, K. Kariko, T. Palanche, B. Fischer, A. Schultz, P. Y. Shi, C. Fontes-Garfias, J. L. Perez, K. A. Swanson, J. Loschko, I. L. Scully, M. Cutler, W. Kalina, C. A. Kyratsous, D. Cooper, P. R. Dormitzer, K. U. Jansen and O. Tureci (2020). "COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses." Nature **586**(7830): 594-599.

Sallusto, F., A. Cassotta, D. Hoces, M. Foglierini and A. Lanzavecchia (2018). "Do Memory CD4 T Cells Keep Their Cell-Type Programming: Plasticity versus Fate Commitment? T-Cell Heterogeneity, Plasticity, and Selection in Humans." Cold Spring Harb Perspect Biol **10**(3).

Sallusto, F., J. Geginat and A. Lanzavecchia (2004). "Central memory and effector memory T cell subsets: function, generation, and maintenance." Annu Rev Immunol **22**: 745-763.

Sartorius, B., J. Cano, H. Simpson, L. S. Tusting, L. B. Marczak, M. K. Miller-Petrie, B. Kinvi, H. Zoure, P. Mwinzi, S. I. Hay, M. Rebollo and R. L. Pullan (2021). "Prevalence and intensity of soil-transmitted helminth infections of children in sub-Saharan Africa, 2000–18: a geospatial analysis." The Lancet Global Health **9**(1): e52-e60.

Schaerli, P., K. Willimann, A. B. Lang, M. Lipp, P. Loetscher and B. Moser (2000). "CXC chemokine receptor 5 expression defines follicular homing T cells with B cell helper function." The Journal of experimental medicine **192**(11): 1553-1562.

Schnell, A., L. Huang, M. Singer, A. Singaraju, R. M. Barilla, B. M. Regan, A. Bollhagen, P. I. Thakore, D. Dionne and T. M. Delorey (2021). "Stem-like intestinal Th17 cells give rise to pathogenic effector T cells during autoimmunity." Cell **184**(26): 6281-6298. e6223.

Scott, C. L. and K. D. Omilusik (2019). "ZEBs: Novel Players in Immune Cell Development and Function." Trends Immunol **40**(5): 431-446.

Scott, C. L., B. Soen, L. Martens, N. Skrypek, W. Saelens, J. Taminau, G. Blancke, G. Van Isterdael, D. Huylebroeck, J. Haigh, Y. Saeys, M. Guilliams, B. N. Lambrecht and G. Berx (2016). "The transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing Id2." J Exp Med **213**(6): 897-911.

Seder, R. A., L.-J. Chang, M. E. Enama, K. L. Zephir, U. N. Sarwar, I. J. Gordon, L. A. Holman, E. R. James, P. F. Billingsley and A. Gunasekera (2013). "Protection against malaria by

intravenous immunization with a nonreplicating sporozoite vaccine." Science **341**(6152): 1359-1365.

Sette, A., M. Moutaftsi, J. Moyron-Quiroz, M. M. McCausland, D. H. Davies, R. J. Johnston, B. Peters, M. Rafii-El-Idrissi Benhnia, J. Hoffmann, H. P. Su, K. Singh, D. N. Garboczi, S. Head, H. Grey, P. L. Felgner and S. Crotty (2008). "Selective CD4+ T cell help for antibody responses to a large viral pathogen: deterministic linkage of specificities." Immunity **28**(6): 847-858.

Shaffer, A., X. Yu, Y. He, J. Boldrick, E. P. Chan and L. M. Staudt (2000). "BCL-6 represses genes that function in lymphocyte differentiation, inflammation, and cell cycle control." Immunity **13**(2): 199-212.

Shahinian, A., K. Pfeffer, K. P. Lee, T. M. Kündig, K. Kishihara, A. Wakeham, K. Kawai, P. S. Ohashi, C. B. Thompson and T. W. Mak (1993). "Differential T cell costimulatory requirements in CD28-deficient mice." Science **261**(5121): 609-612.

Sheiner, L., A. B. Vaidya and G. I. McFadden (2013). "The metabolic roles of the endosymbiotic organelles of Toxoplasma and Plasmodium spp." Curr Opin Microbiol **16**(4): 452-458.

Sidjanski, S. and J. P. Vanderberg (1997). "Delayed migration of Plasmodium sporozoites from the mosquito bite site to the blood." The American journal of tropical medicine and hygiene **57**(4): 426-429.

Singer, M. and F. Frischknecht (2021). "Fluorescent tagging of Plasmodium circumsporozoite protein allows imaging of sporozoite formation but blocks egress from oocysts." Cell Microbiol **23**(5): e13321.

Sissoko, M. S., S. A. Healy, A. Katile, F. Omaswa, I. Zaidi, E. E. Gabriel, B. Kamate, Y. Samake, M. A. Guindo and A. Dolo (2017). "Safety and efficacy of PfSPZ Vaccine against Plasmodium falciparum via direct venous inoculation in healthy malaria-exposed adults in Mali: a randomised, double-blind phase 1 trial." The Lancet infectious diseases **17**(5): 498-509.

Sissoko, M. S., S. A. Healy, A. Katile, F. Omaswa, I. Zaidi, E. E. Gabriel, B. Kamate, Y. Samake, M. A. Guindo, A. Dolo, A. Niangaly, K. Niaré, A. Zeguime, K. Sissoko, H. Diallo, I. Thera, K. Ding, M. P. Fay, E. M. O'Connell, T. B. Nutman, S. Wong-Madden, T. Murshedkar, A. J. Ruben, M. Li, Y. Abebe, A. Manoj, A. Gunasekera, S. Chakravarty, B. K. L. Sim, P. F. Billingsley, E. R. James, M. Walther, T. L. Richie, S. L. Hoffman, O. Doumbo and P. E. Duffy (2017). "Safety and efficacy of PfSPZ Vaccine against Plasmodium falciparum via direct venous inoculation in healthy malaria-exposed adults in Mali: a randomised, double-blind phase 1 trial." The Lancet Infectious Diseases **17**(5): 498-509.

Song, W., O. Q. Antao, E. Condiff, G. M. Sanchez, I. Chernova, K. Zembrzuski, H. Steach, K. Rubtsova, D. Angeletti, A. Lemenze, B. J. Laidlaw, J. Craft and J. S. Weinstein (2022). "Development of Tbet- and CD11c-expressing B cells in a viral infection requires T follicular helper cells outside of germinal centers." Immunity **55**(2): 290-307 e295.

Spellberg, B. and J. E. Edwards Jr (2001). "Type 1/Type 2 immunity in infectious diseases." Clinical Infectious Diseases **32**(1): 76-102.

Spensieri, F., E. Borgogni, L. Zedda, M. Bardelli, F. Buricchi, G. Volpini, E. Fragapane, S. Tavarini, O. Finco, R. Rappuoli, G. Del Giudice, G. Galli and F. Castellino (2013). "Human circulating influenza-CD4+ ICOS1+IL-21+ T cells expand after vaccination, exert helper function, and predict antibody responses." Proc Natl Acad Sci U S A **110**(35): 14330-14335.

Stewart, A., J. C. Ng, G. Wallis, V. Tsioligka, F. Fraternali and D. K. Dunn-Walters (2021). "Single-Cell Transcriptomic Analyses Define Distinct Peripheral B Cell Subsets and Discrete Development Pathways." Front Immunol **12**: 602539.

Stoeckius, M., C. Hafemeister, W. Stephenson, B. Houck-Loomis, P. K. Chattopadhyay, H. Swerdlow, R. Satija and P. Smibert (2017). "Simultaneous epitope and transcriptome measurement in single cells." Nat Methods **14**(9): 865-868.

Sturm, A., R. Amino, C. Van de Sand, T. Regen, S. Retzlaff, A. Rennenberg, A. Krueger, J. r.-M. Pollok, R. Menard and V. T. Heussler (2006). "Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids." science **313**(5791): 1287-1290.

Sullivan, R. T., C. C. Kim, M. F. Fontana, M. E. Feeney, P. Jagannathan, M. J. Boyle, C. J. Drakeley, I. Ssewanyana, F. Nankya, H. Mayanja-Kizza, G. Dorsey and B. Greenhouse (2015). "FCRL5 Delineates Functionally Impaired Memory B Cells Associated with Plasmodium falciparum Exposure." PLoS Pathog **11**(5): e1004894.

Sun, Z., D. Unutmaz, Y.-R. Zou, M. J. Sunshine, A. Pierani, S. Brenner-Morton, R. E. Mebius and D. R. Littman (2000). "Requirement for RORγ in Thymocyte Survival and Lymphoid Organ Development." Science **288**(5475): 2369.

Sundling, C., C. Ronnberg, V. Yman, M. Asghar, P. Jahnmatz, T. Lakshmikanth, Y. Chen, J. Mikes, M. N. Forsell, K. Sonden, A. Achour, P. Brodin, K. E. Persson and A. Farnert (2019). "B cell profiling in malaria reveals expansion and remodelling of CD11c+ B cell subsets." JCI Insight **5**.

Sutton, H. J., R. Aye, A. H. Idris, R. Vistein, E. Nduati, O. Kai, J. Mwacharo, X. Li, X. Gao, T. D. Andrews, M. Koutsakos, T. H. O. Nguyen, M. Nekrasov, P. Milburn, A. Eltahla, A. A. Berry, N. Kc, S. Chakravarty, B. K. L. Sim, A. K. Wheatley, S. J. Kent, S. L. Hoffman, K. E. Lyke, P. Bejon, F. Luciani, K. Kedzierska, R. A. Seder, F. M. Ndungu and I. A. Cockburn (2021).

"Atypical B cells are part of an alternative lineage of B cells that participates in responses to vaccination and infection in humans." Cell Rep **34**(6): 108684.

Suzuki, K., I. Grigorova, T. G. Phan, L. M. Kelly and J. G. Cyster (2009). "Visualizing B cell capture of cognate antigen from follicular dendritic cells." J Exp Med **206**(7): 1485-1493.

Swearingen, K. E., S. E. Lindner, L. Shi, M. J. Shears, A. Harupa, C. S. Hopp, A. M. Vaughan, T. A. Springer, R. L. Moritz, S. H. Kappe and P. Sinnis (2016). "Interrogating the Plasmodium Sporozoite Surface: Identification of Surface-Exposed Proteins and Demonstration of Glycosylation on CSP and TRAP by Mass Spectrometry-Based Proteomics." PLoS Pathog **12**(4): e1005606.

Tang, F., C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Lao and M. A. Surani (2009). "mRNA-Seq whole-transcriptome analysis of a single cell." Nat Methods **6**(5): 377-382.

Tauzin, A., M. Nayrac, M. Benlarbi, S. Y. Gong, R. Gasser, G. Beaudoin-Bussières, N. Brassard, A. Laumaea, D. Vézina and J. Prévost (2021). "A single dose of the SARS-CoV-2 vaccine BNT162b2 elicits Fc-mediated antibody effector functions and T cell responses." Cell host & microbe **29**(7): 1137-1150. e1136.

Thai, E., G. Costa, A. Weyrich, R. Murugan, D. Oyen, Y. Flores-Garcia, K. Prieto, A. Bosch, A. Valleriani, N. C. Wu, T. Pholcharee, S. W. Scally, I. A. Wilson, H. Wardemann, J. P. Julien and E. A. Levashina (2020). "A high-affinity antibody against the CSP N-terminal domain lacks Plasmodium falciparum inhibitory activity." J Exp Med **217**(11).

Thompson, E., K. Roznik, A. H. Karaba, K. Cascino, L. Biavati, L. A. Sena, S. Dahkal, W. Ngecu, A. T. Abedon and J. L. Alejo "Lipid-Oxidizing B Cells Enable Successful Vaccine Response Despite Immunosuppression."

Tsai, L. M. and D. Yu (2014). "Follicular helper T-cell memory: establishing new frontiers during antibody response." Immunol Cell Biol **92**(1): 57-63.

Turner, J. S., M. Marthi, Z. L. Benet and I. Grigorova (2017). "Transiently antigen-primed B cells return to naive-like state in absence of T-cell help." Nat Commun **8**: 15072.

Turner, J. S., J. Q. Zhou, J. Han, A. J. Schmitz, A. A. Rizk, W. B. Alsoussi, T. Lei, M. Amor, K. M. McIntire, P. Meade, S. Strohmeier, R. I. Brent, S. T. Richey, A. Haile, Y. R. Yang, M. K. Klebert, T. Suessen, S. Teefey, R. M. Presti, F. Krammer, S. H. Kleinstein, A. B. Ward and A. H. Ellebedy (2020). "Human germinal centres engage memory and naive B cells after influenza vaccination." Nature **586**(7827): 127-132.

Van Damme, P., M. Dionne, G. Leroux-Roels, O. Van Der Meeren, E. Di Paolo, B. Salaun, P. Surya Kiran and N. Folschweiller (2019). "Persistence of HBsAg-specific antibodies and

immune memory two to three decades after hepatitis B vaccination in adults." J Viral Hepat **26**(9): 1066-1075.

van Helden, M. J., S. Goossens, C. Daussy, A. L. Mathieu, F. Faure, A. Marcais, N. Vandamme, N. Farla, K. Mayol, S. Viel, S. Degouve, E. Debien, E. Seuntjens, A. Conidi, J. Chaix, P. Mangeot, S. de Bernard, L. Buffat, J. J. Haigh, D. Huylebroeck, B. N. Lambrecht, G. Berx and T. Walzer (2015). "Terminal NK cell maturation is controlled by concerted actions of T-bet and Zeb2 and is essential for melanoma rejection." J Exp Med **212**(12): 2015-2025.

Victora, G. D. and M. C. Nussenzweig (2022). "Germinal Centers." Annu Rev Immunol **40**: 413-442.

Victora, G. D., T. A. Schwickert, D. R. Fooksman, A. O. Kamphorst, M. Meyer-Hermann, M. L. Dustin and M. C. Nussenzweig (2010). "Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter." Cell **143**(4): 592-605.

Vijay, R., J. J. Guthmiller, A. J. Sturtz, F. A. Surette, K. J. Rogers, R. R. Sompallae, F. Li, R. L. Pope, J. A. Chan, F. de Labastida Rivera, D. Andrew, L. Webb, W. J. Maury, H. H. Xue, C. R. Engwerda, J. S. McCarthy, M. J. Boyle and N. S. Butler (2020). "Infection-induced plasmablasts are a nutrient sink that impairs humoral immunity to malaria." Nat Immunol **21**(7): 790-801.

Vinuesa, C. G., M. A. Linterman, D. Yu and I. C. MacLennan (2016). "Follicular Helper T Cells." Annu Rev Immunol **34**: 335-368.

Wahl, I., A. S. Obraztsova, J. Puchan, R. Hundsdorfer, S. Chakravarty, B. K. L. Sim, S. L. Hoffman, P. G. Kremsner, B. Mordmüller and H. Wardemann (2022). "Clonal evolution and TCR specificity of the human TFH cell response to Plasmodium falciparum CSP." Science Immunology **7**(72): eabm9644.

Walker, L. S. (2013). "Treg and CTLA-4: two intertwining pathways to immune tolerance." J Autoimmun **45**: 49-57.

Wang, S., J. Wang, V. Kumar, J. L. Karnell, B. Naiman, P. S. Gross, S. Rahman, K. Zerrouki, R. Hanna, C. Morehouse, N. Holoweckyj, H. Liu, T. Autoimmunity Molecular Medicine, Z. Manna, R. Goldbach-Mansky, S. Hasni, R. Siegel, M. Sanjuan, K. Streicher, M. P. Cancro, R. Kolbeck and R. Ettinger (2018). "IL-21 drives expansion and plasma cell differentiation of autoreactive CD11c(hi)T-bet(+) B cells in SLE." Nat Commun **9**(1): 1758.

Weber, J. P., F. Fuhrmann and A. Hutloff (2012). "T-follicular helper cells survive as long-term memory cells." Eur J Immunol **42**(8): 1981-1988.

Wei, C., J. Anolik, A. Cappione, B. Zheng, A. Pugh-Bernard, J. Brooks, E. H. Lee, E. C. Milner and I. Sanz (2007). "A new population of cells lacking expression of CD27 represents a notable

component of the B cell memory compartment in systemic lupus erythematosus." J Immunol **178**(10): 6624-6633.

Weinstein, J. S., S. G. Hernandez and J. Craft (2012). "T cells that promote B-Cell maturation in systemic autoimmunity." Immunological reviews **247**(1): 160-171.

Weisel, F. and M. Shlomchik (2017). "Memory B Cells of Mice and Humans." Annu Rev Immunol **35**: 255-284.

Weiss, D. J., T. C. D. Lucas, M. Nguyen, A. K. Nandi, D. Bisanzio, K. E. Battle, E. Cameron, K. A. Twohig, D. A. Pfeffer, J. A. Rozier, H. S. Gibson, P. C. Rao, D. Casey, A. Bertozzi-Villa, E. L. Collins, U. Dalrymple, N. Gray, J. R. Harris, R. E. Howes, S. Y. Kang, S. H. Keddie, D. May, S. Rumisha, M. P. Thorn, R. Barber, N. Fullman, C. K. Huynh, X. Kulikoff, M. J. Kutz, A. D. Lopez, A. H. Mokdad, M. Naghavi, G. Nguyen, K. A. Shackelford, T. Vos, H. Wang, D. L. Smith, S. S. Lim, C. J. L. Murray, S. Bhatt, S. I. Hay and P. W. Gething (2019). "Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study." The Lancet **394**(10195): 322-331.

Weiss, G. E., P. D. Crompton, S. Li, L. A. Walsh, S. Moir, B. Traore, K. Kayentao, A. Ongoiba, O. K. Doumbo and S. K. Pierce (2009). "Atypical memory B cells are greatly expanded in individuals living in a malaria-endemic area." J Immunol **183**(3): 2176-2182.

White, M. T., R. Verity, J. T. Griffin, K. P. Asante, S. Owusu-Agyei, B. Greenwood, C. Drakeley, S. Gesase, J. Lusingu, D. Ansong, S. Adjei, T. Agbenyega, B. Ogutu, L. Otieno, W. Otieno, S. T. Agnandji, B. Lell, P. Kremsner, I. Hoffman, F. Martinson, P. Kamthunzu, H. Tinto, I. Valea, H. Sorgho, M. Oneko, K. Otieno, M. J. Hamel, N. Salim, A. Mtoro, S. Abdulla, P. Aide, J. Sacarlal, J. J. Aponte, P. Njuguna, K. Marsh, P. Bejon, E. M. Riley and A. C. Ghani (2015). "Immunogenicity of the RTS,S/AS01 malaria vaccine and implications for duration of vaccine efficacy: secondary analysis of data from a phase 3 randomised controlled trial." Lancet Infect Dis **15**(12): 1450-1458.

Woodruff, M. C., R. P. Ramonell, D. C. Nguyen, K. S. Cashman, A. S. Saini, N. S. Haddad, A. M. Ley, S. Kyu, J. C. Howell, T. Ozturk, S. Lee, N. Suryadevara, J. B. Case, R. Bugrovsky, W. Chen, J. Estrada, A. Morrison-Porter, A. Derrico, F. A. Anam, M. Sharma, H. M. Wu, S. N. Le, S. A. Jenks, C. M. Tipton, B. Staitieh, J. L. Daiss, E. Ghosn, M. S. Diamond, R. H. Carnahan, J. E. Crowe, Jr., W. T. Hu, F. E. Lee and I. Sanz (2020). "Extrafollicular B cell responses correlate with neutralizing antibodies and morbidity in COVID-19." Nat Immunol **21**(12): 1506-1516.

Wragg, K. M., W. S. Lee, M. Koutsakos, H.-X. Tan, T. Amarasena, A. Reynaldi, G. Gare, P. Konstandopoulos, K. R. Field and R. Esterbauer (2022). "Establishment and recall of SARS-CoV-2 spike epitope-specific CD4+ T cell memory." Nature Immunology **23**(5): 768-780.

Wu, X., C. G. Briseno, G. E. Grajales-Reyes, M. Haldar, A. Iwata, N. M. Kretzer, W. Kc, R. Tussiwand, Y. Higashi, T. L. Murphy and K. M. Murphy (2016). "Transcription factor Zeb2 regulates commitment to plasmacytoid dendritic cell and monocyte fate." Proc Natl Acad Sci U S A **113**(51): 14775-14780.

Wu, X., J. Tian and S. Wang (2018). "Insight Into Non-Pathogenic Th17 Cells in Autoimmune Diseases." Front Immunol **9**: 1112.

Xu, H., L. N. Liew, I. C. Kuo, C. H. Huang, D. L. Goh and K. Y. Chua (2008). "The modulatory effects of lipopolysaccharide-stimulated B cells on differential T-cell polarization." Immunology **125**(2): 218-228.

Yaari, G. and S. H. Kleinstein (2015). "Practical guidelines for B-cell receptor repertoire sequencing analysis." Genome Med **7**: 121.

Yang, R., D. T. Avery, K. J. Jackson, M. Ogishi, I. Benhsaien, L. Du, X. Ye, J. Han, J. Rosain and J. N. Peel (2022). "Human T-bet governs the generation of a distinct subset of CD11chighCD21low B cells." Science Immunology **7**(73): eabq3277.

Yao, Y., C. L. Chen, D. Yu and Z. Liu (2021). "Roles of follicular helper and regulatory T cells in allergic diseases and allergen immunotherapy." Allergy **76**(2): 456-470.

Yao, Y., Z. Chen, H. Zhang, C. Chen, M. Zeng, J. Yunis, Y. Wei, Y. Wan, N. Wang, M. Zhou, C. Qiu, Q. Zeng, H. S. Ong, H. Wang, F. V. Makota, Y. Yang, Z. Yang, N. Wang, J. Deng, C. Shen, Y. Xia, L. Yuan, Z. Lian, Y. Deng, C. Guo, A. Huang, P. Zhou, H. Shi, W. Zhang, H. Yi, D. Li, M. Xia, J. Fu, N. Wu, J. B. de Haan, N. Shen, W. Zhang, Z. Liu and D. Yu (2021). "Selenium-GPX4 axis protects follicular helper T cells from ferroptosis." Nat Immunol **22**(9): 1127-1139.

Yokosuka, T., K. Takase, M. Suzuki, Y. Nakagawa, S. Taki, H. Takahashi, T. Fujisawa, H. Arase and T. Saito (2002). "Predominant role of T cell receptor (TCR)-alpha chain in forming preimmune TCR repertoire revealed by clonal TCR reconstitution system." J Exp Med **195**(8): 991-1001.

Yost, K. E., A. T. Satpathy, D. K. Wells, Y. Qi, C. Wang, R. Kageyama, K. L. McNamara, J. M. Granja, K. Y. Sarin, R. A. Brown, R. K. Gupta, C. Curtis, S. L. Bucktrout, M. M. Davis, A. L. S. Chang and H. Y. Chang (2019). "Clonal replacement of tumor-specific T cells following PD-1 blockade." Nat Med **25**(8): 1251-1259.

Young, B., S. Sadarangani, L. Jiang, A. Wilder-Smith and M. I. Chen (2018). "Duration of Influenza Vaccine Effectiveness: A Systematic Review, Meta-analysis, and Meta-regression of Test-Negative Design Case-Control Studies." J Infect Dis **217**(5): 731-741.

Yu, D., S. Rao, L. M. Tsai, S. K. Lee, Y. He, E. L. Sutcliffe, M. Srivastava, M. Linterman, L. Zheng and N. Simpson (2009). "The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment." Immunity **31**(3): 457-468.

Yu, D., L. S. K. Walker, Z. Liu, M. A. Linterman and Z. Li (2022). "Targeting TFH cells in human diseases and vaccination: rationale and practice." Nat Immunol.

Yuseff, M. I., A. Reversat, D. Lankar, J. Diaz, I. Fanget, P. Pierobon, V. Randrian, N. Larochette, F. Vascotto, C. Desdouets, B. Jauffred, Y. Bellaiche, S. Gasman, F. Darchen, C. Desnos and A. M. Lennon-Dumenil (2011). "Polarized secretion of lysosomes at the B cell synapse couples antigen extraction to processing and presentation." Immunity **35**(3): 361-374.

Zhang, C., T. X. Zhang, Y. Liu, D. Jia, P. Zeng, C. Du, M. Yuan, Q. Liu, Y. Wang and F. D. Shi (2021). "B-Cell Compartmental Features and Molecular Basis for Therapy in Autoimmune Disease." Neurol Neuroimmunol Neuroinflamm **8**(6).

Zhang, F., K. Wei, K. Slowikowski, C. Y. Fonseka, D. A. Rao, S. Kelly, S. M. Goodman, D. Tabechian, L. B. Hughes, K. Salomon-Escoto, G. F. M. Watts, A. H. Jonsson, J. Rangel-Moreno, N. Meednu, C. Rozo, W. Apruzzese, T. M. Eisenhaure, D. J. Lieb, D. L. Boyle, A. M. Mandelin, 2nd, A. Accelerating Medicines Partnership Rheumatoid, C. Systemic Lupus Erythematosus, B. F. Boyce, E. DiCarlo, E. M. Gravallese, P. K. Gregersen, L. Moreland, G. S. Firestein, N. Hacohen, C. Nusbaum, J. A. Lederer, H. Perlman, C. Pitzalis, A. Filer, V. M. Holers, V. P. Bykerk, L. T. Donlin, J. H. Anolik, M. B. Brenner and S. Raychaudhuri (2019). "Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry." Nat Immunol **20**(7): 928-942.

Zhang, M., R. Mandraju, U. Rai, T. Shiratsuchi and M. Tsuji (2017). "Monoclonal Antibodies against Plasmodium falciparum Circumsporozoite Protein." Antibodies (Basel) **6**(3).

Zhang, W., H. Zhang, S. Liu, F. Xia, Z. Kang, Y. Zhang, Y. Liu, H. Xiao, L. Chen, C. Huang, N. Shen, H. Xu and F. Li (2019). "Excessive CD11c(+)Tbet(+) B cells promote aberrant TFH differentiation and affinity-based germinal center selection in lupus." Proc Natl Acad Sci U S A **116**(37): 18550-18560.

Zhao, J., Y. Chen, Q. Zhao, J. Shi, W. Yang, Z. Zhu, W. Yu, J. Guan, Y. Song, H. Wu, W. Jin, Y. Zhou and J. Liu (2018). "Increased circulating Tfh17 and PD-1+Tfh cells are associated with autoantibodies in Hashimoto's thyroiditis." Autoimmunity **51**(7): 352-359.

Zhao, J., S. Zhang, Y. Liu, X. He, M. Qu, G. Xu, H. Wang, M. Huang, J. Pan, Z. Liu, Z. Li, L. Liu and Z. Zhang (2020). "Single-cell RNA sequencing reveals the heterogeneity of liver-resident immune cells in human." Cell Discov **6**: 22.

Zhao, Z., C. Li, F. Tong, J. Deng, G. Huang and Y. Sang (2021). "Review of applications of CRISPR-Cas9 gene-editing technology in cancer research." Biol Proced Online **23**(1): 14.

Zotos, D., J. M. Coquet, Y. Zhang, A. Light, K. D'Costa, A. Kallies, L. M. Corcoran, D. I. Godfrey, K. M. Toellner, M. J. Smyth, S. L. Nutt and D. M. Tarlinton (2010). "IL-21 regulates germinal center B cell differentiation and proliferation through a B cell-intrinsic mechanism." J Exp Med **207**(2): 365-378.

Zotos, D. and D. M. Tarlinton (2012). "Determining germinal centre B cell fate." Trends in immunology **33**(6): 281-288.

Zuccarino-Catania, G. and M. Shlomchik (2015). "Adoptive Transfer of Memory B Cells." Bio-protocol **5**(16): e1563.

Code for Figure 3.10, 3.11

```
#Open seurat and load the Seurat samples (can be downloaded from
https://drive.google.com/drive/folders/1aMo8RgZt9bZy5tc94Kr1x-
wuzUoRbOCz)
library(Seurat)
load("~/Desktop/CD4T6_seurat.rdata")
mycells = UpdateSeuratObject(object = mycells)
virus<-mycells@meta.data
#Also load the annotation files including sample source and TCR
information (can be downloaded from
https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE152522)
library(readr)
cd4t6_annotation <- read_table2("Desktop/cd4t6_annotation.txt")
TRA.aa.chains.tag<-cd4t6_annotation$TRA.aa.chains.tag
TRB.aa.chains.tag<-cd4t6_annotation$TRB.aa.chains.tag
mycells<-AddMetaData(mycells,metadata =TRA.aa.chains.tag,col.name =
"TRA.aa.chains.tag")
mycells<-AddMetaData(mycells,metadata =TRB.aa.chains.tag,col.name =
"TRB.aa.chains.tag")
#Extract cells that belongs to influenza vaccine cohort
mycells<-subset(x=mycells, subset=orig.virus2=="FLU")
#Extract singlets
mycells<-subset(x=mycells,
subset=orig.MULTI_classification.global=="Singlet")
#Recluster the cells
mycells <- RunPCA(mycells, verbose = FALSE)
ElbowPlot(mycells, ndims =50)
mycells <- FindNeighbors(mycells, dims = 1:9, reduction="pca")
mycells <- RunUMAP(mycells, reduction = "pca", dims = 1:9)
mycells <- FindClusters(mycells, resolution = 0.7,
reduction.type="pca")
FeaturePlot(mycells,"CXCR5",order = T,pt.size = 2)
FeaturePlot(mycells,"CD69",order = T,pt.size = 2,min.cutoff = 3) #There
seems to have a more activated population, perhaps due to the antigen
stimulation
DimPlot(mycells,label = T)
#Extract cTfh population
mycells <- subset(x = mycells, idents=c("2","3","5","4"))
#Calculate Tfh1/Tfh17 signature gene scores on each cell, the genes
were showed in Figure S5A
Tfh1geneset <- read.csv("~/Desktop/Tfh1_up.csv")
mycells <- AddModuleScore(object = mycells,features = Tfh1geneset,name
= 'Tfh1score', assay = "RNA")
Tfh17geneset <- read.csv("~/Desktop/Tfh17_up.csv")
mycells <- AddModuleScore(object = mycells,features = Tfh17geneset,name
= 'Tfh17score')
#Exporting values used for plotting
aCDR3<-mycells@meta.data[["TRA.aa.chains.tag"]]
bCDR3<-mycells@meta.data[["TRB.aa.chains.tag"]]
donor<-mycells@meta.data[["orig.donor"]]
Tfh1score<-mycells@meta.data[["Tfh1score1"]]
Tfh17score<-mycells@meta.data[["Tfh17score1"]]
pre_post<-mycells@meta.data[["orig.virus"]]
#Generation of table used for plotting
TCRanalysis<-cbind(Tfh1score,Tfh17score,aCDR3,bCDR3,donor,pre_post)
write.csv(TCRanalysis,file="TCRanalysis.csv")
#Calculate the clonal abundance and to identify most abundant clones
```

```
TCRanalysis<-as.data.frame(TCRanalysis)
table1<-table(TCRanalysis$aCDR3,TCRanalysis$donor)
View(table1)
write.csv(table1,file="TCRanalysisabundance.csv")
#Then use excel to separate different blood donors; separate
before/after vaccination; select the Tfh1/Tfh17 scores of the highly
abundant clones and use Prism for plotting
```

Code for Figure 4.2, 4.3, 4.4

```
#Open seurat and load the samples
library(Seurat)
data_dir <-
"~/Desktop/Sophie_GEX_VDJ_Feature_1/outs/count/filtered_feature_bc_matr
ix/"
list.files(data_dir)
expression_matrix1<- Read10X(data.dir = data_dir)
sample1 = CreateSeuratObject(counts = expression_matrix1$`Gene
Expression`,project = "sample1project")
sample1[['Protein']] = CreateAssayObject(counts =
expression_matrix1$`Antibody Capture`)
#will need to use hashtag to filter later Refer to seurat HTO tutorial
#tradeSeq
data_dir <-
"~/Desktop/Sophie_GEX_VDJ_Feature_2/outs/count/filtered_feature_bc_matr
ix/"
list.files(data_dir)
expression_matrix2<- Read10X(data.dir = data_dir)
sample2 = CreateSeuratObject(counts = expression_matrix2$`Gene
Expression`,project = "sample2project")
sample2[['Protein']] = CreateAssayObject(counts =
expression_matrix2$`Antibody Capture`)
rm(expression_matrix1, expression_matrix2)
#Perform QC to filter dead and not singlets
library("cowplot")
library("dplyr")
library("ggplot2")
sample1[["percent.mt"]] <- PercentageFeatureSet(sample1, pattern =
"^MT-")
sample2[["percent.mt"]] <- PercentageFeatureSet(sample2, pattern =
"^MT-")
#Normalize 2 samples and merge the dataset
sample1 <- NormalizeData(sample1, assay = "Protein",
normalization.method = "CLR")
sample1 <- ScaleData(sample1, assay = "Protein")
sample2 <- NormalizeData(sample2, assay = "Protein",
normalization.method = "CLR")
sample2 <- ScaleData(sample2, assay = "Protein")
DefaultAssay(sample1) <- "RNA"
DefaultAssay(sample2) <- "RNA"
sample1 <- SCTransform(sample1, method = "glmGamPoi",verbose = FALSE)
sample2 <- SCTransform(sample2, method = "glmGamPoi", verbose = FALSE)
sample.list<-c(sample1,sample2)
features <- SelectIntegrationFeatures(object.list = sample.list,
nfeatures = 3000)
sample.list <- PrepSCTIntegration(object.list = sample.list,
anchor.features = features)
immune.anchors <- FindIntegrationAnchors(object.list = sample.list,
normalization.method = "SCT",
                                         anchor.features = features)
sample.combined <- IntegrateData(anchorset = immune.anchors,
normalization.method = "SCT")
sample.combinedbackup<-sample.combined
#sample.combined<-sample.combinedbackup
rm(sample.anchors,sample1,sample2,immune.anchors,sample.list)
```

```
#Remove some annoying genes
all.genes<-rownames(sample.combined@assays[["RNA"]])
rps <- all.genes[grep('^RPS|^RPL|MRPS', all.genes)]
igh <- all.genes[grep('^IGHV', all.genes)]
igl <- all.genes[grep('^IGL|^IGK', all.genes)]
s.genes <- cc.genes$s.genes
g2m.genes <- cc.genes$g2m.genes
var.genes.no.riboigg <-
setdiff(sample.combined@assays[["integrated"]]@var.features, s.genes)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, g2m.genes)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, igh)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, igl)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, rps)
#Downstream analysis-remove nonB cell and dead and doublets
sample.combined <- RunPCA(sample.combined, verbose = FALSE,features
=var.genes.no.riboigg)
ElbowPlot(sample.combined, ndims =50)
DimHeatmap(sample.combined, dims = 1:15, cells = 500, balanced = TRUE,
ncol=5)
DimHeatmap(sample.combined, dims = 16:30, cells = 500, balanced = TRUE,
ncol=5)
sample.combined <- FindNeighbors(sample.combined, dims = 1:20,
reduction="pca")
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:20)
sample.combined <- FindClusters(sample.combined, resolution = 0.3,
reduction.type="pca")
DimPlot(sample.combined,label = T)
VlnPlot(sample.combined, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3,pt.size = 0)
DimPlot(sample.combined,split.by = "orig.ident")
FeaturePlot(sample.combined,"CD3E")
sample.markers <- FindAllMarkers(sample.combined, only.pos = TRUE)
top10markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 10, wt
= avg_logFC)
sample.combined.small<- subset(sample.combined, downsample = 100)
DoHeatmap(sample.combined.small, features = unique(top10markers$gene),
angle = 90) + NoLegend() + theme(text = element_text(size = 7.5))
sample.combined <- subset(x = sample.combined,
idents=c("7","9","10"),invert = TRUE) #9 T cell;7,10 Dead cells
VlnPlot(sample.combined, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
sample.combined <- subset(sample.combined, subset = nFeature_RNA > 500
& nFeature_RNA < 6000 &nCount_RNA<35000 & percent.mt < 10)
#Demultiplex
DefaultAssay(sample.combined) <- "Protein"
sample.combined <- ScaleData(sample.combined, assay = "Protein")
sample.combined <- RunPCA(sample.combined, features = c("Hashtag1 -
TotalSeqC","Hashtag2 -TotalSeqC","Hashtag3 -TotalSeqC") ,
reduction.name = "pca_protein", reduction.key = "pca_protein_", verbose
= FALSE)
sample.combined <- FindNeighbors(sample.combined, features =
c("Hashtag1 -TotalSeqC","Hashtag2 -TotalSeqC","Hashtag3 -TotalSeqC"),
dims = NULL)
sample.combined <- FindClusters(sample.combined, resolution = 0.03)
DimPlot(sample.combined, reduction = "pca_protein")
sample.combined <- subset(x = sample.combined, idents="3",invert =
TRUE)
```

```
DefaultAssay(sample.combined) <- "integrated"
#run PCA with the new gene list to further grouping
sample.combined <- RunPCA(object = sample.combined, features =
var.genes.no.riboigg,verbose = FALSE)
ElbowPlot(sample.combined, ndims =50)
DimHeatmap(sample.combined, dims = 1:15, cells = 500, balanced = TRUE,
ncol=5)
DimHeatmap(sample.combined, dims = 16:30, cells = 500, balanced = TRUE,
ncol=5)
sample.combined <- FindNeighbors(sample.combined, dims =
1:18,reduction="pca")
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:30)
sample.combined <- FindClusters(sample.combined, resolution = 0.9,
reduction.type="pca")
DimPlot(sample.combined,pt.size = 0.3,label=T)
DimPlot(sample.combined,pt.size = 0.3,label=F)
DimPlot(sample.combined,split.by = "orig.ident")
#plot <- DimPlot(sample.atyBfocus, reduction = "umap")
#select.cells <- CellSelector(plot = plot)
#Idents(sample.atyBfocus, cells = select.cells) <- "NewCells"
#intermakers<-FindMarkers(sample.atyBfocus, ident.1 = "NewCells")
VlnPlot(sample.combined, features = c("nFeature_Protein",
"nCount_Protein"), ncol = 3)
sample.combined <- subset(sample.combined, subset =
nCount_Protein<15000)
#Build tree
sample.combined<-BuildClusterTree(sample.combined,
features=var.genes.no.riboigg,dim=T)
PlotClusterTree(sample.combined)
levels(sample.combined)
#Reset the names to make heatmap more reasonable
#my_names <- c(4,8,9,0,3,5,1,7,2,6)
#names(my_names)<-levels(sample.combined)
#sample.combined<-RenameIdents(sample.combined,my_names)
#my_newlevels<-c(1,8,6,0,12,15,5,10,11,2,3,4,7,9,13,14)
sample.combined<-
RenameIdents(sample.combined,"0"="3","1"="0","2"="9","3"="10","4"="11",
"5"="8","6"="2","7"="12","8"="1","9"="13","10"="6","11"="7","12"="4","1
3"="14","14"="15","15"="5")
my_newlevels<-c(0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15)
sample.combined@active.ident<-
factor(x=sample.combined@active.ident,levels=my_newlevels)
sample.combined<-AddMetaData(object=sample.combined,metadata
=sample.combined@active.ident ,col.name = "Newcluster")
sample.combined<-
RenameIdents(sample.combined,"0"="Naive","1"="IgDmem","2"="Activatedmem
","3"="Classicalmem","4"="Atypicalmem","5"="PreGCmem","6"="PB/PC","7"="
PrememGC","8"="LZGC","9"="GC1","10"="GC2","11"="DZGC","12"="ProGC","13"
="ProGC","14"="ProGC","15"="ProGC")
#sample.combined<-
RenameIdents(sample.combined,"0"="Naive","1"="IgDmem","2"="Activatedmem
","3"="Classicalmem","4"="Atypicalmem","5"="PreGCmem","6"="PB/PC","7"="
PrememGC","8"="LZGC","9"="GC1","10"="GC2","11"="DZGC","12"="ProGC1","13
"="ProGC2","14"="ProGC3","15"="ProGC4")
sample.markers <- FindAllMarkers(sample.combined,assay = "SCT",only.pos
= T,min.pct = 0.05)
```

```
top3markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 3, wt =
avg_logFC)
top5markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 5, wt =
avg_logFC)
sample.combined.small<- subset(sample.combined, downsample = 50)
DoHeatmap(sample.combined.small, features = unique(top5markers$gene),
angle = 90) + NoLegend() + theme(text = element_text(size = 8))
DotPlot(sample.combined,features = unique(top3markers$gene),assay =
"SCT",scale.by = "radius",dot.scale = 3)+RotatedAxis()+coord_flip()#+
scale_size()+scale_radius(limits = c(min.pct=0, max.pct=100))
plot<-DotPlot(sample.combined,features = unique(top3markers$gene),assay
= "SCT",scale.by = "radius",dot.scale = 3)
atypicalvsclassical<-FindMarkers(sample.combined,ident.1 =
"Atypicalmem",ident.2 = "Classicalmem",min.pct = 0,assay =
"SCT",min.diff.pct = 0,logfc.threshold = 0)
write.csv(atypicalvsclassical,file="atypicalvsclassicaltonsilmoregenes.
csv")
classicalvsatypical<-FindMarkers(sample.combined,ident.1 =
"Classicalmem",ident.2 = "Atypicalmem",min.pct = 0,assay =
"SCT",min.diff.pct = 0,logfc.threshold = 0)
#Perform correlation test on TFs
Atypicalsubset<- subset(x = sample.combined, idents="Atypicalmem")
Atypicalmemgeneset_RNA<-Atypicalsubset@meta.data[["Atypicalscore1"]]
CD11c_protein<-FetchData(Atypicalsubset,vars="CD11c-TotalSeqC")
#FCRL4_protein<-FetchData(Atypicalsubset,vars="FCRL4-TotalSeqC")
#CD19_protein<-FetchData(Atypicalsubset,vars="CD19-TotalSeqC")
#FCRL5_protein<-FetchData(Atypicalsubset,vars="FCRL5-TotalSeqC")
Classicalmemgeneset_RNA<-Atypicalsubset@meta.data[["Classicalscore1"]]
#CD21_protein<-FetchData(Atypicalsubset,vars="CD21-TotalSeqC")
CXCR5_protein<-FetchData(Atypicalsubset,vars="CXCR5-TotalSeqC")
CD27_protein<-FetchData(Atypicalsubset,vars="CD27-TotalSeqC")
ZEB2_RNA<-FetchData(Atypicalsubset,vars="ZEB2")
SOX5_RNA<-FetchData(Atypicalsubset,vars="SOX5")
ZBED2_RNA<-FetchData(Atypicalsubset,vars="ZBED2")
BHLHE40_RNA<-FetchData(Atypicalsubset,vars="BHLHE40")
ZBTB32_RNA<-FetchData(Atypicalsubset,vars="ZBTB32")
###Correlationdataset<-
cbind(Atypicalmemgeneset_RNA,CD11c_protein,FCRL4_protein,CD19_protein,F
CRL5_protein,Classicalmemgeneset_RNA,CD21_protein,CXCR5_protein,CD27_pr
otein,ZEB2_RNA,SOX5_RNA,ZBED2_RNA,BHLHE40_RNA,ZBTB32_RNA)
Correlationdataset<-
cbind(Atypicalmemgeneset_RNA,CD11c_protein,Classicalmemgeneset_RNA,CD27
_protein,ZEB2_RNA,BHLHE40_RNA,SOX5_RNA,ZBED2_RNA,ZBTB32_RNA)
library("Hmisc")
library("corrplot")
correlationassay<-rcorr(as.matrix(Correlationdataset),type = "pearson")
M<-correlationassay$r
M<--M
M<-M[1:4,5:9]
M<-t(M)
p_mat<-correlationassay$P
p_mat<-p_mat[1:4,5:9]
p_mat<-t(p_mat)
corrplot(M,p.mat = p_mat,sig.level = 0.05,method = "color")
#Violin plot for top5 TFs
library(patchwork)
modify_vlnplot<- function(obj,
                          feature,
```

```
                           pt.size = 0,
                           plot.margin = unit(c(-0.75, 0, -0.75, 0),
"cm"),
                           ...) {
        p<- VlnPlot(obj, features = feature, pt.size = pt.size, ... )
+
                xlab("") + ylab(feature) + ggtitle("") +
                theme(legend.position = "none",
                      axis.text.x = element_blank(),
                      axis.ticks.x = element_blank(),
                      axis.title.y = element_text(size = rel(1), angle
= 0),
                      axis.text.y = element_text(size = rel(1)),
                      plot.margin = plot.margin )
        return(p)
}
## extract the max value of the y axis
extract_max<- function(p){
        ymax<-
max(ggplot_build(p)$layout$panel_scales_y[[1]]$range$range)
        return(ceiling(ymax))
}
StackedVlnPlot<- function(obj, features,
                           pt.size = 0,
                           plot.margin = unit(c(-0.75, 0, -0.75, 0),
"cm"),
                           ...) {

        plot_list<- purrr::map(features, function(x) modify_vlnplot(obj
= obj,feature = x, ...))

        # Add back x-axis title to bottom plot. patchwork is going to
support this?
        plot_list[[length(plot_list)]]<- plot_list[[length(plot_list)]]
+
                theme(axis.text.x=element_text(), axis.ticks.x =
element_line())

        # change the y-axis tick to only max value
        ymaxs<- purrr::map_dbl(plot_list, extract_max)
        plot_list<- purrr::map2(plot_list, ymaxs, function(x,y) x +
                                        scale_y_continuous(breaks =
c(y)) +
                                        expand_limits(y = y))

        p<- patchwork::wrap_plots(plotlist = plot_list, ncol = 1)
        return(p)
}
StackedVlnPlot(sample.combined,features =
c("ZEB2","BHLHE40","SOX5","ZBED2","ZBTB32"))
#sample.markersciteseq <- FindAllMarkers(sample.combined,assay =
"Protein",features =
sample.combined@assays[["Protein"]]@counts@Dimnames[[1]][1:14])
#top3markers<-sample.markersciteseq %>% group_by(cluster) %>% top_n(n =
3, wt = avg_logFC)
#top5markers<-sample.markersciteseq %>% group_by(cluster) %>% top_n(n =
5, wt = avg_logFC)
#sample.combined.small<- subset(sample.combined, downsample = 50)
```

```
#DoHeatmap(sample.combined.small, features = unique(top3markers$gene),
angle = 90,assay = "Protein") + NoLegend() + theme(text =
element_text(size = 8))
#DotPlot(sample.combined,features = unique(top3markers$gene),assay =
"Protein",dot.scale = 15,scale.by = "radius",scale = F,col.max =
1.5)+RotatedAxis()+coord_flip()+ scale_size()+scale_radius(limits =
c(min.pct=0, max.pct=100))
#Show summary heatmap
avg<-AverageExpression(sample.combined,assays = "Protein",features =
sample.combined@assays[["Protein"]]@counts@Dimnames[[1]][1:14],return.s
eurat = T )
avg<-AverageExpression(sample.combined,assays = "Protein",features =
sample.combined@assays[["Protein"]]@counts@Dimnames[[1]][15:17],return.
seurat = T )
library(dittoSeq)
dittoHeatmap(avg,show_colnames = T,scale = "none",cluster_rows=F)
dittoHeatmap(avg,show_colnames = T)
FeaturePlot(sample.combined,
features=c("sct_CR2","sct_CD27","sct_IGHD","sct_IGHM","sct_CXCR4","sct_
CD83","sct_ZEB2","sct_ITGAX","sct_SDC1","sct_CCR6","sct_CD38","sct_CD69
"),ncol=4,order=T,pt.size=0.01,min.cutoff = 0)
FeaturePlot(sample.combined, features=c("sct_CD38"))
#Do some genesets
library(readr)
LZgeneset<- read_csv("Downloads/GSE38712.top.table.tsv")
sample.combined <- AddModuleScore(object = sample.combined,features =
LZgeneset,name = 'LZscore',assay="SCT")
FeaturePlot(sample.combined,features="LZscore1",pt.size =
0.1,min.cutoff = 0.1,order=T)
FeaturePlot(sample.combined,features = "CD83",min.cutoff = 1)
DZgeneset<- read_csv("Downloads/GSE38712.top.table-2.tsv")
sample.combined <- AddModuleScore(object = sample.combined,features =
DZgeneset,name = 'DZscore',assay="SCT")
FeaturePlot(sample.combined,features="DZscore1",pt.size =
0.1,min.cutoff = 0,max.cutoff = 0.2,order = T)
FeaturePlot(sample.combined,features = "CXCR4",min.cutoff = 1)
library(readr)
Atypicalgeneset <- read_csv("Downloads/GSE64493.top.table-
atypical.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Atypicalgeneset,name = 'Atypicalscore',assay="SCT")
FeaturePlot(sample.combined,features="Atypicalscore1",pt.size =
0.1,min.cutoff = 0,max.cutoff = 0.08,order=T)
FeaturePlot(sample.combined,features = "ITGAX",min.cutoff = 0)
classicalBgeneset <- read_csv("Downloads/GSE64493.top.classical.tsv")
sample.combined <- AddModuleScore(object = sample.combined,features =
classicalBgeneset,name = 'Classicalscore',assay="SCT")
FeaturePlot(sample.combined,features="Classicalscore1",pt.size =
0.1,min.cutoff = 0.05,order = T)
FeaturePlot(sample.combined,features = "CD27",min.cutoff = 0)
FeaturePlot(sample.combined,features = "IGHD",min.cutoff = 0,order=T)
library(readr)
Activated_B_cell_gene <- read_csv("Downloads/b act-2.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Activated_B_cell_gene,name = 'Activatedscore',assay="SCT")
FeaturePlot(sample.combined,features="Activatedscore1",pt.size =
0.1,min.cutoff = 0.01,max.cutoff = 0.165)
FeaturePlot(sample.combined,features = "CD69",min.cutoff = 1)
```

```
Becans_mouseatyBtop200gene <- read_csv("Downloads/becandata.csv")
sample.combined <- AddModuleScore(object = sample.combined,features =
Becans_mouseatyBtop200gene,name = 'mouseatyB',assay="SCT")
FeaturePlot(sample.combined,features="mouseatyB1",pt.size =
0.1,min.cutoff = 0)
#Calculate TGFb
TGFBsignalingfeatures<-list(c(
                              "ACVR1",
                              "APC",
                              "ARID4B",
                              "BCAR3",
                              "BMP2",
                              "BMPR1A",
                              "BMPR2",
                              "CDH1",
                              "CDK9",
                              "CDKN1C",
                              "CTNNB1",
                              "ENG",
                              "FKBP1A",
                              "FNTA",
                              "FURIN",
                              "HDAC1",
                              "HIPK2",
                              "ID1",
                              "ID2",
                              "ID3",
                              "IFNGR2",
                              "JUNB",
                              "KLF10",
                              "LEFTY2",
                              "LTBP2",
                              "MAP3K7",
                              "NCOR2",
                              "NOG",
                              "PMEPA1",
                              "PPM1A",
                              "PPP1CA",
                              "PPP1R15A",
                              "RAB31",
                              "RHOA",
                              "SERPINE1",
                              "SKI",
                              "SKIL",
                              "SLC20A1",
                              "SMAD1",
                              "SMAD3",
                              "SMAD6",
                              "SMAD7",
                              "SMURF1",
                              "SMURF2",
                              "SPTBN1",
                              "TGFB1",
                              "TGFBR1",
                              "TGIF1",
                              "THBS1",
                              "TJP1",
                              "TRIM33",
```

```
                                "UBE2D3",
                                "WWTR1",
                                "XIAP"))
Stat3pathway<-
list(c("JAK1","JAK2","JAK3","MAPK1","MAPK3","MTOR","STAT3","TYK2"))
sample.combined <- AddModuleScore(object = sample.combined,features =
Stat3pathway,name = 'BIOCARTA_STAT3_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="BIOCARTA_STAT3_SIGNALING_PATHWAY1
",min.cutoff = 0, pt.size = 0.2,order=TRUE)
my_comparisons1 <- list( c("Atypicalmem",
"Classicalmem"),c("Atypicalmem", "PreGCmem"),c("Atypicalmem",
"Activatedmem"),c("Atypicalmem", "IgDmem"),c("Atypicalmem", "Naive"))
VlnPlot(sample.combined,"BIOCARTA_STAT3_SIGNALING_PATHWAY1",pt.size=0.1
,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=18)+ stat_compare_means(label =
"p.signif",ref.group = "Atypicalmem")+scale_y_continuous()+
theme(legend.position = 'none')
#Greyzonefeatures<-
list(c("CENPM","BUB1B","TUBA1C","SKA1","CCNB1","SPC24","CDCA5","NSL1","
CDK1","AURKB","PLK1","BIRC5","CDCA8","CENPP","UBE2C","SPDL1","TUBB4B","
CCNE1","RAD21","FBXO5","HMMR","BUB3","ATP5D","TOP2A","POU2AF1"))
#sample.combined <- AddModuleScore(object = sample.combined,features =
Greyzonefeatures,name = 'GZscore',assay="SCT")
#FeaturePlot(sample.combined,features="GZscore1",pt.size =
0.1,order=TRUE)
library(ggpubr)
sample.combined <- AddModuleScore(object = sample.combined,features =
TGFBsignalingfeatures,name = 'HALLMARK_TGF_BETA_SIGNALING',assay="SCT")
FeaturePlot(sample.combined,features="HALLMARK_TGF_BETA_SIGNALING1",min
.cutoff = 0, pt.size = 0.2,order=TRUE)
my_comparisons <- list( c("Atypicalmem", "Classicalmem"))
VlnPlot(sample.combined,"HALLMARK_TGF_BETA_SIGNALING1",pt.size=0.1,adju
st=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=18)+ stat_compare_means(label =
"p.signif",ref.group = "Atypicalmem")+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-2.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_1_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_1_MEDIATED_SIGNALI
NG_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_1_MEDIATED_SIGNALING_PATHWAY1",
pt.size=1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-4.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_21_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_21_MEDIATED_SIGNAL
ING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_21_MEDIATED_SIGNALING_PATHWAY1"
,pt.size=0.1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=18)+ stat_compare_means(label =
"p.signif",ref.group = "Atypicalmem")+scale_y_continuous()+
theme(legend.position = 'none')
```

```
Geneset<- read_csv("Downloads/cytokines geneset/geneset-5.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_23_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_23_MEDIATED_SIGNAL
ING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_23_MEDIATED_SIGNALING_PATHWAY1"
,pt.size=1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-6.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_27_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_27_MEDIATED_SIGNAL
ING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_27_MEDIATED_SIGNALING_PATHWAY1"
,pt.size=1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-7.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_35_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_35_MEDIATED_SIGNAL
ING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset-8.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_4_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_4_MEDIATED_SIGNALI
NG_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_4_MEDIATED_SIGNALING_PATHWAY1",
pt.size=1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-9.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_6_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_6_MEDIATED_SIGNALI
NG_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_INTERLEUKIN_6_MEDIATED_SIGNALING_PATHWAY1",
pt.size=1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-10.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_7_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_7_MEDIATED_SIGNALI
NG_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset-12.txt")
```

```
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_12',assay="SCT")
FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_121",p
t.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset-13.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_15',assay="SCT")
FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_151",p
t.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GO_RESPONSE_TO_INTERLEUKIN_151",pt.size=1,adju
st=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=20)+
stat_compare_means(comparisons=my_comparisons)+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/geneset-14.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_17',assay="SCT")
FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_171",p
t.size = 0.1,min.cutoff = 0)
#Geneset<- read_csv("Downloads/cytokines geneset/geneset-15.txt")
#sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_18',assay="SCT")
#FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_181",
pt.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset-16.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_2',assay="SCT")
FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_21",pt
.size = 0.1,min.cutoff = 0)
#Geneset<- read_csv("Downloads/cytokines geneset/geneset-17.txt")
#sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_3',assay="SCT")
#FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_31",p
t.size = 0.1,min.cutoff = 0)
#Geneset<- read_csv("Downloads/cytokines geneset/geneset-18.txt")
#sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_4',assay="SCT")
#FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_41",p
t.size = 0.1,min.cutoff = 0)
#Geneset<- read_csv("Downloads/cytokines geneset/geneset-19.txt")
#sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_6',assay="SCT")
#FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_61",p
t.size = 0.1,min.cutoff = 0)
#Geneset<- read_csv("Downloads/cytokines geneset/geneset-20.txt")
#sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_7',assay="SCT")
#FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_71",p
t.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset-21.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GO_RESPONSE_TO_INTERLEUKIN_9',assay="SCT")
FeaturePlot(sample.combined,features="GO_RESPONSE_TO_INTERLEUKIN_91",pt
.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines geneset/geneset.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GO_INTERLEUKIN_18_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
```

```
FeaturePlot(sample.combined,features="GO_INTERLEUKIN_18_MEDIATED_SIGNAL
ING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
Geneset<- read_csv("Downloads/cytokines
geneset/GOBP_INTERFERON_GAMMA_MEDIATED_SIGNALING_PATHWAY.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name =
'GOBP_INTERFERON_GAMMA_MEDIATED_SIGNALING_PATHWAY',assay="SCT")
FeaturePlot(sample.combined,features="GOBP_INTERFERON_GAMMA_MEDIATED_SI
GNALING_PATHWAY1",pt.size = 0.1,min.cutoff = 0)
VlnPlot(sample.combined,"GOBP_INTERFERON_GAMMA_MEDIATED_SIGNALING_PATHW
AY1",pt.size=0.1,adjust=1.2)+stat_summary(fun.y=median,geom="point",
shape=95,color="white",size=18)+ stat_compare_means(label =
"p.signif",ref.group = "Atypicalmem")+scale_y_continuous()+
theme(legend.position = 'none')
Geneset<- read_csv("Downloads/cytokines geneset/GOMF_SMAD_BINDING.txt")
sample.combined <- AddModuleScore(object = sample.combined,features =
Geneset,name = 'GOMF_SMAD_BINDING',assay="SCT")
FeaturePlot(sample.combined,features="GOMF_SMAD_BINDING1",pt.size =
0.1,min.cutoff = 0)
cytokinepathway<-sample.combined@meta.data[14:30]
cytokinepathway<-as.matrix(cytokinepathway)
Clusternames<-as.matrix(sample.combined@active.ident)
cytokine_merge<- cbind(cytokinepathway, Clusternames)
cytokine_merge<-as.data.frame(cytokine_merge)
Check<-cytokine_merge["V18"]
cytokine_merge[,18]<-NULL
cytokine_merge<-t(cytokine_merge)
Pathwayobject<-CreateSeuratObject(counts = cytokine_merge )
Pathwayobject<-AddMetaData(Pathwayobject,metadata
=sample.combined@active.ident ,col.name = "Pathwayexpression" )
Pathwayobject@active.ident<-sample.combined@active.ident
library(dittoSeq)
avg<-AverageExpression(Pathwayobject,assays = "RNA",features =
list(Pathwayobject@assays[["RNA"]]@counts@Dimnames[[1]][1:9]),return.se
urat = T )
avg<-AddMetaData(avg,metadata
=c("Naive","IgDmem","Activatedmem","Classicalmem","Atypicalmem","PreGCm
em","PB/PC","PrememGC","LZGC","GC1","GC2","DZGC","ProGC1","ProGC2","Pro
GC3","ProGC4") ,col.name = "Subset" )
dittoHeatmap(avg,show_colnames = T)
#Export the values of pathway analysis
pathwayassayresult<-avg@assays[["RNA"]]
pathwayassayresult <-as.data.frame(pathwayassayresult)
write.csv(avg@assays[["RNA"]]@scale.data, file = "pathwayscores.csv")
#Pathwayobjectmemory<-subset(x = Pathwayobject,
idents=c("Naive","IgDmem","Activatedmem","Classicalmem","Atypicalmem","
PreGCmem"),invert = F)
#avg<-AverageExpression(Pathwayobjectmemory,assays = "RNA",features =
list(Pathwayobject@assays[["RNA"]]@counts@Dimnames[[1]][1:9]),return.se
urat = T )
#dittoHeatmap(avg,show_colnames = T,scale = "row")
#my_oldlevels<-c(0,1,2,3,4,5,6,7,8,9)
#sample.combined@active.ident<-
factor(x=sample.combined@active.ident,levels=my_oldlevels)
#newclusters<-sample.combined@active.ident
#sample.combined<-
AddMetaData(sample.combined,metadata=newclusters,col.name="newclusters"
)
```

```
#Redo the heatmap
sample.markers <- FindAllMarkers(sample.combined,only.pos = T,assay =
"SCT",min.pct = 0.05)
top10markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 10, wt
= avg_logFC)
sample.combined.small<- subset(sample.combined, downsample = 100)
DoHeatmap(sample.combined.small, features = unique(top10markers$gene),
angle = 90) + NoLegend() + theme(text = element_text(size =7))
#Export antibody or Clustering based on antibody detection
Abfeatures<-sample.combined@assays[["Protein"]]@data #no@data means raw
count
write.csv(Abfeatures,file="Abdetection.csv")
write.csv(sample.combined@active.ident,file="clusterinf.csv")
reductionscore<-sample.combined@reductions[["umap"]]
write.csv(reductionscore@cell.embeddings,file="umap.csv")
mtratio<-sample.combined@meta.data[["percent.mt"]]
write.csv(mtratio,file="mtratio.csv")
#proteincount<-sample.combined@meta.data[["nCount_Protein"]]
#write.csv(proteincount,file="proteincount.csv")
#Idents(sample.combined) <-
sample.combined@meta.data[["seurat_clusters"]]
#DefaultAssay(sample.combined,"Protein")
saveRDS(sample.combined, file = file.path("~/Desktop/humantonsil.RDS"))
#Do monocle trajectory
library(monocle3)
library(Seurat)
library(SeuratData)
library(SeuratWrappers)
library(ggplot2)
library(patchwork)
library(magrittr)
cds <- as.cell_data_set(sample.combined)
cds <- cluster_cells(cds)
#replace monocle dimensions with seurat
#cds@reducedDims$TSNE <-  suerat_object@reductions$tsne@cell.embeddings
#cds@reducedDims$UMAP <-  suerat_object@reductions$umap@cell.embeddings
plot_cells(cds,show_trajectory_graph = T)
cds <-
learn_graph(cds,use_partition=FALSE,close_loop=TRUE,learn_graph_control
= list(minimal_branch_len=6,geodesic_distance_ratio=0.3,rann.k=50))
plot_cells(cds,show_trajectory_graph = T,color_cells_by =
"ident",label_roots = F,label_branch_points=FALSE,
label_leaves=FALSE,label_cell_groups = F,
           label_groups_by_cluster = TRUE)
cds<-order_cells(cds)
plot_cells(cds, color_cells_by = "pseudotime",show_trajectory_graph =
TRUE, label_branch_points=FALSE,
label_leaves=FALSE,trajectory_graph_color="blue")
sample.combined<-AddMetaData(object=sample.combined,metadata =
cds@principal_graph_aux@listData$UMAP$pseudotime,col.name = "PSDtime")
#select trajectory for atypical b cell
#cds_sub<-choose_graph_segments(cds, clear_cds = FALSE)
#plot_cells(cds_sub, color_cells_by =
"pseudotime",show_trajectory_graph = FALSE, label_branch_points=FALSE,
label_leaves=FALSE,cell_size = 1)
#Specifically check atypical B cell generation
plot <- DimPlot(sample.combined, reduction = "umap")
select.cells <- CellSelector(plot = plot)
```

```
Idents(sample.combined, cells = select.cells) <- "NewCells"
sample.combinedatyBfocus1<-
subset(x=sample.combined,idents=c("NewCells"))
sample.combinedatyBfocus <- subset(x = sample.combinedatyBfocus1,
subset=PSDtime>21&PSDtime<30) #adjust to fit in both population
FeaturePlot(sample.combinedatyBfocus,"PSDtime")
FeaturePlot(sample.combinedatyBfocus,"sct_ITGAX")
plotdatac <- FetchData(object = sample.combinedatyBfocus, vars =
c("TGFB_signaling1","IL21_signaling1","sct_ZEB2","PSDtime","Atypicalsco
re1","Classicalscore1","sct_ITGAX","sct_STAT1","sct_STAT2","sct_STAT3",
"sct_STAT4","sct_STAT5A","sct_STAT5B","sct_STAT6","sct_CD38","sct_TGFB1
","SMADbinding1","sct_CCR6","sct_CD27","Interferong1","Allmutation"))
require(methods)
p <- ggplot(plotdatac, aes(x = PSDtime, y = TGFB_signaling1))
p + geom_point(colour="grey",size=1)+geom_smooth( span=0.9,method =
'loess',se=F)+xlim(22.5,30) #method gam+scale_y_sqrt()
p + geom_point(colour="grey",size=1)+geom_smooth( span=0.7,method =
'gam',se=F)+xlim(22.5,30) #method gam
p  +geom_smooth(span = 0.5, method = 'gam')+xlim(22.5,30)
p + geom_smooth( span=1.0,method = 'loess')+xlim(22.5,30) #method gam
## plot base + points
#some funky code here to plot mutation in cluster 14
library(ggpubr)
p<- ggplot(plotdatac, aes(x = Atypicalscore1, y = Allmutation))
p + geom_point(colour="grey",size=1)+geom_smooth( method =
'lm')+stat_cor(label.y = 20)
head(plotdatac)
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_ITGAX))
p <- ggplot(plotdatac, aes(x = PSDtime, y = Atypicalscore1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = TGFB_signaling1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_TGFB1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = IL21_signaling1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_ZEB2))
p <- ggplot(plotdatac, aes(x = PSDtime, y = Classicalscore1))
p <- ggplot(plotdatac, aes(x = PSDtime, y =sct_STAT1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_STAT2))
p <- ggplot(plotdatac, aes(x = PSDtime, y=sct_STAT3))
p <- ggplot(plotdatac, aes(x = PSDtime, y =sct_STAT4))
p <- ggplot(plotdatac, aes(x = PSDtime, y =sct_STAT5A))
p <- ggplot(plotdatac, aes(x = PSDtime, y =sct_STAT5B))
p <- ggplot(plotdatac, aes(x = PSDtime, y =sct_STAT6))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_CD27))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_CCR6))
p <- ggplot(plotdatac, aes(x = PSDtime, y = sct_CD38))
p <- ggplot(plotdatac, aes(x = PSDtime, y = SMADbinding1))
p <- ggplot(plotdatac, aes(x = PSDtime, y = Interferong1))
#p <- ggplot(plotdatac, aes(x = monoclepseduotime, y =
rna_CR2))+geom_point(colour="grey",size=1)
#p + geom_point(colour="grey",size=1) +geom_smooth(span = 0.3, method =
'loess')+xlim(15, 45)
#Try combined some
p <- ggplot()+geom_smooth(data=plotdatac,aes(x = PSD1opt, y =
Atypicalscore1),span = 0.1, method = 'loess',colour="blue")+xlim(5,20)
p<-p+geom_smooth(data=plotdatac,aes(x = PSD1opt, y =
TGFB_signaling1),span = 0.1, method = 'loess',colour="red")+xlim(5,20)
p
p <- ggplot()+geom_smooth(data=plotdatac,aes(x = PSD1opt, y =
ZEB2),span = 0.1, method = 'loess',colour="blue")+xlim(5,20)
```

```
p<-p+geom_smooth(data=plotdatac,aes(x = PSD1opt, y = ITGAX),span = 0.1,
method = 'loess',colour="red")+xlim(5,20)
p
p <- ggplot()+geom_smooth(data=plotdatac,aes(x = PSD1opt, y =
ZEB2),span = 0.1, method = 'loess',colour="blue")+xlim(5,20)
p<-p+geom_smooth(data=plotdatac,aes(x = PSD1opt, y = rna_STAT3),span =
0.1, method = 'loess',colour="red")+xlim(5,20)
p
#Add result back to seurat
sample.combined <- AddMetaData(object = sample.combined,metadata =
cds@principal_graph_aux@listData$UMAP$pseudotime,col.name =
"monoclepseduotime")
#Use slingshot to do trajectory again
library(slingshot)
library(BUSpaRse)
library(tidyverse)
library(tidymodels)
library(Seurat)
library(scales)
library(viridis)
library(Matrix)
#https://bustools.github.io/BUS_notebooks_R/slingshot.html
sds <- slingshot(Embeddings(sample.combined, "umap"), clusterLabels =
sample.combined$Newcluster,
                 start.clus = 0, stretch = 0,approx_points=100)
cell_pal <- function(cell_vars, pal_fun,...) {
        if (is.numeric(cell_vars)) {
                pal <- pal_fun(100, ...)
                return(pal[cut(cell_vars, breaks = 100)])
        } else {
                categories <- sort(unique(cell_vars))
                pal <- setNames(pal_fun(length(categories), ...),
categories)
                return(pal[cell_vars])
        }
}
cell_colors_clust <- cell_pal(sample.combined$Newcluster, hue_pal())
plot(reducedDim(sds), col = cell_colors_clust, pch = 16, cex = 0.3)
lines(sds, lwd = 1,col = 'black',type = 'lineages')
lines(sds, lwd = 2,col = 'black',linInd=2,type = 'lineages')
lines(sds, lwd = 2,col = 'black',linInd=2)
nc <- 3
pt <- slingPseudotime(sds)
nms <- colnames(pt)
nr <- ceiling(length(nms)/nc)
pal <- viridis(100, end = 0.95)
for (i in nms) {
        colors <- pal[cut(pt[,i], breaks = 100)]
        plot(reducedDim(sds), col = colors, pch = 16, cex = 0.5, main =
i)
        lines(sds, lwd = 2, col = 'black')
}
par(mfrow = c(1, 1))
colors <- pal[cut(pt[,1], breaks = 100)]
plot(reducedDim(sds), col = colors, pch = 16, cex = 0.5, main = 1)
lines(sds, lwd = 2, col = 'black',linInd=1)
colors <- pal[cut(pt[,2], breaks = 100)]
plot(reducedDim(sds), col = colors, pch = 16, cex = 0.5, main = 2)
```

```
lines(sds, lwd = 2, col = 'black',linInd=2)
colors <- pal[cut(pt[,3], breaks = 100)]
plot(reducedDim(sds), col = colors, pch = 16, cex = 0.5, main = 3)
lines(sds, lwd = 2, col = 'black',linInd=3)
#Do again for changing the origin
sds <- slingshot(Embeddings(sample.combined, "umap"), clusterLabels =
sample.combined$seurat_clusters,
                 start.clus = 6, stretch = 0,approx_points=100)
cell_pal <- function(cell_vars, pal_fun,...) {
        if (is.numeric(cell_vars)) {
                pal <- pal_fun(100, ...)
                return(pal[cut(cell_vars, breaks = 100)])
        } else {
                categories <- sort(unique(cell_vars))
                pal <- setNames(pal_fun(length(categories), ...),
categories)
                return(pal[cell_vars])
        }
}
cell_colors_clust <- cell_pal(sample.combined$seurat_clusters,
hue_pal())
nc <- 3
pt <- slingPseudotime(sds)
nms <- colnames(pt)
nr <- ceiling(length(nms)/nc)
pal <- viridis(100, end = 0.95)
plot(reducedDim(sds), col = cell_colors_clust, pch = 16, cex = 0.5)
lines(sds, lwd = 2,col = 'black',type = 'lineages')
lines(sds, lwd = 2,col = 'black',linInd=1)
lines(sds, lwd = 2,col = 'black',linInd=2)
par(mfrow = c(1, 1))
colors <- pal[cut(pt[,1], breaks = 100)]
plot(reducedDim(sds), col = colors, pch = 16, cex = 0.5, main = 1)
lines(sds, lwd = 2, col = 'black',linInd=1)
#Export pseudotime to further optimize
PSD1<-pt
PSD1<-PSD1[,1]
as.data.frame(PSD1)
sample.combined<-AddMetaData(sample.combined, metadata=PSD1,col.name =
"PSD1")
write.csv(PSD1,file="PSD1opt.csv")
#Add optimized pseudotime back
library(readr)
PSD1opt <- read.csv("Desktop/Tonsil VDJ H chain
only/PSD1opt.csv",row.names = 1)
sample.combined<-AddMetaData(sample.combined,metadata =
PSD1opt,col.name = "PSD1opt")
#Findmarkers and export for GSEA
DimPlot(sample.combined,label = T)
markers_cluster7<-FindMarkers(sample.combined,ident.1 = "7",ident.2 =
c("2","0"))
write.csv(markers_cluster7,file="cluster7marker.csv")
#Run pseudobulk rnaseq on cluster1, 9, 0
sample.combinedsub1<-subset(x = sample.combined, idents=c("1"),invert =
F)
write.csv(sample.combinedsub1@meta.data[["seurat_clusters"]],file="seur
atclustersub1.csv")
```

```
rnacount<-
as.matrix(GetAssayData(sample.combinedsub1,assay="RNA",slot="counts"))
write.csv(rnacount,file="rawrnacountssub1.csv",sep=",",row.names =
T,col.names = NA)
sample.combinedsub0<-subset(x = sample.combined, idents=c("0"),invert =
F)
write.csv(sample.combinedsub0@meta.data[["seurat_clusters"]],file="seur
atclustersub0.csv")
rnacount<-
as.matrix(GetAssayData(sample.combinedsub0,assay="RNA",slot="counts"))
write.csv(rnacount,file="rawrnacountssub0.csv",sep=",",row.names =
T,col.names = NA)
sample.combinedsub9<-subset(x = sample.combined, idents=c("9"),invert =
F)
write.csv(sample.combinedsub9@meta.data[["seurat_clusters"]],file="seur
atclustersub9.csv")
rnacount<-
as.matrix(GetAssayData(sample.combinedsub9,assay="RNA",slot="counts"))
write.csv(rnacount,file="rawrnacountssub9.csv",sep=",",row.names =
T,col.names = NA)
#Run fgsea
ranks <- read.csv("Desktop/cluster9marker.csv",header=TRUE)
ranks <- setNames(ranks$avg_logFC, ranks$Gene_symbol)
str(ranks)
pathways <- gmtPathways("Desktop/h.all.v7.2.symbols.gmt")
fgseaRes <- fgsea(pathways, ranks, minSize=2, maxSize=30000)
#manually get some population
plot <- DimPlot(sample.combined, reduction = "umap")
select.cells <- CellSelector(plot = plot)
Idents(sample.combined, cells = select.cells) <- "11"
Idents(sample.combined, cells = select.cells) <- "12"
#write.csv(sample.combined$seurat_clusters, file =
paste("seuratcluster.csv",sep=","))
#write.csv(sample.combined$hashtagClusterID, file =
paste("hashtag.csv",sep=","))
#write.csv(clustering.table, file = paste("clusterstat.csv",sep=","))
#find markers between different RAS and CSP
Idents(sample_RC) <- sample_RC@meta.data$'orig.ident'
FindMarkers(sample_RC,ident.1="sample3project",ident.2="sample4project"
)
Idents(sample_RCTfh) <- sample_RCTfh@meta.data$'orig.ident'
FindMarkers(sample_RCTfh,ident.1="sample3project",ident.2="sample4proje
ct")
#Jon s code for adding metadata information into seurat
object~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
d1 <- sample.combined$seurat_clusters%>% data.frame()
d1 <- tibble::rownames_to_column(d1, 'barcode')
#Do statistic for IgG subtype informaiton
library(readr)
ln1 <- read_csv("Desktop/Tonsil VDJ H chain
only/filtered_contig_annotations-1.csv")
ln1['barcode_ln'] <- paste(ln1$barcode, '_1', sep='')
#sample.combined$hashtagClusterID %>% View()
ln2 <- read_csv("Desktop/Tonsil VDJ H chain
only/filtered_contig_annotations-2.csv")
ln2['barcode_ln'] <- paste(ln2$barcode, '_2', sep='')
d1_merged <- merge(x=d1, y=ln1, by.x='barcode', by.y = 'barcode_ln')
d2_merged <- merge(x=d1, y=ln2, by.x='barcode', by.y = 'barcode_ln')
```

```
BCR1<-d1_merged[,c("barcode","c_gene")]
View(BCR1)
BCR1<-BCR1[!duplicated(BCR1$barcode),]
rownames(BCR1)<-BCR1[,1]
BCR1[,1]<-NULL
BCR2<-d2_merged[,c("barcode","c_gene")]
View(BCR2)
BCR2<-BCR2[!duplicated(BCR2$barcode),]
rownames(BCR2)<-BCR2[,1]
BCR2[,1]<-NULL
BCR<<-rbind(BCR1,BCR2)
sample.combined<-AddMetaData(sample.combined, BCR1,col.name =
"sample1isotype")
sample.combined<-AddMetaData(sample.combined, BCR2,col.name =
"sample2isotype")
sample.combined<-AddMetaData(sample.combined, BCR,col.name =
"Allisotype")
#get your samples divided by treatment
#to do statistic (can also change other collumns for more information)
table1<-
table(sample.combined@meta.data[["sample1isotype"]],sample.combined@act
ive.ident)
table2<-
table(sample.combined@meta.data[["sample2isotype"]],sample.combined@act
ive.ident)
View(table1)
View(table2)
#Export IgGsubtype inf
write.csv(table1,file="sample1constant.csv")
write.csv(table2,file="sample2constant.csv")
#Calculate clonotypes-seperate2 people
clonotypeinf1<-d1_merged[,c("barcode","raw_clonotype_id")]
clonotypeinf1<-clonotypeinf1[!duplicated(clonotypeinf1$barcode),]
clonotypeinf1$raw_clonotype_id<-paste(clonotypeinf1$raw_clonotype_id,
'_1', sep='')
rownames(clonotypeinf1)<-clonotypeinf1[,1]
clonotypeinf1[,1]<-NULL
clonotypeinf2<-d2_merged[,c("barcode","raw_clonotype_id")]
clonotypeinf2<-clonotypeinf2[!duplicated(clonotypeinf2$barcode),]
clonotypeinf2$raw_clonotype_id<-paste(clonotypeinf2$raw_clonotype_id,
'_2', sep='')
rownames(clonotypeinf2)<-clonotypeinf2[,1]
clonotypeinf2[,1]<-NULL
Clonotype<<-rbind(clonotypeinf1,clonotypeinf2)
sample.combined<-AddMetaData(sample.combined, Clonotype,col.name =
"Clonotype")
tableclonotype<-
table(sample.combined@meta.data[["Clonotype"]],sample.combined@active.i
dent)
view(tableclonotype)
write.csv(tableclonotype,file="clonotypesummary.csv")
Alltheinformation<-
rbind(sample.combined@meta.data[["Clonotype"]],sample.combined@active.i
dent,sample.combined@meta.data[["Allisotype"]],sample.combined@meta.dat
a[["Allmutation"]])
Alltheinformation<-t(Alltheinformation)
write.csv(Alltheinformation,file="BCRcloneallinformation.csv")
activeident<-sample.combined@active.ident
```

```
Idents(sample.combined) <- sample.combined@meta.data[["Clonotype"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = c("clonotype1_1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = c("clonotype1_2")))
#putbackident
Idents(sample.combined) <- activeident

#calculate mutations
library(readr)
ln1 <- read.csv("Desktop/Tonsil VDJ H chain
only/sample1BCRmutation.csv",sep = ",")
ln1['Idsample'] <- paste(ln1$Idsample, '_1', sep='')
#sample.combined$hashtagClusterID %>% View()
ln2 <- read.csv("Desktop/Tonsil VDJ H chain
only/sample2BCRmutation.csv",sep = ",")
ln2['Idsample'] <- paste(ln2$Idsample, '_2', sep='')
d1_merged <- merge(x=d1, y=ln1, by.x='barcode', by.y = 'Idsample')
d2_merged <- merge(x=d1, y=ln2, by.x='barcode', by.y = 'Idsample')
BCR1<-d1_merged[,c("barcode","Mutationpercentage")]
View(BCR1)
BCR1<-BCR1[!duplicated(BCR1$barcode),]
rownames(BCR1)<-BCR1[,1]
BCR1[,1]<-NULL
BCR2<-d2_merged[,c("barcode","Mutationpercentage")]
View(BCR2)
BCR2<-BCR2[!duplicated(BCR2$barcode),]
rownames(BCR2)<-BCR2[,1]
BCR2[,1]<-NULL
BCR<<-rbind(BCR1,BCR2)
sample.combined<-AddMetaData(sample.combined, BCR1,col.name =
"sample1mutation")
sample.combined<-AddMetaData(sample.combined, BCR2,col.name =
"sample2mutation")
sample.combined<-AddMetaData(sample.combined, BCR,col.name =
"Allmutation")
FeaturePlot(sample.combined,"Allmutation")
#get your samples divided by treatment
#to do statistic (can also change other collumns for more information)
table1<-sample.combined$sample1mutation
table2<-sample.combined$sample2mutation
table3<-sample.combined$Allmutation
View(table1)
write.csv(table1,file="sample1mutation.csv")
write.csv(table2,file="sample2mutation.csv")
Idents(sample.combined) <-
sample.combined@meta.data[["laneCSPclonotype"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = "clonotype52"))
Idents(sample.combined) <-
sample.combined@meta.data[["laneRASclonotype"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = "clonotype1"))
#psedotime analysis
library(SingleCellExperiment)
seurat_object <- RunPCA(sample.combined,  npcs = 30, verbose = FALSE)
sce <- as.SingleCellExperiment(seurat_object, assay="SCT")
```

```
colData(sce)$clusters <-
as.character(seurat_object@meta.data[["Newcluster"]])
pca <- reducedDim(sce, "PCA")
#library(slingshot)
#sce <- slingshot(sce, clusterLabels = 'Seurat_clusters', reducedDim =
'PCA', allow.breaks=FALSE)
#PSD1<-sce$slingPseudotime_1
#PSD2<-sce$slingPseudotime_2
#PSD3<-sce$slingPseudotime_3
#PSD4<-sce$slingPseudotime_4
#PSD5<-sce$slingPseudotime_5
library(destiny)
library(Biobase)
#ct <- as.ExpressionSet(as.data.frame(PCA_Cluster))
#num_cells <- gsub("[[:lower:]]", "", ct$CellCluster)
#ct$num_cells <- as.integer(num_cells)
#Calculation for plotting
dm <- DiffusionMap(pca)
plot(dm,1:2,col.by="clusters")
dpt <- DPT(dm)
#Get value for diffusion pseudotime
DPT0 = dpt$dpt
DPT1 = dpt$DPT1
DPT2 = dpt$DPT2
DPT3 = dpt$DPT3
DPT4 = dpt$DPT4
DPT5 = dpt$DPT5
DC1 = eigenvectors(dm)[, 1]
DC2 = eigenvectors(dm)[, 2]
DC3 = eigenvectors(dm)[, 3]
#Generating the figures
library(plotly)
seurat_object <-
AddMetaData(seurat_object,metadata=PSD1,col.name="PSD1" )
seurat_object <-
AddMetaData(seurat_object,metadata=PSD2,col.name="PSD2" )
seurat_object <-
AddMetaData(seurat_object,metadata=PSD3,col.name="PSD3" )
seurat_object <-
AddMetaData(seurat_object,metadata=PSD4,col.name="PSD4" )
seurat_object <-
AddMetaData(seurat_object,metadata=PSD5,col.name="PSD5" )
#Can add more
seurat_object <-
AddMetaData(seurat_object,metadata=DPT0,col.name="DPT0" )
seurat_object <-
AddMetaData(seurat_object,metadata=DPT1,col.name="DPT1" )
seurat_object <-
AddMetaData(seurat_object,metadata=DPT2,col.name="DPT2" )
seurat_object <-
AddMetaData(seurat_object,metadata=DPT3,col.name="DPT3" )
seurat_object <-
AddMetaData(seurat_object,metadata=DPT4,col.name="DPT4" )
seurat_object <-
AddMetaData(seurat_object,metadata=DPT5,col.name="DPT5" )
seurat_object <-
AddMetaData(seurat_object,metadata=DC1,col.name="DC1" )
```

```
seurat_object <-
AddMetaData(seurat_object,metadata=DC2,col.name="DC2" )
seurat_object <-
AddMetaData(seurat_object,metadata=DC3,col.name="DC3" )
umap_1 <- seurat_object[["umap"]]@cell.embeddings[,1]
umap_2 <- seurat_object[["umap"]]@cell.embeddings[,2]
plot.data <- FetchData(object = seurat_object, vars = c("UMAP_1",
"AUCscore","UMAP_2","Stat5","Stat6","Stat1",
"seurat_clusters","Prdm1","Zeb2","Itgax","Cd38","Smad3","Mki67","Cxcr3"
,"Cd27","Bcl6","Stat3",
"DPT0","DPT1","DPT2","DPT3","DPT4","DC1","DC2","DC3","PSD1","PSD2","PSD
3","PSD4","PSD5"))
colorcode<-hue_pal()(13)
fig<-plot_ly(data = plot.data,
        x = ~DC1, y = ~DC2, z = ~DC3,
        color = seurat_object@active.ident,
        colors=colorcode,
        type = "scatter3d",
        mode = "markers",
        marker = list(size = 2, width=2), # controls size of points
        #text=~label, #This is that extra column we made earlier for
which we will use for cell ID
        hoverinfo="text") #When you visualize your plotly object,
hovering your mouse pointer over a point shows
fig<-fig%>%layout(legend=list(itemsizing="constant"))
fig
fig<-plot_ly(data = plot.data,
        x = ~DC1, y = ~DC2,
        color = seurat_object@active.ident,
        colors=colorcode,
        mode = "markers",
        marker = list(size = 3, width=3), # controls size of points
        #text=~label, #This is that extra column we made earlier for
which we will use for cell ID
        hoverinfo="text") #When you visualize your plotly object,
hovering your mouse pointer over a point shows
fig<-fig%>%layout(legend=list(itemsizing="constant"))
fig
```

Code for Figure 4.5, 4.7

```
library(edgeR)
library(limma)
library(tidyr)
library(Glimma)
# Process the single count matrix
RNAseqsamples <- read.delim("D:/ANU Research/data/20210729-atyB RNAseq
files and analysis/RNAseqsamplesinvivo.txt")
factorinformation <- read.delim("D:/ANU Research/data/20210729-atyB
RNAseq files and analysis/factorinformationinvivo.txt")
genes<- read.delim("D:/ANU Research/data/20210729-atyB RNAseq files and
analysis/Annotation.tabular")

groupCondition<-factorinformation$Condition
RIN<-factorinformation$RIN
design <- model.matrix(~0 + groupCondition)
colnames(design) <- gsub("groupCondition", "", colnames(design))

x2 <-RNAseqsamples
x2$Geneid<-genes$SYMBOL
x2<-drop_na(x2)

x <- x2[,-1]
rownames(x) <- x2[,1]

cpm <- cpm(x)
lcpm <- cpm(x, log=TRUE)
keep.exprs <- filterByExpr(x, group=groupCondition)
x <- x[keep.exprs,]

norm.factors <- calcNormFactors(x, method = "TMM")
cpm <- cpm(x)
lcpm <- cpm(x, log=TRUE)


lcpmPCA<-removeBatchEffect(lcpm,covariates=RIN,design = design)

#write.csv(cpm,file = "D:/ANU Research/data/20210729-atyB RNAseq files
and analysis/cpm.csv")

plotMDS(lcpmPCA,dim=c(1,2),gene.selection = "common")

mds <- plotMDS(lcpmPCA,dim=c(1,2),gene.selection = "common")

toplot <- data.frame(Dim1 = mds$x, Dim2 = mds$y)
#toplot2 <- data.frame(Dim3 = mds2$x, Dim4 = mds2$y)
library(ggplot2)
ggplot(toplot, aes(Dim1, Dim2, colour = groupCondition)) +
geom_point(size =5)+stat_ellipse()+theme_light()

glMDSPlot(lcpmPCA,gene.selection = "common")

#Plot more genes with dimensions
#logcpm<-t(lcpm)
#logcpm<-as.data.frame(logcpm)
#ggplot(toplot, aes(Dim1, logcpm$Cd72, colour = groupCondition)) +
geom_point(size =3)
```

```
#check genes in PCA1 and 2
#PCADimension1 <- mds$x
#PCADimension2 <- mds$y
#PCADimension3 <- mds2$x
#PCADimension4 <- mds2$y

#testdim1<-rbind(lcpm,PCADimension1)
#testdim2<-rbind(lcpm,PCADimension2)
#testdim3<-rbind(lcpm,PCADimension3)
#testdim4<-rbind(lcpm,PCADimension4)

#testdim1<-t(testdim1)
#testdim2<-t(testdim2)
#testdim3<-t(testdim3)
#testdim4<-t(testdim4)

#testdim1<-as.data.frame(testdim1)
#testdim2<-as.data.frame(testdim2)
#testdim3<-as.data.frame(testdim3)
#testdim4<-as.data.frame(testdim4)

#library("Hmisc")
#library(corrplot)
#dim1genes <- cor(x = testdim1[13450:13450], y = testdim1[1:13450],
use="complete.obs")
#dim1genes<-t(dim1genes)

#dim2genes <- cor(x = testdim2[13450:13450], y = testdim2[1:13450],
use="complete.obs")
#dim2genes<-t(dim2genes)

#dim3genes <- cor(x = testdim3[13450:13450], y = testdim3[1:13450],
use="complete.obs")
#dim3genes<-t(dim3genes)

#dim4genes <- cor(x = testdim4[13450:13450], y = testdim4[1:13450],
use="complete.obs")
#dim4genes<-t(dim4genes)

#Differentially expressed gene
contr.matrix <- makeContrasts(
  Day4compare = ABday4-CBday4,
  Day10compare =ABday10-CBday10,
  Day21compare =ABday21-CBday21,
  levels = colnames(design))
contr.matrix
v <- voom(x, design, plot=TRUE)
vfit <- lmFit(v, design)
vfit <- contrasts.fit(vfit, contrasts=contr.matrix)
efit <- eBayes(vfit)
plotSA(efit, main="Final model: Mean-variance trend")

summary(decideTests(efit,adjust.method="none",p.value=0.05)) #or
method=none
tfit <- treat(vfit, lfc=0)
dt <- decideTests(tfit,adjust.method="none",p.value=0.05)
summary(dt)
```

```
de.common <- which(dt[,1]!=0 & dt[,2]!=0 & dt[,3]!=0)
de.common <- which(dt[,2]!=0 & dt[,3]!=0)
de.common <- which(dt[,2]!=0)
length(de.common)
head(vfit[["Amean"]][de.common], n=96)
commongenes<-as.data.frame(vfit[["Amean"]][de.common], n=96)
#commongenes<-as.data.frame(vfit[["Amean"]][de.common], n=373)
#commongenes<-as.data.frame(vfit[["Amean"]][de.common], n=1952)
commongenes<-rownames(commongenes)
vennDiagram(dt[,1:3], circle.col=c("red", "blue","orange"))
vennDiagram(dt[,2:3], circle.col=c("red", "blue"))

Day4compare <- topTreat(tfit, coef=1, n=Inf)
Day10compare <- topTreat(tfit, coef=2, n=Inf)
Day21compare <- topTreat(tfit, coef=3, n=Inf)

#write.csv(Day4compare,file = "D:/ANU Research/data/20210729-atyB
RNAseq files and analysis/Day4compare.csv")
#write.csv(Day10compare,file = "D:/ANU Research/data/20210729-atyB
RNAseq files and analysis/Day10compare.csv")
#write.csv(Day21compare,file = "D:/ANU Research/data/20210729-atyB
RNAseq files and analysis/Day21compare.csv")
#plotMD(tfit, column=1, status=dt[,1], main=colnames(tfit)[1],
xlim=c(-8,13))

#make heatmap
library(gplots)
mycol <- colorpanel(1000,"blue","white","red")
heatmap.2(lcpm[commongenes,], scale="row",
          col=mycol, trace="none", density.info="none",
dendrogram="column",cexRow = 0.8,cexCol = 0.8,
reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean),
          distfun=function(x) as.dist(1-cor(t(x))),
          hclustfun=function(x) hclust(x,
method="complete"),labRow=commongenes)

#Perform fgsea
library(fgsea)
ranks<-geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq
files and analysis/Day10compare_rank.csv",header=TRUE, colClasses =
c("character", "numeric"))
ranks<-geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq
files and analysis/Day4compare_rank.csv",header=TRUE, colClasses =
c("character", "numeric"))
ranks<-geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq
files and analysis/Day21compare_rank.csv",header=TRUE, colClasses =
c("character", "numeric"))

ranks <- setNames(ranks$t, ranks$X)
str(ranks)
geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq files
and analysis/Analyze Becans data/CD11c_up_fdr005.txt")
fgseaRes <- fgsea(pathways = geneset,
                  stats    = ranks,
                  minSize  = 15,
                  maxSize  = 500)
library(ggplot2)
```

```
plotEnrichment(geneset[["CD11c_up_fdr005"]],ranks) + labs(title =
paste(title=fgseaRes$padj),subtitle=paste(fgseaRes$NES))

#Repeat DEG analysis on different days
factorinformation <- read.delim("D:/ANU Research/data/20210729-atyB
RNAseq files and analysis/factorinformationinvivo_days.txt")
groupCondition<-factorinformation$Condition
RIN<-factorinformation$RIN
design <- model.matrix(~0 + groupCondition)
colnames(design) <- gsub("groupCondition", "", colnames(design))

#Differentially expressed gene
contr.matrix <- makeContrasts(
  Day4_10_21compare_AB = ABday4-ABday10_21,
  Day4_10_21compare_CB =CBday4-CBday10_21,
  levels = colnames(design))
contr.matrix
v <- voom(x, design, plot=TRUE)
vfit <- lmFit(v, design)
vfit <- contrasts.fit(vfit, contrasts=contr.matrix)
efit <- eBayes(vfit)
plotSA(efit, main="Final model: Mean-variance trend")

summary(decideTests(efit,adjust.method="BF",p.value=0.05))
tfit <- treat(vfit, lfc=0)
dt <- decideTests(tfit,adjust.method="BF",p.value=0.05)
summary(dt)
de.common <- which(dt[,1]!=0 & dt[,2]!=0)
length(de.common)
head(vfit[["Amean"]][de.common], n=1365)
commongenes<-as.data.frame(vfit[["Amean"]][de.common], n=1365)
commongenes<-rownames(commongenes)
vennDiagram(dt[,1:2], circle.col=c("red", "blue"))

Day4_10_21compare_AB <- topTreat(tfit, coef=1, n=Inf)
Day4_10_21compare_CB <- topTreat(tfit, coef=2, n=Inf)

#write.csv(Day4_10_21compare_AB,file = "D:/ANU Research/data/20210729-
atyB RNAseq files and analysis/Day4_10_21compare_AB.csv")
#write.csv(Day4_10_21compare_CB,file = "D:/ANU Research/data/20210729-
atyB RNAseq files and analysis/Day4_10_21compare_CB.csv")
#plotMD(tfit, column=1, status=dt[,1], main=colnames(tfit)[1],
xlim=c(-8,13))

#make heatmap
library(gplots)
mycol <- colorpanel(1000,"blue","white","red")
heatmap.2(lcpm[commongenes,], scale="row",
          col=mycol, trace="none", density.info="none",
dendrogram="column",cexRow = 0.8,cexCol = 0.8,
reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean),
          distfun=function(x) as.dist(1-cor(t(x))),
          hclustfun=function(x) hclust(x,
method="complete"),labRow=commongenes)


#Perform fgsea
library(fgsea)
```

```
ranks<-geneset <- read.csv("D:/AtyBMS/OneDrive - Australian National
University/Figure2/20210729-atyB RNAseq files and
analysis/Day10compare_rank.csv",header=TRUE, colClasses =
c("character", "numeric"))
ranks<-geneset <- read.csv("D:/AtyBMS/OneDrive - Australian National
University/Figure2/20210729-atyB RNAseq files and analysis/Analyze
Becans data/CD11cvsDNcompare.csv",header=TRUE, colClasses =
c("character", "numeric"))

ranks <- setNames(ranks$t, ranks$X)
str(ranks)
geneset <- read.csv("C:/Users/xiong/Downloads/GSE150124_diffexpr-
results_geneIDCD11.csv/GSE150124 CD11+ gene list.txt")
geneset <- read.csv("C:/Users/xiong/Downloads/GSE150124_diffexpr-
results_geneIDCD11.csv/Plasmodium-specific atypical memory B cells are
short-lived activated B cells_genelist.txt")
fgseaRes <- fgsea(pathways = geneset,
                  stats    = ranks,
                  minSize  = 15,
                  maxSize  = 5000000)
library(ggplot2)
plotEnrichment(geneset[["GeneID"]],ranks) + labs(title =
paste(title=fgseaRes$padj),subtitle=paste(fgseaRes$NES))


#Check geneset overall expression
geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq files
and analysis/Analyze Becans data/BecanCD71_up_fdr025.txt")
geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq files
and analysis/Analyze Becans data/CD11c_up_fdr005.txt")
geneset <- read.csv("D:/ANU Research/data/20210729-atyB RNAseq files
and analysis/Analyze Becans data/CD11c_down_fdr005.txt")

#find common genes in case any bug
rownamesvalue<-as.data.frame(rownames(lcpm))
common <- intersect(rownamesvalue$`rownames(lcpm)`,
geneset$BecanCD71_Upfdr025)
common <- intersect(rownamesvalue$`rownames(lcpm)`,
geneset$CD11c_up_fdr005)
common <- intersect(rownamesvalue$`rownames(lcpm)`,
geneset$CD11c_down_fdr005)

mycol <- colorpanel(1000,"blue","white","red")
heatmap.2(lcpm[common,], scale="row",
          col=mycol, trace="none", density.info="none",
dendrogram="column",cexRow = 0.8,cexCol = 0.8,
reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean),
          distfun=function(x) as.dist(1-cor(t(x))),
          hclustfun=function(x) hclust(x,
method="complete"),labRow=common)


#Check some but not all samples
lcpmselect<-lcpm[,c("CBday4_1","CBday4_2","ABday4_1","ABday4_2")]
lcpmselect<-lcpm[,c("CBday21_1","CBday21_2","ABday21_1","ABday21_2")]
lcpmselect<-
lcpm[,c("CBday10_1","CBday10_2","CBday10_3","ABday10_1","ABday10_2","AB
day10_3")]
```

```
mycol <- colorpanel(1000,"blue","white","red")
heatmap.2(lcpmselect[common,], scale="row",
          col=mycol, trace="none", density.info="none",
dendrogram="column",cexRow = 0.8,cexCol = 0.8,
reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean),
          distfun=function(x) as.dist(1-cor(t(x))),
          hclustfun=function(x) hclust(x,
method="complete"),labRow=common)


#Make selected genes heatmap
Atypicalgenes<-
c("Itgax","Tbx21","Zeb2","Hck","Fcrl5","Bhlhe40","Cd19","Fas","Tcf7","Z
eb1","Cr2","Cxcr5","Cd55","Ccr7","Fcer2a","Myc","Itgam","Sell","Lgals1"
)
heatmap.2(lcpm[Atypicalgenes,], scale="row",
          col=mycol, trace="none", density.info="none",
dendrogram="column",cexRow = 0.8,cexCol = 0.8,
reorderfun=function(d,w) reorder(d, w, agglo.FUN=mean),
          distfun=function(x) as.dist(1-cor(t(x))),
          hclustfun=function(x) hclust(x,
method="complete"),labRow=Atypicalgenes)
```

Code for Figure 5.3

```
library(Seurat)
data_dir <-
"~/Desktop/cellranger_outs/Ab_Henry_3/filtered_feature_bc_matrix/"
list.files(data_dir)
expression_matrix3 <- Read10X(data.dir = data_dir)
sample3 = CreateSeuratObject(counts = expression_matrix3$`Gene
Expression`,project = "sample3project")
sample3[['Protein']] = CreateAssayObject(counts =
expression_matrix3$`Antibody Capture`)
library("cowplot")
library("dplyr")
library("ggplot2")
sample3[["percent.mt"]] <- PercentageFeatureSet(sample3, pattern =
"^mt-")
VlnPlot(sample3, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
plot1 <- FeatureScatter(sample3, feature1 = "nCount_RNA", feature2 =
"percent.mt")
plot2 <- FeatureScatter(sample3, feature1 = "nCount_RNA", feature2 =
"nFeature_RNA")
plot1 + plot2
sample3 <- subset(sample3, subset = nFeature_RNA > 200 & nFeature_RNA <
4000 &nCount_RNA<20000 & percent.mt < 7)
sample3 <- NormalizeData(sample3, assay = "Protein",
normalization.method = "CLR")
sample3 <- ScaleData(sample3, assay = "Protein")
DefaultAssay(sample3) <- "Protein"
sample3 <- RunPCA(sample3, features = c("Mouse-hashtag-5-
TotalSeqC","Mouse-hashtag-4-TotalSeqC","Mouse-hashtag-3-
TotalSeqC","Mouse-hashtag-2-TotalSeqC","Mouse-hashtag-1-TotalSeqC") ,
reduction.name = "pca_protein", reduction.key = "pca_protein_", verbose
= FALSE)
DimPlot(sample3, reduction = "pca_protein")
sample3 <- FindNeighbors(sample3, features=c("Mouse-hashtag-5-
TotalSeqC","Mouse-hashtag-4-TotalSeqC","Mouse-hashtag-3-
TotalSeqC","Mouse-hashtag-2-TotalSeqC","Mouse-hashtag-1-TotalSeqC"),
dims = NULL)
sample3 <- FindClusters(sample3, resolution = 0.4)
DimPlot(sample3)
DimPlot(sample3,split.by = "ident",ncol=6)
FeaturePlot(sample3,"Mouse-hashtag-5-TotalSeqC")
FeaturePlot(sample3,"Mouse-hashtag-4-TotalSeqC")
FeaturePlot(sample3,"Mouse-hashtag-3-TotalSeqC")
FeaturePlot(sample3,"Mouse-hashtag-2-TotalSeqC")
FeaturePlot(sample3,"Mouse-hashtag-1-TotalSeqC")
sample3 <- subset(x = sample3,
idents=c("5","6","7","8","9","10","11"),invert = TRUE)
DimPlot(sample3)
current.cluster.ids <- c(0, 1, 2, 3, 4)
new.cluster.ids <- c("CSP1", "CSP3", "no2a104", "no2a105","CSP2")
sample3@active.ident <- plyr::mapvalues(x = sample3@active.ident, from
= current.cluster.ids, to = new.cluster.ids)
sample3[["hashtagClusterID"]] <- Idents(sample3)
#-------------------------------------------
#------ now the secound sample
```

```
data_dir <-
"~/Desktop/cellranger_outs/Ab_Henry_4/filtered_feature_bc_matrix/"
list.files(data_dir)
expression_matrix4 <- Read10X(data.dir = data_dir)
sample4 = CreateSeuratObject(counts = expression_matrix4$`Gene
Expression`,project="sample4project")
sample4[['Protein']] = CreateAssayObject(counts =
expression_matrix4$`Antibody Capture`)
sample4[["percent.mt"]] <- PercentageFeatureSet(sample4, pattern =
"^mt-")
VlnPlot(sample4, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
plot1 <- FeatureScatter(sample4, feature1 = "nCount_RNA", feature2 =
"percent.mt")
plot2 <- FeatureScatter(sample4, feature1 = "nCount_RNA", feature2 =
"nFeature_RNA")
plot1 + plot2
sample4 <- subset(sample4, subset = nFeature_RNA > 200 & nFeature_RNA <
4000 &nCount_RNA<20000 & percent.mt < 7)
sample4 <- NormalizeData(sample4, assay = "Protein",
normalization.method = "CLR")
sample4 <- ScaleData(sample4, assay = "Protein")
DefaultAssay(sample4) <- "Protein"
sample4 <- RunPCA(sample4, features = c("Mouse-hashtag-5-
TotalSeqC","Mouse-hashtag-4-TotalSeqC","Mouse-hashtag-3-
TotalSeqC","Mouse-hashtag-2-TotalSeqC","Mouse-hashtag-1-TotalSeqC") ,
reduction.name = "pca_protein", reduction.key = "pca_protein_", verbose
= FALSE)
DimPlot(sample4, reduction = "pca_protein")
FeaturePlot(sample4,"Mouse-hashtag-5-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-4-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-3-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-2-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-1-TotalSeqC")
sample4 <- FindNeighbors(sample4, features = c("Mouse-hashtag-5-
TotalSeqC","Mouse-hashtag-4-TotalSeqC","Mouse-hashtag-3-
TotalSeqC","Mouse-hashtag-2-TotalSeqC","Mouse-hashtag-1-TotalSeqC"),
dims = NULL)
sample4 <- FindClusters(sample4, resolution = 0.2)
DimPlot(sample4)
FeaturePlot(sample4,"Mouse-hashtag-5-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-4-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-3-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-2-TotalSeqC")
FeaturePlot(sample4,"Mouse-hashtag-1-TotalSeqC")
sample4 <- subset(x = sample4,
idents=c("0","1","4","6","8","10"),invert = TRUE)
DimPlot(sample4)
current.cluster.ids <- c(2, 3, 5, 7, 9)
new.cluster.ids <- c("RAS1", "RAS3", "NT5", "RAS2","NT4")
sample4@active.ident <- plyr::mapvalues(x = sample4@active.ident, from
= current.cluster.ids, to = new.cluster.ids)
sample4[["hashtagClusterID"]] <- Idents(sample4)
#----------Run integration and downstream analysis
VlnPlot(sample3, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
sample3 <- subset(sample3, subset = nCount_RNA > 1000 & percent.mt < 5)
```

```
VlnPlot(sample4, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
sample4 <- subset(sample4, subset = nCount_RNA > 1000 & percent.mt < 5)
#-----------
DefaultAssay(sample3) <- "RNA"
DefaultAssay(sample4) <- "RNA"
sample3 <- NormalizeData(sample3)
sample4 <- NormalizeData(sample4)
sample.anchors <- FindIntegrationAnchors(object.list = list(sample3,
sample4), anchor.features = 2000) #Max number of intergrated features
by 40000
sample.combined <- IntegrateData(anchorset = sample.anchors)
sample.combinedbackup<-sample.combined
#sample.combined<-sample.combinedbackup
rm(expression_matrix3,
expression_matrix4,sample.anchors,sample3,sample4,plot1,plot2)
sample.combined <- ScaleData(sample.combined,features =
rownames(sample.combined))
sample.combined <- RunPCA(sample.combined, verbose = FALSE)
ElbowPlot(sample.combined, ndims =50)
sample.combined <- FindNeighbors(sample.combined, dims = 1:30)
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:30)
sample.combined <- FindClusters(sample.combined, resolution = 0.3)
DimPlot(sample.combined)
VlnPlot(sample.combined, features = c("nFeature_RNA", "nCount_RNA",
"percent.mt"), ncol = 3)
DimPlot(sample.combined,split.by = "orig.ident")
DimPlot(sample.combined,split.by = "hashtagClusterID",ncol=5)
FeaturePlot(sample.combined,features
=c("rna_Ncr1","rna_Cd19","rna_Itgax","rna_Itgam","rna_Ly6c1"))
FeaturePlot(sample.combined,features
=c("rna_Cxcr5","rna_Cd40lg","rna_Il21","rna_Icos","rna_Bcl6","rna_Cd4",
"rna_Foxp3","rna_Mki67","rna_Sell","rna_Cd44"),ncol=5)
sample.markers <- FindAllMarkers(sample.combined, only.pos = TRUE)
top10markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 10, wt
= avg_logFC)
sample.combined.small<- subset(sample.combined, downsample = 100)
DoHeatmap(sample.combined.small, features = unique(top10markers$gene),
angle = 90) + NoLegend() + theme(text = element_text(size = 8))
sample.combined <- subset(x = sample.combined,
idents=c("6","7","8","9","10"),invert = TRUE) #6naive,
7dead,8NK,9dcPdcepi ifit3,10 neutrophil
#recluster and remove ribosome trv genes
all.genes<-rownames(sample.combined@assays[["RNA"]])
Ribosome <- all.genes[grep('^Rps|^Rpl', all.genes)]
TRAV <- all.genes[grep('^Trav', all.genes)]
TRBV <- all.genes[grep('^Trbv', all.genes)]
igh <- all.genes[grep('^Ighv', all.genes)]
igl <- all.genes[grep('^Igl|^Igl', all.genes)]
var.genes.no.riboigg <-
setdiff(sample.combined@assays[["integrated"]]@var.features, Ribosome)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, TRAV)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, TRBV)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, igh)
var.genes.no.riboigg <- setdiff(var.genes.no.riboigg, igl)
sample.combined <- RunPCA(sample.combined, verbose = FALSE,features
=var.genes.no.riboigg )
```

```
ElbowPlot(sample.combined, ndims =50)
sample.combined <- FindNeighbors(sample.combined, dims = 1:15)
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:15)
sample.combined <- FindClusters(sample.combined, resolution = 0.18)
DimPlot(sample.combined)
DimPlot(sample.combined,split.by = "hashtagClusterID",ncol = 5)
FeaturePlot(sample.combined,features =
c("Cxcr5","Sell","Foxp3","Mki67"),min.cutoff = 0)
#only choose memory population (no code here)
sample.combined <- subset(x = sample.combined, idents=c("0"),invert =
FALSE)
sample.combined <- RunPCA(sample.combined, verbose = FALSE,features
=var.genes.no.riboigg)
ElbowPlot(sample.combined, ndims =50)
sample.combined <- FindNeighbors(sample.combined, dims = 1:20)
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:20)
sample.combined <- FindClusters(sample.combined, resolution = 1.2)
DimPlot(sample.combined)
DimPlot(sample.combined,split.by = "hashtagClusterID",ncol = 5)
#remove a strange population 10 or something, no code here
sample.combined <- subset(x = sample.combined, idents=c("12"),invert =
TRUE)
sample.combined <- RunPCA(sample.combined, verbose = FALSE,features
=var.genes.no.riboigg)
ElbowPlot(sample.combined, ndims =50)
sample.combined <- FindNeighbors(sample.combined, dims = 1:16)
sample.combined <- RunUMAP(sample.combined, reduction = "pca", dims =
1:16)
sample.combined <- FindClusters(sample.combined, resolution = 0.55)
DimPlot(sample.combined,label = T)
DimPlot(sample.combined,split.by = "hashtagClusterID",ncol = 5)
#then manualy gate on ras1-ras2 specific populations
plot <- DimPlot(sample.combined, reduction = "umap")
select.cells <- CellSelector(plot = plot)
Idents(sample.combined, cells = select.cells) <- "RASTfh2"
#FindMarkers(sample.combined, ident.1 = "NewCells", ident.2 =
c("0","1","3"))
sample.markers <- FindAllMarkers(sample.combined)
top20markers<-sample.markers %>% group_by(cluster) %>% top_n(n = 20, wt
= avg_logFC)
sample.combined.small<- subset(sample.combined, downsample = 100)
DoHeatmap(sample.combined.small, features = unique(top20markers$gene),
angle = 90) + NoLegend() + theme(text = element_text(size = 7))
DimPlot(sample.combined,split.by = "orig.ident")
DimPlot(sample.combined,split.by = "hashtagClusterID",ncol=5,pt.size=1)
FeaturePlot(sample.combined,"rna_Il21")
FeaturePlot(sample.combined,"rna_Cd40lg")
FeaturePlot(sample.combined,"rna_Cd69")
FeaturePlot(sample.combined,"rna_Icos")
FeaturePlot(sample.combined,"rna_Pdcd1")
FeaturePlot(sample.combined,"rna_Bcl6")
FeaturePlot(sample.combined,"rna_Cxcr5")
FeaturePlot(sample.combined,"rna_Ascl2")
FeaturePlot(sample.combined,"rna_Cxcr3")
FeaturePlot(sample.combined,"rna_Tbx21")
FeaturePlot(sample.combined,"rna_Cxcr6")
```

```
FeaturePlot(sample.combined,"rna_Lag3")
FeaturePlot(sample.combined,"rna_Cfl1")
FeaturePlot(sample.combined,"rna_Hif1a")
FeaturePlot(sample.combined,"CXCR3-TotalSeqC")
FeaturePlot(sample.combined,"AntiAPC-TotalSeqC")
FeaturePlot(sample.combined,features =
c("rna_Il21","rna_Bcl6","rna_Hif1a","rna_Ascl2","rna_Pdcd1","rna_Cxcr5"
,"rna_Lag3","rna_Cd69","rna_Cxcr6","rna_Cxcr3"),ncol = 5)
sample.combinedbackup<-sample.combined
#Check the selected RAS specific population
plot <- DimPlot(sample.combined, reduction = "umap")
select.cells <- CellSelector(plot = plot)
Idents(sample.combined, cells = select.cells) <- "RASTfh2"
FindMarkers(sample.combined, ident.1 = "NewCells", ident.2 =
c("0","1","3"))
FeaturePlot(sample.combined,"Actg1", min.cutoff = 3)
DimPlot(sample.combined)
#S G phases
library(dplyr)
s.genes <- cc.genes$s.genes
g2m.genes <- cc.genes$g2m.genes
sample.combined <- CellCycleScoring(sample.combined, s.features =
s.genes, g2m.features = g2m.genes, ignore_case = TRUE,set.ident = TRUE)
DimPlot(sample.combined)
Idents(sample.combined) <-
sample.combined@meta.data[["seurat_clusters"]]
sample_CSPRAS<-subset(x = sample.combined, subset
=hashtagClusterID=="no2a104",invert = TRUE)
sample_CSPRAS<-subset(x = sample_CSPRAS, subset
=hashtagClusterID=="no2a105",invert = TRUE)
sample_CSPRAS<-subset(x = sample_CSPRAS, subset
=hashtagClusterID=="NT4",invert = TRUE)
sample_CSPRAS<-subset(x = sample_CSPRAS, subset
=hashtagClusterID=="NT5",invert = TRUE)
sample_matureTfh <- subset(x = sample_CSPRAS, idents=c("1"),invert =
FALSE)
sample_preTfh <- subset(x = sample_CSPRAS, idents=c("0"),invert =
FALSE)
sample_newlygeneratedTfh <- subset(x = sample_CSPRAS,
idents=c("3"),invert = FALSE)
sample_TFR <- subset(x = sample_CSPRAS, idents=c("2"),invert = FALSE)
Idents(sample_preTfh) <- sample_preTfh@meta.data$'orig.ident'
Idents(sample_newlygeneratedTfh) <-
sample_newlygeneratedTfh@meta.data$'orig.ident'
Idents(sample_TFR) <- sample_TFR@meta.data$'orig.ident'
Idents(sample_matureTfh) <- sample_matureTfh@meta.data$'orig.ident'
sample_CSP <- subset(x = sample.combined, subset
=hashtagClusterID=="no2a104",invert = TRUE)
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="no2a105",invert = TRUE)
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="NT4",invert = TRUE)
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="NT5",invert = TRUE)
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="RAS1",invert = TRUE)
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="RAS2",invert = TRUE)
```

```
sample_CSP <- subset(x = sample_CSP, subset
=hashtagClusterID=="RAS3",invert = TRUE)
sample_CSP1 <- subset(x = sample_CSP, subset
=hashtagClusterID=="CSP1",invert = FALSE)
sample_CSP2 <- subset(x = sample_CSP, subset
=hashtagClusterID=="CSP2",invert = FALSE)
sample_CSP3 <- subset(x = sample_CSP, subset
=hashtagClusterID=="CSP3",invert = FALSE)
sample_RAS <- subset(x = sample.combined, subset
=hashtagClusterID=="no2a104",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="no2a105",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="NT4",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="NT5",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="CSP1",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="CSP2",invert = TRUE)
sample_RAS <- subset(x = sample_RAS, subset
=hashtagClusterID=="CSP3",invert = TRUE)
sample_RAS1 <- subset(x = sample_RAS, subset
=hashtagClusterID=="RAS1",invert = FALSE)
sample_RAS2 <- subset(x = sample_RAS, subset
=hashtagClusterID=="RAS2",invert = FALSE)
sample_RAS3 <- subset(x = sample_RAS, subset
=hashtagClusterID=="RAS3",invert = FALSE)
sample_no2a10 <- subset(x = sample.combined, subset
=hashtagClusterID=="NT4",invert = TRUE)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="NT5",invert = TRUE)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="CSP1",invert = TRUE)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="CSP2",invert = TRUE)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="CSP3",invert = TRUE)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="RAS1",invert = T)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="RAS2",invert = T)
sample_no2a10 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="RAS3",invert = T)
sample_no2a104 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="no2a104",invert = F)
sample_no2a105 <- subset(x = sample_no2a10, subset
=hashtagClusterID=="no2a105",invert = F)
sample_NT <- subset(x = sample.combined, subset
=hashtagClusterID=="no2a104",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="no2a105",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="CSP1",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="CSP2",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="CSP3",invert = TRUE)
```

```
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="RAS1",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="RAS2",invert = TRUE)
sample_NT <- subset(x = sample_NT, subset
=hashtagClusterID=="RAS3",invert = TRUE)
sample_NT4 <- subset(x = sample_NT, subset
=hashtagClusterID=="NT4",invert = F)
sample_NT5 <- subset(x = sample_NT, subset
=hashtagClusterID=="NT5",invert = F)
sample.combined<-BuildClusterTree(sample.combined, dim=T)
PlotClusterTree(sample.combined)
clustering.table <- table(Idents(sample.combined),
sample.combined$hashtagClusterID)
clustering.table
#write.csv(sample.combined$seurat_clusters, file =
paste("seuratcluster.csv",sep=","))
#write.csv(sample.combined$hashtagClusterID, file =
paste("hashtag.csv",sep=","))
#write.csv(clustering.table, file = paste("clusterstat.csv",sep=","))
#compare more differences between spz and alum (such as IL4, cxcl10...)
Idents(sample.combined) <-
sample.combined@meta.data[["seurat_clusters"]]
sample_RC <- subset(x = sample.combined, subset
=hashtagClusterID=="no2a104",invert = TRUE)
sample_RC <- subset(x = sample_RC, subset
=hashtagClusterID=="no2a105",invert = TRUE)
sample_RC <- subset(x = sample_RC, subset
=hashtagClusterID=="NT4",invert = TRUE)
sample_RC <- subset(x = sample_RC, subset
=hashtagClusterID=="NT5",invert = TRUE)
sample_RCTfh <- subset(x = sample_RC, idents="1",invert = FALSE)
sample_Tfh <- subset(x = sample.combined, idents="1",invert = FALSE)
sample_activatedTfh <- subset(x = sample.combined, idents="7",invert =
FALSE)
#find markers between different RAS and CSP
Idents(sample_RC) <- sample_RC@meta.data$'orig.ident'
FindMarkers(sample_RC,ident.1="sample3project",ident.2="sample4project"
)
Idents(sample_RCTfh) <- sample_RCTfh@meta.data$'orig.ident'
FindMarkers(sample_RCTfh,ident.1="sample3project",ident.2="sample4proje
ct")
FeaturePlot(sample_RCTfh,"rna_Cxcl10",split.by = "orig.ident",
pt.size=0.5)
#Jon s code for adding metadata information into seurat
object~~~~~~~~~~~~~~~~~~~~~~~~~~~~
d1 <- sample.combined$seurat_clusters%>% data.frame()
d1 <- tibble::rownames_to_column(d1, 'barcode')
library(readr)
ln1 <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_3/filtered_contig_annotatio
ns.csv")
ln1['barcode_ln'] <- paste(ln1$barcode, '_1', sep='')
ln1a <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_3/filtered_contig_annotatio
ns3a.csv")
ln1a['barcode_ln'] <- paste(ln1a$barcode, '_1', sep='')
```

```
ln1b <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_3/filtered_contig_annotatio
ns3b.csv")
ln1b['barcode_ln'] <- paste(ln1b$barcode, '_1', sep='')
#sample.combined$hashtagClusterID %>% View()
ln2 <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_4/filtered_contig_annotatio
ns.csv")
ln2['barcode_ln'] <- paste(ln2$barcode, '_2', sep='')
ln2a <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_4/filtered_contig_annotatio
ns4a.csv")
ln2a['barcode_ln'] <- paste(ln2a$barcode, '_2', sep='')
ln2b <-
read_csv("Desktop/cellranger_outs/VDJ_Henry_4/filtered_contig_annotatio
ns4b.csv")
ln2b['barcode_ln'] <- paste(ln2b$barcode, '_2', sep='')
d1_merged <- merge(x=d1, y=ln1, by.x='barcode', by.y = 'barcode_ln')
d2_merged <- merge(x=d1, y=ln2, by.x='barcode', by.y = 'barcode_ln')
d1a_merged <- merge(x=d1, y=ln1a, by.x='barcode', by.y = 'barcode_ln')
d2a_merged <- merge(x=d1, y=ln2a, by.x='barcode', by.y = 'barcode_ln')
d1b_merged <- merge(x=d1, y=ln1b, by.x='barcode', by.y = 'barcode_ln')
d2b_merged <- merge(x=d1, y=ln2b, by.x='barcode', by.y = 'barcode_ln')
tcr1<-d1_merged[,c("barcode","raw_clonotype_id")]
View(tcr1)
tcr1<-tcr1[!duplicated(tcr1$barcode),]
rownames(tcr1)<-tcr1[,1]
tcr1[,1]<-NULL
tcr2<-d2_merged[,c("barcode","raw_clonotype_id")]
View(tcr2)
tcr2<-tcr2[!duplicated(tcr2$barcode),]
rownames(tcr2)<-tcr2[,1]
tcr2[,1]<-NULL
tcr3a<-d1a_merged[,c("barcode","v_gene")]
tcr3a<-tcr3a[!duplicated(tcr3a$barcode),]
rownames(tcr3a)<-tcr3a[,1]
tcr3a[,1]<-NULL
tcr3b<-d1b_merged[,c("barcode","v_gene")]
tcr3b<-tcr3b[!duplicated(tcr3b$barcode),]
rownames(tcr3b)<-tcr3b[,1]
tcr3b[,1]<-NULL
tcr4a<-d2a_merged[,c("barcode","v_gene")]
tcr4a<-tcr4a[!duplicated(tcr4a$barcode),]
rownames(tcr4a)<-tcr4a[,1]
tcr4a[,1]<-NULL
tcr4b<-d2b_merged[,c("barcode","v_gene")]
tcr4b<-tcr4b[!duplicated(tcr4b$barcode),]
rownames(tcr4b)<-tcr4b[,1]
tcr4b[,1]<-NULL
sample.combined<-AddMetaData(sample.combined, tcr1,col.name =
"laneCSPclonotype")
sample.combined<-AddMetaData(sample.combined, tcr2,col.name =
"laneRASclonotype")
sample.combined<-AddMetaData(sample.combined, tcr3a,col.name =
"laneCSPVA")
sample.combined<-AddMetaData(sample.combined, tcr3b,col.name =
"laneCSPVB")
```

```
sample.combined<-AddMetaData(sample.combined, tcr4a,col.name =
"laneRASVA")
sample.combined<-AddMetaData(sample.combined, tcr4b,col.name =
"laneRASVB")
#get your samples divided by treatment
#to do statistic (can also change other collumns for more information)
table1<-
table(sample.combined@meta.data[["laneCSPclonotype"]],sample.combined$h
ashtagClusterID)
table2<-
table(sample.combined@meta.data[["laneRASclonotype"]],sample.combined$h
ashtagClusterID)
table3a<-
table(sample.combined@meta.data[["laneCSPVA"]],sample.combined$hashtagC
lusterID)
table3b<-
table(sample.combined@meta.data[["laneCSPVB"]],sample.combined$hashtagC
lusterID)
table4a<-
table(sample.combined@meta.data[["laneRASVA"]],sample.combined$hashtagC
lusterID)
table4b<-
table(sample.combined@meta.data[["laneRASVB"]],sample.combined$hashtagC
lusterID)
View(table1)
View(table2)
View(table3a)
write.csv(table3a,file="3VAgene.csv")
write.csv(table3b,file="3VBgene.csv")
write.csv(table4a,file="4VAgene.csv")
write.csv(table4b,file="4VBgene.csv")

Idents(sample.combined) <-
sample.combined@meta.data[["laneCSPclonotype"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = "clonotype52"))
Idents(sample.combined) <-
sample.combined@meta.data[["laneRASclonotype"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample.combined,
idents = "clonotype1"))
table1<-
table(sample_RCTfh@meta.data[["laneCSPclonotype"]],sample_RCTfh$hashtag
ClusterID)
table2<-
table(sample_RCTfh@meta.data[["laneRASclonotype"]],sample_RCTfh$hashtag
ClusterID)
View(table1)
View(table2)
#try to combined TCR information and put into seurat meta
suppressMessages(library(scRepertoire))
library(readr)
filtered_contig_annotations3 <-
read.csv("Desktop/cellranger_outs/VDJ_Henry_3/filtered_contig_annotatio
ns.csv", stringsAsFactors=F)
filtered_contig_annotations4 <-
read.csv("Desktop/cellranger_outs/VDJ_Henry_4/filtered_contig_annotatio
ns.csv", stringsAsFactors=F)
```

```
combinedTCR3a <- combineTCR(filtered_contig_annotations3 , samples =
"", ID = "",cells="T-AB", filterMulti = T)
combinedTCR4a <- combineTCR(filtered_contig_annotations4 , samples =
"", ID = "",cells="T-AB", filterMulti = T)
combinedTCR3<-as.data.frame(combinedTCR3a[["_"]])
combinedTCR4<-as.data.frame(combinedTCR4a[["_"]])
combinedTCR3$barcode<-sub(".","",combinedTCR3$barcode) #Run twice,
sometime doesnt work
combinedTCR4$barcode<-sub(".","",combinedTCR4$barcode)
combinedTCR3$barcode <- paste(combinedTCR3$barcode, '_1', sep='')
combinedTCR4$barcode <- paste(combinedTCR4$barcode, '_2', sep='')
mergedTCRlane3<-combinedTCR3[,c("barcode","CTgene")]
mergedTCRlane3<-mergedTCRlane3[!duplicated(mergedTCRlane3$barcode),]
rownames(mergedTCRlane3)<-mergedTCRlane3[,1]
mergedTCRlane3[,1]<-NULL
mergedTCRlane4<-combinedTCR4[,c("barcode","CTgene")]
mergedTCRlane4<-mergedTCRlane4[!duplicated(mergedTCRlane4$barcode),]
rownames(mergedTCRlane4)<-mergedTCRlane4[,1]
mergedTCRlane4[,1]<-NULL
sample.combined<-AddMetaData(sample.combined, mergedTCRlane3,col.name =
"laneCSPmergedTCR")
sample.combined<-AddMetaData(sample.combined, mergedTCRlane4,col.name =
"laneRASmergedTCR")
table5<-
table(sample.combined@meta.data[["laneCSPmergedTCR"]],sample.combined$h
ashtagClusterID)
table6<-
table(sample.combined@meta.data[["laneRASmergedTCR"]],sample.combined$h
ashtagClusterID)
View(table5)
View(table6)
write.csv(table5,file="laneCSPTCRcombined.csv")
write.csv(table6,file="laneRASTCRcombined.csv")
#Highlight cells that use TRAV10J31
Idents(sample_CSP1) <- sample_CSP1@meta.data[["laneCSPmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP1,
idents = c("TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-
1.None.TRBC2","TRAV10N.TRAJ31.TRAC_TRBV16.TRBJ2-
5.None.TRBC2","TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-
1.None.TRBC1","TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-
4.None.TRBC2","TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-
5.None.TRBC2","TRAV10.TRAJ31.TRAC_TRBV14.TRBJ2-
5.None.TRBC2","TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-
1.None.TRBC1","TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-
3.None.TRBC1","TRAV10.TRAJ31.TRAC_TRBV3.TRBJ2-3.None.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP1,
idents = c("TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-
1.None.TRBC2","TRAV10N.TRAJ31.TRAC_TRBV16.TRBJ2-
5.None.TRBC2","TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-
1.None.TRBC1","TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-
4.None.TRBC2","TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-
5.None.TRBC2","TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-
1.None.TRBC1","TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-3.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP1,
idents ="TRAV14D-1.TRAJ49.TRAC_TRBV5.TRBJ2-7.None.TRBC2" ))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP1,
idents ="TRAV5D-4.TRAJ7.TRAC_TRBV5.TRBJ1-1.None.TRBC1" ))
Idents(sample_CSP2) <- sample_CSP2@meta.data[["laneCSPmergedTCR"]]
```

```
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP2,
idents = c("TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2",
"TRAV10N.TRAJ31.TRAC_TRBV3.TRBJ2-5.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV12-2.TRBJ2-7.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV31.TRBJ2-5.TRBD2.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP2,
idents = c("TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP2,
idents ="TRAV5-4.TRAJ52.TRAC_TRBV3.TRBJ1-1.None.TRBC1" ))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP2,
idents ="TRAV4-3.TRAJ40.TRAC_TRBV3.TRBJ2-7.None.TRBC2" ))
Idents(sample_CSP3) <- sample_CSP3@meta.data[["laneCSPmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP3,
idents = c("TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",
"TRAV10N.TRAJ31.TRAC_TRBV3.TRBJ1-1.None.TRBC1",
"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV3.TRBJ2-1.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV3.TRBJ2-7.None.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP3,
idents = c("TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",

"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP3,
idents = "TRAV7-5.TRAJ40.TRAC_TRBV12-2+TRBV13-2.TRBJ1-6.None.TRBC1"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_CSP3,
idents = "TRAV19.TRAJ56.TRAC_TRBV15.TRBJ1-4.None.TRBC1"))
Idents(sample_RAS1) <- sample_RAS1@meta.data[["laneRASmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS1,
idents = c("TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",
"TRAV10N.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1",
"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-2.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-4.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV3.TRBJ1-1.None.TRBC1",
"TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS1,
idents = c("TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",

"TRAV10N.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1",

"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-2.None.TRBC2",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-4.None.TRBC2",


"TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1")))
```

```
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS1,
idents = "TRAV7-3.TRAJ12.TRAC_TRBV5.TRBJ1-5.None.TRBC1"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS1,
idents = "TRAV14-1.TRAJ40.TRAC_TRBV16.TRBJ2-7.None.TRBC2"))
Idents(sample_RAS2) <- sample_RAS2@meta.data[["laneRASmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS2,
idents =c( "TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",
"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV13-1.TRBJ2-4.None.TRBC2",
"TRAV10D.TRAJ31.TRAC_TRBV3.TRBJ2-4.None.TRBC2",
"TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-3.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS2,
idents =c( "TRAV10.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1",

"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2",



"TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-3.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS2,
idents = "TRAV7-5.TRAJ40.TRAC_TRBV5.TRBJ2-1.None.TRBC2"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS2,
idents = "TRAV4-2.TRAJ9.TRAC_TRBV5.TRBJ2-5.None.TRBC2"))
Idents(sample_RAS3) <- sample_RAS3@meta.data[["laneRASmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS3,
idents = c("TRAV10.TRAJ31.TRAC_TRBV19.TRBJ2-7.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2",
"TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS3,
idents = c(

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2",

"TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-4.None.TRBC2",

"TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2",

"TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS3,
idents = "TRAV10.TRAJ31.TRAC_TRBV26.TRBJ2-1.None.TRBC2"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS3,
idents = "TRAV4-3.TRAJ34.TRAC_TRBV15.TRBJ1-1.None.TRBC1"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_RAS3,
idents = "TRAV4D-4.TRAJ33.TRAC_TRBV12-2.TRBJ2-5.TRBD2.TRBC2"))
Idents(sample_no2a104) <-
sample_no2a104@meta.data[["laneCSPmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a104,
idents = c("TRAV10.TRAJ31.TRAC_TRBV16.TRBJ2-5.None.TRBC2")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a104,
idents = "TRAV4-4-DV10.TRAJ49.TRAC_TRBV16.TRBJ2-3.None.TRBC2"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a104,
idents = "TRAV4-4-DV10.TRAJ43.TRAC_TRBV19.TRBJ2-5.None.TRBC2"))
Idents(sample_no2a105) <-
sample_no2a105@meta.data[["laneCSPmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a105,
idents = c("TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1",
```

```
"TRAV10.TRAJ31.TRAC_TRBV20.TRBJ1-4.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a105,
idents = c("TRAV10D.TRAJ31.TRAC_TRBV26.TRBJ1-2.None.TRBC1"

)))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a105,
idents = "TRAV6-6.TRAJ17.TRAC_TRBV26.TRBJ1-3.None.TRBC1"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_no2a105,
idents = "TRAV9N-4.TRAJ33.TRAC_TRBV10.TRBJ2-7.None.TRBC2"))
Idents(sample_NT) <- sample_NT@meta.data[["laneRASmergedTCR"]]
DimPlot(sample.combined,cells.highlight = WhichCells(sample_NT, idents
= c("TRAV10.TRAJ31.TRAC_TRBV12-2+TRBV13-2.TRBJ2-3.None.TRBC2",
"TRAV10.TRAJ31.TRAC_TRBV20.TRBJ1-1.None.TRBC1",
"TRAV10D.TRAJ31.TRAC_TRBV10.TRBJ2-4.None.TRBC2",
"TRAV10N.TRAJ31.TRAC_TRBV13-1.TRBJ2-5.None.TRBC2",
"TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_NT, idents
= c("TRAV10N.TRAJ31.TRAC_TRBV26.TRBJ1-1.None.TRBC1")))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_NT5, idents
= "TRAV6-6.TRAJ17.TRAC_TRBV26.TRBJ1-3.None.TRBC1"))
DimPlot(sample.combined,cells.highlight = WhichCells(sample_NT5, idents
= "TRAV9N-4.TRAJ33.TRAC_TRBV10.TRBJ2-7.None.TRBC2"))
```

Code for Figure 5.4

```
#------TCR analysis

suppressMessages(library(scRepertoire))

library(readr)
library(readr)
SPZ1 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/SPZ1.csv",stringsAsFactors = F)
SPZ2 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/SPZ2.csv",stringsAsFactors = F)
SPZ3 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/SPZ3.csv",stringsAsFactors = F)
CSP1 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/CSP1.csv",stringsAsFactors = F)
CSP2 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/CSP2.csv",stringsAsFactors = F)
CSP3 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/CSP3.csv",stringsAsFactors = F)
NoB1 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/NoB_1.csv",stringsAsFactors = F)
NoB2 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/NoB_2.csv",stringsAsFactors = F)
NT1 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/NT1.csv",stringsAsFactors = F)
NT2 <- read.csv("E:/ANU Research/data/20200917-cellranger_outs/TCR
reptoire_Tfh/analyze1-all T cell for most
abundant/NT2.csv",stringsAsFactors = F)



contig_list <- list(SPZ1,SPZ2,SPZ3,CSP1,CSP2,CSP3,NoB1,NoB2,NT1,NT2)

combined <- combineTCR(contig_list, samples =
c("SPZ1","SPZ2","SPZ3","CSP1","CSP2","CSP3","NoB1","NoB2","NT1","NT2"),
                       ID =
c("SPZ","SPZ","SPZ","CSP","CSP","CSP","NoB","NoB","NT","NT"), cells
="T-AB",removeNA = T,
                       filterMulti = T)

quantContig(combined, cloneCall="gene+nt", scale = T)

quantContig_output <- quantContig(combined, cloneCall="gene+nt", scale
= T, exportTable = T)
quantContig_output

abundanceContig(combined, cloneCall = "aa", scale = F)
```

```
abundance_scaled<-abundanceContig(combined, cloneCall = "gene", scale =
T,exportTable = T)
abundance_nonscaled<-abundanceContig(combined, cloneCall = "gene",
scale = F,exportTable = T)
write.csv(abundance_nonscaled,file="abundancenonscale.csv")
write.csv(abundance_scaled,file="abundancescale.csv")


lengthContig(combined, cloneCall="aa", chains = "combined")

lengthContig(combined, cloneCall="nt", chains = "single")

compareClonotypes(combined, numbers=20, cloneCall="gene", graph =
"alluvial")

vizVgenes(combined, TCR="TCR1",  facet.y = "ID")
tcra_usage<-vizVgenes(combined, TCR="TCR1",  facet.y = "ID",exportTable
= T)
tcrb_usage<-vizVgenes(combined, TCR="TCR2",  facet.y = "ID",exportTable
= T)
write.csv(tcra_usage,file="tcra.csv")
write.csv(tcrb_usage,file="tcrb.csv")


clonalHomeostasis(combined, cloneCall = "gene")

clonalProportion(combined, cloneCall = "gene")

install.packages('Rcpp')
library(Rcpp)
clonalDiversity(combined, cloneCall = "gene", group = "ID")

clonalOverlap(combined, cloneCall = "gene", method = "morisita")
```